Pan-Canadian ABRACADABRA follow-up: What do we know four years later about students' and teachers' responsiveness to being part of an intervention study?

Sukhbinder (Bindy) Sanghera-Sidhu, MA

Department of Educational and Counselling Psychology

McGill University, Montreal

PhD in Educational Psychology

December 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Doctor of Philosophy in Educational Psychology, Human Development specialization

Abstract

This dissertation is a Canadian-based 4-year follow-up study that examines the long-term effectiveness of the ABRACADABRA (ABRA) web-based literacy intervention on students' (*n* = 467) reading progress and teachers' (*n* = 22) long-term use of a new intervention. This mixed-methods study is a quantitative study with a nested qualitative component at the teacher level of data analysis. This dissertation identifies factors influencing both students' and teachers' responses to being part of a randomized control trial (RCT) intervention study that examined the effectiveness of teacher-implemented ABRA lessons during classroom-level instruction. Framed within a response to an intervention (R*t*I) context, this study broadens the scope of the R*t*I literature from primarily focusing on pupil-level R*t*I variations to also considering the R*t*I effects on teachers.

At the pupil level, this study examines the enduring effectiveness of the ABRA intervention and investigates if the short-term reading gains obtained by students at immediate posttesting (T2), who received the ABRA intervention, were maintained up to 4 years later at follow-up (T3). The added contribution of demographic variables in predicting students' short- and long-term likelihood of being at risk of reading difficulties is also examined. A series of binary stepwise logistic regressions were run to examine the interaction and main effects of ABRA and the demographic variables on the variance of students' reading. An Ethnicity effect evident at T2 found students of Asian background having a raised risk of not responding to the intervention and remaining in the at risk of reading difficulties group in comparison to their White peers. A Sex effect in favour of female students was evident at T3. While a SES effect at both T2 and T3 showed that the odds of having stronger reading skills increased for students with mothers with some post-secondary education. When examining the students' long-term reading intervention response, no support was found for the inoculation hypothesis model, as the positive short-term reading gains made at T2 by the students identified at risk were not maintained at T3. No interaction effects between the ABRA intervention condition and the demographic variables of interest were found at T2 or at T3.

At the teacher level, a deductive thematic analysis (TA) approach is employed to examine factors influencing teachers' response to being part of an intervention study (R*t*I) and their subsequent long-term integration of a new resource into their teaching practice. At T3, over 70% of the teacher respondents reported that the ABRA program continued to be part of their literacy practice repertoire. A significant relationship was found between teachers' level of implementation (IFM) during the intervention phase and teachers' continued use of the ABRA tool. The findings from this study may have implications for how teachers are trained and supported during a classroom based intervention study, and how teachers can be included in the process to facilitate greater buy-in and improve their quality of implementation fidelity of new technology-based resources.

Keywords: reading intervention, follow-up study, longitudinal effects, randomized control trial, response-to-intervention, demographic variables, teacher change, technology integration, ABRACADABRA, implementation fidelity, mixed methods

Résumé

Cette thèse est une étude de suivi canadienne effectuée sur quatre ans examinant l’efficacité à long terme de l’outil Web ABRACADABRA (ABRA) sur les progrès en lecture des élèves (n = 467). Par ailleurs, cette étude examine également de quelle façon les enseignants utilisent une nouvelle ressource, à long terme, dans leur pratique (n = 22). Cette recherche à méthodologie mixte utilise principalement une approche quantitative mais comporte également une composante qualitative imbriquée au niveau de l’analyse de données des enseignants. En contexte de réponse à l’intervention (R*a*I), cette thèse identifie les facteurs influençant les réactions des élèves et des enseignants suite à leur participation à un essai randomisé contrôlé au sein d’une vaste étude d’intervention portant sur l’efficacité des leçons ABRA données par l’enseignant en classe (T1). Cette étude élargit la portée de la littérature scientifique portant sur le R*a*I, en mettant principalement l’accent sur les variations du R*a*I chez les élèves et en s’intéressant aux effets du R*a*I sur les enseignants.

En ce qui concerne les élèves, cette étude examine l’efficacité à long terme de l’intervention ABRA et tente de déterminer si les gains en lecture à court terme observés lors du posttest immédiat (T2) chez les élèves ayant reçu l’intervention ABRA ont été maintenus après quatre ans, lors du suivi (T3). Cette recherche examine également le rôle des variables démographiques dans le but de prédire la probabilité qu’ont les élèves d’éprouver des difficultés en lecture à court et à long termes. Une série de régressions logistiques binomiales séquentielles a été effectuée afin de vérifier les effets d’interaction et principaux d’ABRA et des variables démographiques sur la variance en lecture des élèves. Un effet principal d’ethnicité évident au T2 a montré que les élèves asiatiques présentaient un risque accru de ne pas répondre à l’intervention et de demeurer dans le groupe à risque d’éprouver des difficultés en lecture, comparativement aux élèves caucasiens. Au T3, un effet principal de genre favorable aux filles a également été observé alors qu'aux T2 et T3, un effet principal du statut socio-économique a montré que les chances d'avoir de meilleures compétences en lecture augmentaient chez les élèves dont la mère avait effectué des études post secondaires. En examinant la réponse à l'intervention en lecture des élèves, aucun appui n'a été trouvé en faveur du modèle de "l'inoculation hypothesis," puisque les gains en lecture à court terme enregistrés par les élèves à risque au T2 n'ont pas été maintenus au T3. Aucune interaction entre ABRA et les variables démographiques à l'étude n'a été observée au T2 ou au T3.

En ce qui a trait aux enseignants, une approche d'analyse thématique déductive a été utilisée afin d'étudier les facteurs influençant la réaction des enseignants à leur participation à une étude d'intervention. La même méthode a également été utilisée afin d'investiguer les facteurs affectant l'intégration subséquente, à long terme, d'une nouvelle ressource à leur pratique. Au T3, plus de 70% des enseignants ont mentionné qu'ABRA faisait toujours partie de leurs pratiques enseignantes en matière de littératie. Aussi, une relation significative a été observée entre le niveau d'implantation des enseignants pendant la phase d'intervention et leur utilisation continue de l'outil ABRA par la suite. Ces résultats pourraient avoir des conséquences sur la façon dont les enseignants sont formés et soutenus pendant une étude d'intervention en classe, ainsi que sur la manière dont ils peuvent être intégrés au processus visant à favoriser leur adhésion et la qualité de l'implantation de nouvelles ressources technologiques.

Mots-clés : intervention en lecture, étude de suivi, effets longitudinaux, essai randomisé contrôlé, réponse à l'intervention, variables démographiques, changement chez les enseignants, intégration de la technologie, ABRACADABRA, fidélité d'implantation, méthodes mixtes

Dedication

I dedicate this dissertation to my beloved family who shared in the sacrifices necessary to complete this degree.

To my husband, Amandeep, who has had a profound influence on my life since the day we met. You have supported me in this journey since the beginning (*by throwing me into it!*). Without you, this would not have been possible. You believed in me, way before I even believed in myself. I share this degree with you.

To Kartar, my spunky, intelligent and adorable daughter, I am so blessed to be your mother. Having you in my life did make this journey a bit more challenging -- but it was SO worth it. I had never imagined the depths of joy you would bring into my life. I am so proud to be your '*momma*' and now that this dissertation is finally completed I look forward to spending countless hours playing with you and watching you blossom. XO

I also dedicate this dissertation to my mother, who taught me to work hard and love completely. Your unconditional love, words of encouragement and countless hours logged-in as Kartar's caretaker over the last two years have been instrumental to the successful completion of this degree.

I am forever blessed to have all of you in my life. Love you all.
#83 Always, XO

Acknowledgements

'Little drops of water make the mighty ocean'

- Modification of Julia A. F. Carney's (1844) quote.

This dissertation would not have been possible without the help of so many people in so many ways. I begin by thanking the students who participated in this follow-up study, as well as the teachers, administrators and school staff for allowing me into their schools and assisting me in my data collection process. I am especially grateful to the teacher participants of this study who took time away from their busy schedules to share their experiences and stories.

In the past decade that I have been associated with McGill University, I have had the privilege to work with many amazing individuals that have enabled me to reach many academic and personal goals. This dissertation would not have been possible without the support and guidance of these talented individuals.

I begin by expressing my deepest gratitude to my supervisor Dr. Robert Savage for his unwavering support and guidance. Thank you for believing in me and taking me on as your PHD student. I would never have reached this point without your willingness to share your knowledge, wisdom and TIME! Most of all, thank you for your patience and understanding that each student has their own unique process of learning and writing (*And is my process ever unique*!). So many times I questioned if I would forever be your grad student. Seriously, I often thought I would never finish. But you always had faith in me. So thank you Robert, you will always have my deepest gratitude for *finally* getting me to the finish line and successfully obtaining this doctorate.

My sincere appreciation goes out to my Proposal Defence advisory committee members: Drs. Lynn Butler-Kisber, Jessica Ruglis, and Nathan Hall; as well as to my Oral Defence examiners: Drs. Tara Flanagan, Michael Hoover, Adam Dubé, Teresa Strong Wilson, and David Slomp. You are all leaders in your area of specialties, and I thank all of you for taking the time to share your insightful comments and constructive feedback during the initial and final stages of this dissertation. Dr. Ruglis, a very special added acknowledgment to you for your ongoing feedback and valuable insights on the qualitative components of this paper.

A special thank you to Carlos Gomez-Garibello for your invaluable statistical support, to Yasaman Jalai-Kushki, Liane Pereira, and Erika Dabner for sharing your qualitative knowledge and resources, and to Wayne Egers for all of your assistance with the never-ending edits.

I wish to thank all of the faculty and staff in the Department of Educational Counselling and Psychology at McGill University. Specifically, I am appreciative for the guidance and insight bestowed upon me during my graduate studies from the following professors, Dr. Nancy Heath, Dr. Victoria Talwar, Dr. Jake Burack, Dr. Krista Muis, and Dr. Tara Flanagan.

I would like to acknowledge my gratitude to Drs. Gia Deleveaux, Noella Piquette, and Eileen Wood for your guidance and invaluable support and treating me as your equal peer on the ABRA research projects. A special thanks is also extended to the staff at CSLP at Concordia University for all of your assistance during my time on the ABRA project.

My work would also not have been possible without my treasured Montreal friends who ensured my time away from my work was filled with fun and laughter, you are more than school colleagues, you are family. To Heidi Flores, Katie Cohene, Yasaman Jalali-Kushki, Maria Di Stasio, Hanady Bani Hani, Brigitte Weil, Ozlem Erten, Girlie Arcebal De Vera, Liane Pereira, Ramona Arora, Carlos Gomez-Garibello, Alex Nowak, Louise Deault, Adriana Pace, Marie-France Côté, Victoria Budd, Karla Jacobsen, and Vanessa Oude-Reimerink, you are all brilliant, talented and warm individuals whose friendships during my time in Montreal have truly enriched my life. Yogini my cherished SKYPE writing companion, while continents apart yet living parallel lives, you perhaps provided me with the most memorable quotes, from "*In my acknowledgements I'm* *going to write "first I would like to thank myself for tolerating and surviving the thesis writing process*" to "*I think I've become allergic to my thesis.*" *Drop by drop* we did it! No more *tomato-time!* To all of Montreal friends, you are an incredible group. Each of you will always have a special place in my heart for helping me keep my sanity while I wrote, and for being there for me as I began the process of welcoming a new life into this world.

Finally, special recognition goes out to my mother and father for their support, encouragement and patience during my pursuit of this degree. My gratitude is especially extended to my mother, who has always believed in me and supported me through all of my academic endeavours over these last four decades. Your confidence in me, even when I began to wonder and doubt myself, allowed me to achieve. Mom, I am forever indebted to you as all of my accomplishments are the result of your unconditional love and upbringing.

I feel extremely fortunate for having married into such a caring and supportive family, as all of the Sidhus have wholeheartedly supported my endeavors. Dad, Pawan, and especially mom, I sincerely appreciate all of your help taking care of Kartar. Thank you for allowing me to spend time with my studies, as your understanding, patience and providing me the space to write were key variables in the successful completion of this dissertation. Thank you also for helping Kartar build her strong walking legs by taking her out for so many walks so that I would have some quiet time to write. I would also like to express my deepest gratitude to Nancy and Parm. Thank you so much for the unlimited use of chez Serown! Your home became the perfect place to escape to in order to get those final weeks *(…um months)* of writing accomplished. Your incredible hospitality (i.e., midnight brownies and endless supply of java), unwavering positive support and genuine kindness will never be forgotten. To my extended family, I will always be thankful to each of you for the roles you played in the completion of this dissertation.

To my husband, Amandeep, I would like to express my gratitude to you for your love and care. Aman, pursuing this doctorate simply would not have been possible without your understanding, and logistical support. Your belief in my abilities provided me the strength and perseverance necessary to achieve this tremendous goal, and for that I am eternally grateful. Now let us get on with our lives.

To my little girl, Kartar, who over the last year missed out on a lot of Mommy time during my pursuit of this degree. Kartar, above any other title I could earn, I am most proud of the title you gave me, Mother. *"Just one hug Nanni-ji…I want to give momma just one hug!...Momma go work hard…Auja Bindy-yaa"* – words that echoed in my thoughts and warmed my heart when I was away from you. Thank you for the smiles, laughs, hugs, kisses and special dances that always put my life into perspective. You inspire me to be the best possible self I can be.

GRANT INFORMATION

This degree was completed in part with the financial support of a research fellowship provided by Fonds Québécois de la recherche sur la Société et la Culture (FQRSC).

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Purpose of the Study

In educational research, the randomised controlled trial (RCT) is widely accepted as the *gold standard* of effectiveness research (Haynes, Service, Goldacre, & Torgerson, 2012; Torgerson, 2003). Systematically consolidated evidence from well-designed RCT studies has increased our understanding of children's reading acquisition abilities and also provided valuable support regarding early literacy practices considered effective for fostering children's reading development (e.g., National Early Literacy Panel, 2008; Torgerson, 2007). However, results from classic RCT experiments do not answer all questions about the effectiveness of reading interventions. Findings from RCT studies often neglect to address individual variations in responses to interventions considered empirically sound (Al Otaiba & Fuchs, 2006). Thus, adding a response-to-intervention (R*t*I) component of analysis to sound RCT studies enables researchers to extend examinations of effectiveness from is the reading intervention practice effective or not to for whom is the reading intervention effective.

Over the last two decades, the interest, research, and body of literature concerning students' responses to reading interventions has grown (e.g., Foorman & Wanzek, 2016; Lam & McMaster, 2014). The concept of R*t*I has been formalized into stages of early intervention action and also into policy in a number of school districts (McIntosh et al., 2011; Preston, Wood, & Stecker, 2016). In classic R*t*I models, graded *tiers* of early intervention and ongoing assessments are offered. The first tier is assumed to provide quality evidence-based teaching at the whole class level. Students identified as falling behind, or not responding to effective (Tier 1) classroom teaching, are provided with additional support in subsequent tiers of intervention. These subsequent tiered interventions are commonly delivered as intensive small group teaching support (Tier 2), followed by more individualized one-to-one tuition (Tier 3) support for students who continue to demonstrate difficulties (Denton, 2012).

Currently, most R*t*I studies primarily explore the effectiveness of Tier 2 or 3 levels of intervention, with examinations of the implementation and quality of Tier 1 interventions almost entirely neglected (Hill, King, Lemons, & Partanen, 2012). This neglect of detailed examinations of the fidelity of Tier 1 interventions suggests that additional research is needed to better understand the quality of instruction delivered at the whole class level. In addition, more documented studies are needed to examine factors influencing students' responses to whole classroom (Tier 1) instruction deemed to be effective.

Much of the current focus in reading research has examined the degree to which students identified as at risk for reading problems have benefited from their participation in early literacy interventions. This focus in R*t*I is understandable, since the students are the ones receiving the reading intervention; however, rarely acknowledged is the fact that teachers also are responding to the intervention, since they are responding to being part of an intervention study.

Studies have been conducted to investigate predictive factors of *teacher change* in teachers practice (e.g., Clarke & Hollingsworth, 2002; Fullan, 2005; Guskey, 1986), although without an examination of this "change" within a R*t*I study context, which is surprising, since the teacher is one of the key participants in an intervention study. In other words, teachers needs to change their practice to effectively implement the proposed intervention, which is an indication that they are responding (in some way) to the intervention they are implementing. Yet, relatively little is known about how teachers respond (short- or long-term) to being part of an intervention study. This study aims to contribute to the literature by broadening the scope of the R*t*I literature from primarily focusing on pupil-level R*t*I variations to also considering the R*t*I effects on teachers.

This dissertation explores the long-term R*t*I effectiveness of the ABRACADABRA (ABRA) free-access web-based literacy intervention. ABRA as a literacy tool was conceptualized and constructed based on best evidence findings on effective reading interventions for phonics and letter skills, reading fluency, and reading comprehension (Savage, Abrami, Hipps, et al., 2009). By applying an R*t*I analysis approach, both pupil-level and teacher-level within-group variation responses to a Tier 1 intervention are examined. At the teacher level, an analysis is conducted to identify potential mediating factors that may facilitate a teacher's long-term use of an evidence-based program deemed to be effective. At the pupil-level, this follow-up R*t*I study examines the enduring effectiveness of the ABRA intervention and investigates if the short-term reading gains obtained by students, who received the ABRA intervention during the 2007–2009 ABRA RCT study, were maintained up to 4 years later (see Savage et al., 2013).

By conducting a follow-up analysis on the Pan Canadian ABRA early literacy intervention, this follow-up study builds on the existing research conducted on the use of ABRA as an effective early literacy Tier 1 intervention tool (e.g., Abrami, Borohkovski, & Lysenko, 2015; Comaskey, Savage, & Abrami, 2009; Piquette, Savage, & Abrami, 2014; Savage, Abrami, Hipps, & Deault, 2009; Wolgemuth et al., 2013). To the best of this researcher's knowledge, no one in Canada has looked at a large-scale technology-based intervention, and then analyzed that data in detail and followed up to examine the long-term effects on both the teachers' use of the technology and the students' progress. In addition, this present study focuses on the non-cognitive variables of the pupil-level characteristics, such as male or female status (sex-differences), socio-economic status (SES), and ethnicity. As this study looks at both interaction and main effects of ABRA and the demographic variables of interest on students' reading. This analysis adds to the R*t*I literature, since traditionally at the pupil level, cognitive variables have been the focus of R*t*I studies. Finally, this study is not only an empirical investigation of the questions concerning the long-term effects of responses to a classroom level intervention, but also is a testing of the theoretical positions that explain pupil and teacher responses to this intervention. In this examination of Tier 1 R*t*I at the pupil level, two hypotheses put forward by Coyne, Kame'enui, Simmons and Harn (2004) are tested—the *inoculation hypothesis* and the *insulin hypothesis*—with respect to the enduring effects of the ABRA intervention on the reading progress of students at risk of reading difficulties. The models of teacher change and professional growth presented in Clarke and Hollingsworth's (2002) article "Elaborating a Model of Teacher Professional Growth" are employed to examine whether teachers continued to use a new innovation (the ABRA literacy program) after their involvement in the study ended.

Thesis Synopsis

This thesis is presented in nine chapters. Chapter 1 provides the rationale for this study and identifies the theoretical models used to examine the pupil and teacher-level data. Chapter 2 provides a comprehensive review of the related literature. The focus is on how literacy research studies have contributed to our present understanding of effective reading strategies, and discusses the areas where further research is still required. Since this present study is a follow-up on an existing random control trial (RCT) intervention study, the two primary goals of the second chapter are: 1) to provide background about why the original RCT study was needed and (2) to set-up the grounds for how the current study adds value to the earlier findings of the 2007–2009 RCT literacy intervention study by providing a R*t*I framework of analysis.

Chapter 3 explores the value of R*t*I research in the field of reading research and classroom practice. This chapter reviews what has traditionally been the focus of R*t*I literature at the pupil level (i.e., examination of cognitive variables) and addresses the need for researchers to extend their analysis to include examinations of sociological factors that may influence a child's R*t*I. Also introduced are the pupil-level theoretical models (the inoculation and insulin hypotheses).

Chapter 4 details one of the original contributions of this dissertation, which is taking traditional teacher change models and applying them to an R*t*I framework of analysis. Four domains of influence on teacher change are identified (external domain, personal domain, domain of practice, and domain of consequence), and three conceptual teacher change models are introduced. Although each of these models is made up of the same four domains of teacher change, each model reflects a different trajectory of teacher learning across these four domains. One objective of this chapter is to review and identify potential variables that may best support teacher learning and their likelihood of long-term change.

Chapter 5 begins by providing a summary of the original 2007–2009 Pan-Canadian Cluster RCT ABRACADABRA study. Then the chapter outlines the purpose of the present follow-up study, the research objectives, and potential research questions that guided the study.

Chapter 6 provides a description of the research design and methods used in this response-to-intervention follow-up study. It includes CONSORT diagrams that illustrate the flow of the student and teacher samples from the original study to follow-up, a description of the data collection timeline, the materials and methods used to collect the data, and a restatement of the key research questions explored in the present thesis.

Since two distinct levels of data are used in the present study, the data analysis and results section for each level of data was reported in two separate chapters. So, Chapter 7 reports the data findings of the student level analysis, while Chapter 8 reports on the key teacher-level findings. The discussion of findings, conclusions, and recommendations for both the student level and teacher level are presented in Chapter 9.

Chapter 2: Review of Early Literacy Effectiveness Research

Overview

The purpose of this chapter is to review selected literature that outlines the need for this study. This chapter draws attention to the fact that despite the large volume of current literature on effective reading practices, currently, very little high-quality evidence of what interventions work and for whom is available, especially in a Canadian context.

In addition, this chapter's review of effective reading practices illustrates not only the value of well-designed random control trial (RCT) intervention studies for demonstrating group differences, but also the limitations of RCT studies in not identifying individual differences, which establishes the need for a response-to-intervention analysis of sound RCT studies.

Scientifically Based Evidence on Effectiveness Research

Interest is growing in scientific studies of reading processes, instruction, and reading intervention practices as educators and practitioners continue to seek out scientifically-based evidence to guide their literacy programs (Fletcher & Wagner, 2014; Savage & Cloutier, 2016). However, despite a large body of literature on reading (see, for example, National Early Literacy Panel, 2008; National Institute of Child Health and Human Development (NICHD), 2006; National Reading Panel, 2000), Savage (2012) has argued that high quality evidence is still lacking to support the most recommended literacy intervention practices. This deficiency is of concern because when questions about the effectiveness of literacy interventions are raised, it is important that policy, practice, and future research are informed by sound scientifically-based evidence that can best establish genuinely causal links (Savage & Cloutier, 2016).

The wider research community emphasize that whenever possible, evidence from quality randomised control trials (RCT) should be used, and that technology-based tools should be designed and evaluated on the basis of the findings of systematic reviews of RCT studies (National Reading Panel, 2000; Savage, 2008). However, Savage stresses that the RCT studies used within systematic reviews need to be high-quality, well-designed experimental studies that adhere to research quality guidelines, such as those set out in the Consolidated Standards of Reporting Trials (CONSORT) statement (see http://www.consort-statement.org) (Savage, 2008, 2012). If the evidence used to identify effective reading intervention strategies is only as reliable as the quality of the research studies on which they are based, the consolidated evidence from sound RCTs that meet established guidelines will be of higher quality.

Randomised Control Trials (RCT)

With respect to a hierarchical model of evidence reliability, a well-designed RCT study is regarded as the *gold standard* for establishing intervention effectiveness (Haynes et al., 2012; Shelley, Yore, & Hand, 2009). In scientific research, classic experimental designs are used to test causal relationships between variables. The random allocation of participants to either the treatment group or the control group helps to ensure that participants are comparable across the intervention conditions and increases the internal validity of a study by eliminating selection bias (Shelley et al., 2009; Tabachnick & Fidell, 2013). Torgerson and Torgerson (2008) have suggested that if random assignment is used on a sufficient sample size (e.g., $n = 50$ participants per cell), any pre-existing differences between the participants of the treatment and non-treatment groups should even out, since this process best controls for potentially confounding factors (extraneous variables) that may bias the results of an intervention. Randomization with a large sample thus provides an 'unbiased estimate of error' in interpreting IV-DV casual links; as such, quality RCT studies have better control over the effects of extraneous variables that might subsequently lead to false positive or false negative results (Tabachnick & Fidell, 2013).

For example, Savage et al. (2013) ABRACADABRA cluster RCT intervention study addressed the issues associated with intervention effectiveness by paying careful attention to treatment fidelity, implementation integrity, and the control of extraneous variables. Efforts to increase the internal validity of a study, along with a large-enough sample, make it more likely that the effects reported by high-quality RCTs are due to the intervention, rather than chance or other uncontrolled extraneous factors, and thus the findings have a greater likelihood of being repeated and generalized to a wider population. Nonetheless, as with any research methodology, RCT designs have some potential limitations of which readers and researchers need to be aware.

Potential limitations of RCT studies. The science of clinical trials is well established with the CONSORT protocols listing the main elements of design that are needed (see Savage, 2012); however, general pitfalls exist that could produce a badly conducted RCT. For example, Stein et al. (2008) have argued that not all RCTs take the necessary measures to observe and control for the appropriate implementation of the treatment condition. In addition, McIntosh et al. (2011) have pointed out that assumptions should never be made that an evidence-based resource will be implemented with fidelity, even though this resource has been purchased and personnel have been trained to use it. Without the use of treatment implementation fidelity measures, assurance does not exist that the intervention was implemented as intended, and thus accurate conclusions about the effectiveness of intervention practices cannot be drawn (Hill et al., 2012; Kurz, Elliott, & Roach, 2015). Savage (2008) has suggested that even though the RCT design is considered the most robust, poorly designed RCTs or RCTs with small samples may produce less reliable findings.

Another limitation that Jadad and Enkin (2007) have addressed is the ethical dimension. These researchers have highlighted that a complete random allocation, by which an effective treatment is allocated to one group and withheld from another (albeit temporarily), is not always reasonable to do in the real world. So, to avoid such an ethical dilemma, reading researchers at the conclusion of a study may offer the control group an opportunity to access the intervention (for example as offered in Savage et al., 2010). In addition to the ethical issue, Savage (2012) has acknowledged that practical issues—funding, cost, and commitment from all parties involved in running a sound, large-scale RCT study within the classroom environment—also are significant.

Al Otaiba and Fuchs (2006) have identified another problem, perhaps the most critical limitation of the classic RCT experiment design. These authors stress that the findings of even sound RCT studies are not always able to answer all the questions concerning the effectiveness of an intervention. If done well, an RCT can be used to show that an effect occurred (i.e., that A caused B), but the findings of an RCT cannot tell the reader anything about how the intervention works or for whom the intervention was and was not effective. This critique suggests that after an intervention is found to be effective, researchers must continue their investigation beyond the limited scope of the basic RCT design.

Suggestions for avoiding potential pitfalls of RCT research. To avoid some of the potential problems that could arise when implementing an RCT, many journals now require researchers to adhere to the CONSORT guidelines. The CONSORT statement includes a checklist of criteria and detailed guidelines on all aspects of designing and reporting RCTs (Altman et al., 2001; Campbell, Elbourne, & Altman, 2004). Examples of the recommendations made by the CONSORT group include the provision of detailed information of key pupil-level characteristics, explanations of the allocation process of participants (i.e., clustered by classrooms), the inclusion of a flow diagram to illustrate the random allocation process of participants, and the use of appropriate statistical methods to compare groups for analyses—all of these are examples of the quality-criteria needed to ensure an appropriate interpretation of RCT results. Similarly, several researchers (e.g., Anderson, Wood, Piquette-Tomei, Savage, & Mueller, 2011; Lane, Bocian, MacMillan, & Gresham, 2004; Stein et al., 2008) have argued that all classroom-based RCT intervention studies should include treatment integrity measures.

The potential limitations of RCT studies illustrate that even though evidence of practice effectiveness from RCTs is preferred, not all RCTs are a gold standard. Savage (2012) has suggested that in education research, a need exists for more published RCT studies that are rigorously sound and meet the CONSORT criteria. Savage (2012) also found that only a limited number of RCT-only meta-analysis are available that have examined the effectiveness of literacy interventions; these findings are consistent with Seethaler and Fuchs (2005) who reported that quality RCT studies of reading and mathematics represent only a drop in the bucket, since they make up less than 5% of the published papers in leading education and psychology journals; thus, the number of reading researchers conducting quality RCT studies is currently limited. Further, classic experiment designs, even RCT studies, do not answer all the questions about the effectiveness of an intervention. As Al Otaiba and Fuchs (2006) point out, such designs are effective in demonstrating group differences but not in identifying individual differences in pupil-level responses to intervention practices deemed to be effective.

The next section explores the literature on reading intervention practices so to examine the quality of evidence support for the effectiveness of commonly recommended early literacy practices, and to address the methodological importance of follow-up studies.

Value of Early Literacy Intervention Research

The Canadian Council on Learning (2010) has reported that almost half of all Canadian adults (48%) over the age of 16 do not possess the literacy skills required to fully participate in modern society. Reading researchers contend that if by middle school, children with reading difficulties do not receive the preventive or remedial support needed, the majority of them may continue to have reading problems to the end of high school, and many may be at a greater risk of grade repetition, delinquency, and ultimately dropping out of school (Slavin, Lake, Davis, & Madden, 2011). The negative effect of early reading failure on children's educational, social, and affective development also has been well documented (see Cunningham & Stanovich, 1997; Foorman & Wanzek, 2016; Foster & Miller, 2007; Jamieson, 2006; Slavin et al., 2011). For example, Jamieson (2006) has reported that the negative impact of poor literacy development extends beyond success in school, since strong literacy skills are key determinants of future social, educational, and economic success. Jamieson argues that on average, each additional year of education that a person receives corresponds to an increase in earnings by approximately 8% a year, and that high school dropouts earn close to 50% less than those with high school diplomas. However, the causal link from early literacy interventions to later well being has not yet been proven so there is a need to be cautious. However, given the potential cumulative long-term costs of illiteracy, efforts to implement early literacy intervention practices are potentially of great value.

Researchers have found that students at risk of reading difficulties may be helped by early intervention (Lam & McMaster, 2014; Torgesen, 2000). For example, Torgesen's (2000) review found that early identification and literacy intervention by grade 1 may significantly reduce the number of children who might otherwise be eligible for special education services. Torgesen found that after exposure to a reading intervention program, the expected incidence of children identified as at risk for reading disabilities was reduced from between 12 and 18% to between 1.4 and 5.4%.

Linan-Thompson, Vaughn, Prater, and Cirino (2006) found that the majority of the English language learners (ELL) at risk of reading difficulties in their study who received 50 minutes of supplemental reading intervention daily in grade 1 were able to improve their performance on reading-related measures, with a larger percentage of the intervention students maintaining their responder status to the end of the second grade in comparison to the control group.

In another example, Vellutino, Scanlon, Small, and Fanuele (2006) conducted a 5-year longitudinal study of two kindergarten cohorts with an initial sample of students (*n* = 1,373). Children identified as at risk for early reading difficulties were randomly assigned to receive researcher-led small group intervention two to three times a week for the duration of kindergarten. In the first grade, children were reassessed, and those who still had difficulties with reading from the kindergarten project treatment group (approximately 50%) received one-to-one daily tutoring provided by research teachers, while the children from the kindergarten school-based comparison groups who were still at risk and did not participate in any school-based intervention program (approximately 80%) continued on as the comparison group until the end of grade three. The findings of Vellutino et al. (2006) suggest that early intervention in kindergarten alone or combined with first-grade intervention can help prevent early and long-term reading difficulties for most at-risk children, since the majority of children who received some form of intervention generally outperformed the children in the comparison group on measures of early literacy skills. These authors also have claimed that their findings support the argument that early and long-term reading difficulties in most children, but not all, are primarily the result of experiential and instructional deficits rather than cognitive deficits that are biologically based. Similarly, Griffiths and Stuart (2013) in their review of characteristics of effective early intervention programmes add that some older students (aged 7-12) may be instructional casualties of the education system as they may not have received quality classroom-level reading instruction earlier on.

Some notable limitations exist, specifically since the findings of both of these studies (Linan-Thompson, et al., 2006; Vellutino, et al., 2006) were the result of expensive and researcher-led interventions that were implemented with fidelity. Some studies have shown that teacher-delivered intervention studies have generally produced result findings with lower effect sizes than researcher-delivered intervention studies (Archer et al., 2014; Kim, Linan- Thompson, & Misquitta, 2012; Suggate, 2014). This finding trend needs to be addressed, especially if the ultimate aim of testing the value of a reading intervention is to see how effective the intervention tool can be in real-world conditions in the hands of the educators themselves. Moreover, Slavin et al. (2011) have suggested that the maximally effective early literacy interventions are those that aim to improve the quality of regular classroom instruction, such as training in cooperative learning, phonics, and phonemic awareness instruction programs. Concluding that research focus on enhancing the quality of classroom instructional programs and improving the effectiveness of teacher-led practices continues to be of value.

Additional research on the long-term effectiveness of early intervention practices is still required, since many studies similar to the one conducted by Linan-Thompson et al. (2006) may find initial significant effects on early reading skill measures, yet without a follow-up study 3 to 5 years later, it is uncertain if these initial effects were maintained and continued to influence higher level reading skills (Savage & Cloutier, 2016; Suggate, 2014). For example, Hurry and Sylva (2007)—when examining the long-term effectiveness of two models of early literacy intervention, Reading Recovery and a phonological training program—found evidence that both intervention programs significantly improved aspects of the children’s reading in both the short- (immediately post-intervention) and medium-term (1 year later). However, in the longer-term, 3½ years post-intervention, the authors found no significant effects on overall reading measures, which led them to conclude that early intervention alone may not be sufficient to prevent later reading problems and that some children at risk of reading difficulties may require ongoing support, an indicator of the need for further investigation of the long-term effectiveness of interventions.

A recent meta-analysis by Suggate (2014) has provided a rare review of the long-term effectiveness of reading interventions. Suggate addresses the current lack of accumulated research examining the longer-term effects of reading interventions by reporting on experimental and quasi-experimental studies that also measured the impact of intervention programs at delayed post-tests. Overall, Suggate found a good maintenance of effects for comprehension and phonemic awareness interventions, whereas the effects of phonics and fluency interventions diminished at follow-up. Suggate also identified a number of methodological features positively affecting effect size outcomes—including strong treatment integrity, the use of RCT versus non-RCT experimental designs, and dosage referring to how reading interventions were administered—with interventions supplementing existing reading programs faring better than those that replaced existing LA programs already in place.

However, it should be noted that the studies included in Suggate's review were based on relatively short follow-up times (Savage & Cloutier, 2016). The follow-up testing occurred anywhere from 3 to 48 months ($M = 11.17$) immediately following post-testing, with only one study testing over 2 years. Although a 10-year follow-up long-term study by Blachman et al. (2014) was considered for the review, Suggate excluded this study as being untypically long in its follow-up duration. Savage and Cloutier (2016) contend that based on the relatively short follow-up times of the studies in the Suggate review, a need still exists for more long-term studies to genuinely address the long-term effects of sustained intervention on later reading outcomes.

For decades researchers have claimed that the remediation of reading difficulties becomes progressively more difficult after the third grade, and therefore, the earlier that children at risk of reading difficulties receive additional literacy support, the more effective it may be (Juel, 1988; Slavin et al., 2011; Torgesen, 1998). For example, Torgesen's (1998) paper "Catch Them Before They Fall" has argued that the best solution to counter the problem of reading failure is to allocate resources for early identification and prevention because by third grade, successful remediation is both more difficult and costly. Despite educators and policy makers knowing the value of early intervention, some researchers argue that many children at risk of reading difficulties do not receive additional intervention support until third grade when most are officially identified as having a learning disability (Fuchs & Vaughn, 2012; Slavin et al., 2011; Torgesen, 1998). However, this situation is now changing, since many school districts have adopted the use of a response-to-intervention (R*t*I) practice. R*t*I refers to a policy of providing a range of supports, from universal balance literacy practices for all students within a classroom to more individualized intervention instruction for students who demonstrate a greater need for acquiring literacy fundamentals (Tran, Sanchez, Arellano, & Lee Swanson, 2011). More details about the practice of R*t*I are provided in Chapter 3.

Considering the value of early intervention, it also is important to explore the scientific evidence supporting the efficacy of early literacy practices commonly suggested to be effective with children in the early stages of reading development. Arguably, intervention research needs to be grounded in studies of effective early literacy teaching practices so to best address the complex and multidimensional nature of reading development. It is worth noting that systematic evidence continues to be mixed and often limited with respect to literacy intervention practices deemed to be effective, and even more so when focussing on computer-based literacy intervention practices (Savage, 2008). The following section provides an overview of the key findings consolidated by reviewers reporting on the effectiveness of technology-based early literacy research.

Computer-based ICT Literacy Intervention Practices

Since the introduction of technology into the classroom, great speculation has arisen that students' literacy instruction could be enhanced by computer-based information and communications technology (ICT) (Bereiter, 2002; Savage, Abrami, Hipps, et al., 2009; Scardamalia & Bereiter, 1996). In beginning reading instruction, computers are increasingly being used for remediation and skill practice. However, despite the increasing use of computers, the present research literature has provided a mixed picture of the efficacy evidence for ICT literacy interventions. For example, in their narrative review of quasi-experimental studies, MacArthur, Ferretti, Okolo, and Cavalier (2001) argued that technology can be effective, whereas other reviewers such as Blok, Oostdam, Otter, and Overmatt (2002) have suggested that even though computer-assisted instruction generally tends to have a positive effect on beginning readers, the overall effect is small and has been derived from many studies without a control comparison group, and thus the findings should be received with caution. Similarly, Kulik (2003) in his review of 27 controlled evaluation studies on instructional technology and reading found that the studies produced mixed outcomes. With the exception of positive effects being found for one program that integrated computer instruction with non-computer activities, Kulik found that most of the studies reviewed showed few effects on reading performance, and thus he has argued that much remains to be learned about the effectiveness of technology in the classroom before concluding that computers can help improve children's beginning reading performance.

The results of more recent systematic reviews continue to provide mixed support for the degree of added value of technology for literacy practices. For example, Van Daal and Sandvik (2013) in their meta-analytic review of 35 studies examined the effectiveness of ICT intervention on the early literacy outcomes of children's literacy achievement. Their analysis found strong support for the use of multimedia interventions with children at risk, reporting medium to large positive effect sizes on literacy outcomes, including phonological awareness, concepts of print, comprehension, and non-word reading; whereas Takacs, Swart, and Bus (2015) in their recent meta-analysis review of 43 studies (a mix of experimental and quasi-experimental design studies) found only small positive effects of the value added by using technology-enhanced stories with young children in comparison to children simply listening to stories during adult shared book reading. Takacs et al. (2015) review also found that for at-risk children coming from less stimulated home literacy environments the type of ICT features had differing effects on the student outcomes; with multimedia features producing stronger positive effects and ICT interventions with interactive features being more detrimental to the at-risk children's literacy outcomes. The findings of these two reviews highlight that the type and use of ICT tools in schools varies and subsequently the effects on students' literacy and reading skills are likely to be different; especially as the studies within the reviews examine the effectiveness of different ICT tools being used in different ways, at different stages of development by different people.

Another reason to be cautiously optimistic of the use of ICT is that in spite of the widespread use of technology in schools (Cuban, 2001; Ertmer & Ottenbreit-Leftwich, 2010; Musti-Rao, Cartledge, Bennett, & Council, 2015), little firm evidence exists in high-quality randomized control trials (RCT) (Savage, 2012; Savage & Pompey, 2008) for the efficacy of educational technology. For example, at the time of the NRP report, only 21 studies (a mix of both quasi-experimental and RCTs) met the Panel's research methodology criteria (NRP, 2000). The Panel reported that several computer applications—such as the addition of hypertext and word processing functions for writing, and the use of speech to on-screen text technology in reading instruction—showed promise for the teaching of literacy. However, very few specific instructional applications of evidence-based practices for incorporating technology interventions into reading instruction practices could be drawn from their analysis. Savage (2012) has stressed that when looking at ICT literacy interventions, the goal for reading researchers should be to create tools that incorporate as many of the evidence-based practices for reading instruction that were outlined earlier in this chapter.

Since the publication of the NRP report, subsequent comprehensive systematic reviews have been carried out to examine the efficacy evidence for ICT intervention practices on the development of reading skills (see Blok et al., 2002; Ehri et al., 2001; Slavin, Lake, Chambers, Cheung, & Davis, 2009; Torgerson & Zhu, 2003). However, due to the lack of enough quality RCTs for a thorough examination of the literature, researchers such as Blok et al. (2002), Ehri et al. (2001), and more recently Slavin et al. (2009) have continued to include quasi-experiments or matched designs in their reviews. This situation suggests that even though both sets of reviewers reported small effect sizes for ICT interventions on literacy outcomes, some caution should be attached to the claims about established causal connections.

As emphasized earlier, the greatest confidence in the success of an intervention is if the evidence for this success comes from systematic review findings of sound RCT studies (Savage, 2012; Torgerson, Brooks, & Hall, 2006). Therefore, since Torgerson and Zhu's (2003) review adheres to this gold standard principle of effectiveness evidence, their review is considered methodologically more stringent. In their analysis, they examined only the limited number of well-designed RCTs available that studied the efficacy of ICT-based interventions (n = 12 studies selected from an initial 2,319). These authors found some positive effects of technology (e.g., that the use of word processing could effectively benefit weaker writers in improving the quality of their writing), yet overall, they reported a statistically non-significant effect size for the effectiveness of technology as an aid to literacy acquisition. Based on their review of the literature, Torgerson and Zhu (2003) concluded that the need for stronger evidence from well-designed RCT studies continues to exist, and thus policy makers and practitioners should wait for this evidence before embracing the use of technology to aid in literacy acquisition in the classroom. Savage and Pompey's (2008) systematic meta-analysis review of all RCT design studies on reading available in English also concluded that further research on the effectiveness of ICT on reading practices is needed, since in current well-designed RCT trials, they too found very little firm evidence for the use of educational technology.

Slavin et al. (2008) were only able to locate eight quality intervention studies for their systematic review that examined the effects of technology use on the learning of middle and high school students. They too found little support (ES = +0.10) for the claim that computer-based interventions could benefit middle school students' literacy skills development. Moreover, reviews by Slavin and colleagues (see Slavin et al., 2009; Slavin et al., 2011) found few positive effects of computer-assisted instruction programs for elementary children who were struggling to learn to read. For example, Slavin et al. (2011), across the 14 studies they examined ($n = 5$ were RCTs), calculated a weighted means effect size of +0.09, which suggests that technology has minimal impacts on the reading achievement of students at risk of reading difficulties. A more recent review by Cheung and Slavin (2012) also found only small positive effects ($d = +.16$) of educational technology applications on student reading outcomes in comparison to traditional non-technology-based teaching methods.

However, the pessimistic findings of these systematic reviews on the effectiveness of ICT interventions could be attributed to an overarching methodological issue that was identified as potentially troublesome in Savage et al. (2013) literature review of ICT research. The issue of concern has to do with the quality of the implementation of ICT reading-interventions. Evidence from both technology-based reading intervention studies (Chambers et al., 2008; Savage et al., 2010) and non-technology-based reading interventions (Davidson, Fields, & Yang, 2009; Stein et al., 2008) has showed that the variation in program implementation by teachers was the greatest factor influencing the degree of effectiveness of ICT interventions on children's reading measure outcomes. A meta-analysis by Abrami et al. (2015) summarizing the research on the effectiveness of ABRA also found that the larger ABRA effects on students reading competencies were attributable to high-quality implementation of the program. In addition, a tertiary meta-analysis review by Archer et al. (2014) found a significant increase in the overall effectiveness of ICTs when training and support was entered as a moderator variable (an increase from ES = 0.18 to ES = 0.57). This tertiary meta-analysis review highlighted the value of including implementation fidelity factors when examining the relative effectiveness of ICT interventions and technology use in the classroom.

Savage et al. (2013) have argued that even though it may be time consuming and expensive, all ICT intervention studies need to include a rigorously thorough examination of how teachers are implementing (and perhaps not implementing) the reading interventions under investigation. These authors have contended that without a detailed exploration and report of the fidelity of implementation, a guarantee does not exist that the intervention under investigation was implemented as it was intended, and therefore, the findings reported may not be a true reflection of the effectiveness of ICT reading-intervention.

In addition, Savage et al. (2013) have argued that a need continues to exist for quality training on evidence-based ICT programs that can be used to teach a variety of key reading skills in a developmentally appropriate manner, since these authors found that the depth of training was an important feature in the size of effects reported. Subsequently, Savage et al. (2013) have to date published a number of experimental studies (mainly RCT) using the web-based reading intervention ABRACADABRA (ABRA) (see Abrami et al., 2015 and Piquette et al., 2014 for recent reviews of ABRA studies). Most of these studies have shown that in well-implemented intervention conditions, the use of ABRA to supplement existing reading programs can produce medium effect sizes on a range of reading outcome measures, thus providing a more optimistic view of what technology can offer literacy (Savage & Cloutier, 2016).

Taken together, the present research findings suggest that although many individual experimental studies have demonstrated that ICT interventions can have positive effects on children's reading outcomes (e.g., Macaruso, Hook, & McCabe, 2006; Piquette et al., 2014; Savage, Abrami, Hipps, et al., 2009; Wolgemuth et al., 2011), currently, very little high quality evidence is available from systematic reviews suggesting any clear advantage for using technology in reading programs. It is not surprising then that Torgerson and colleagues (Torgerson, 2007; Torgerson & Zhu, 2003) caution teachers not to fully embrace the use of technology to teach literacy skills until several new well-designed RCT studies showing the positive effects of technology on students' reading outcomes are published and evaluated.

What does the evidence say about effective ICT literacy practices? Slavin (2015) has pointed out that we are currently in a strong movement towards an evidence-based reform in education. Slavin argues that for this movement to be successful, there is a need for sustained investment in development, evaluation, and scale-up of proven programs. In addition, he has argued for a change of polices to encourage the use of such proven programs. The preceding review focussed on the evidence base for effective ICT literacy practices. Taken together, the examination of the effectiveness literature on recommended literacy intervention practices has demonstrated that strong long-term systematic evidence is lacking, and thus the need for more long-term follow-up studies (Torgerson et al., 2006).

Currently, the absence of quality RCT-only systematic analyses in areas of reading research means that any generalizations reported concerning the effectiveness of literacy practices must be accepted cautiously. Moreover, this caution is not limited to ICT literacy practices because even for phonics intervention practices, high quality evidence is sparse, and the evidence base for other areas of reading is even less strong (McArthur et al., 2012; Torgerson et al., 2006). A review by Savage (2012) has drawn similar conclusions that despite a large body of literature on most aspects of reading, the highest quality evidence for the effectiveness of many recommended literacy intervention practices is lacking. Savage bases his argument on the fact that presently in reading research, the strongest evidence—from the meta-analysis of random control trial studies—is rare. Other than a few exceptions (e.g., Torgerson & Zhu, 2003), RCT-only meta-analysis examining the effectiveness of literacy intervention approaches continues to be a rarity. Furthermore, the trend appears to be a continuing decrease in intervention and randomized experimental research, yet an increase in recommendations for practice based on non-intervention observational/correlational research (Reinhart, Haring, Levin, Patall, & Robinson, 2013). This continual decrease is of concern, since it suggests that policy recommendations are not being made on sound empirical findings. Recently, Savage and Cloutier (2016) reviewed every article published in the *Scientific Studies of Reading* journal (n = 351) and found only 14% of these studies (n = 48) could be considered an intervention study, with only a portion of these studies unmistakeably being RCT studies. Without the use of RCTs, the opportunity to explore or confirm genuine causal models of effective ICT literacy practices is not possible.

However, results from classic RCT experiments may never be able to answer all questions about the effectiveness of reading interventions. Even though classic experiments are valuable for identifying effective literacy intervention practices that have produced positive group differences with respect to children's reading outcomes, the findings of classic experiments often neglect to address individual variations in the response to interventions considered empirically sound (Al Otaiba & Fuchs, 2006). The next chapter discusses the value of extending the examination of the effectiveness of intervention studies from is the reading intervention practice effective or not to for whom is the reading intervention effective.

Chapter 3: Pupil-level Response to Intervention

Overview

Chapter 3 examines the value of response-to-intervention (R*t*I) studies in the field of reading research. It provides a brief review of what traditionally has been the focus of the R*t*I literature at the pupil level. Also addressed is the need for researchers to extend their analysis of the variables impacting students' R*t*I to include an examination of the sociological factors that may influence a child's R*t*I. This chapter concludes by introducing the two pupil-level R*t*I theoretical models (i.e., the inoculation and insulin hypotheses) that were investigated for this present study.

Response-to-Intervention (R*t*I) Overview

The term *response-to-intervention* (R*t*I) has become an important part of the research lexicon for effective interventions. Researchers have examined the differences between individuals with respect to their responses to an intervention or have assessed the implementation of formal R*t*I models of instruction in the classroom. R*t*I is grounded in the determination of whether an adequate or inadequate academic or behavioural change in performance has been attained as a result of a intervention (Foorman & Wanzek, 2016; Gresham, 2007). Broadly, R*t*I can refer to a process of implementing an intervention that is followed by the collection of data to determine whether the intervention was effective in correcting the problem (McIntosh et al., 2011).

The R*t*I concept suggests students will benefit from exposure to a well-run, evidence-based intervention. As well-designed experiments have been shown to have an effect on the early literacy skills of children, at least in the short-term (Fuchs & Vaughn, 2012). Thus, the general R*t*I concept contends that overall, intervention groups will fare better over control groups that do not receive an intervention in well-run reading studies. However, R*t*I studies highlight that significant variations occur within intervention groups, and thus, R*t*I researchers are interested in examining this within-group variation (Al Otaiba, Calhoon, & Wanzek, 2010).

Gresham (2007) has added that the value of employing an RtI approach is that this method functions as a risk model rather than a deficit model approach, since the R*t*I method emphasizes that all students need be screened for potential learning difficulties early in their school careers (e.g., Kindergarten to Grade 2). This is of value as early intervention has greater efficacy when compared to remedial services provided after the second grade (Catts, Nielsen, Bridges, Liu, & Bontempo, 2015; Greenwood et al., 2014). Since the application of R*t*I stresses the value of early and effective teaching and learning, the International Reading Association (IRA) suggests that a better term for R*t*I could be *response to instruction* (Farstrup, 2007). The IRA also has argued that R*t*I should not be used solely as an extension of learning disabilities or special education programs; rather, it should be integrated as an aspect of all classroom level instruction so to best identify the students who are struggling academically.

R*t*I is considered a policy of providing a continuum of supports—from universal support for all students within a classroom to specialized instruction for those demonstrating more need (Denton, 2012; Fien et al., 2015). A three-tiered R*t*I model is most commonly employed in educational settings, with Tier 1 as the delivery of general classroom instruction, Tier 2 as often involving supplemental small-group instruction, and Tier 3 as more intensive, individualized instruction, such as one-to-one tutoring or an assignment to special education services (Foorman & Wanzek, 2016). Some researchers have suggested that students at risk of reading difficulties may benefit from the more individualized and small-group support provided by the higher R*t*I tiers (Tier 2 and 3) (Al Otaiba, Connor, et al., 2014; Hatcher et al., 2006), while others have proposed that the learning of at-risk students could be enhanced best by simply improving the whole classroom (Tier

1) literacy instruction (Foorman & Wanzek, 2016; Slavin et al., 2011). Slavin et al. (2011) found that low achievers benefited from being in classrooms where teachers changed their practice and adopted cooperative learning or structured phonetic classroom models. In addition, Slavin et al. (2011) have maintained that if enhancements can be made at the Tier 1 level of instruction, this strategy may help to avoid the possible stigmatization, frustration, and demotivation that some children feel when they are singled out to receive additional support. Further, they have argued that enhanced support at the classroom-level instruction may reduce the difficulties, expense, and disruptions that often are inherent when trying to provide supplemental small-group or one-to-one tutoring services. Recently, Lam and McMaster (2014) and Al Otaiba, Wagner, and Miller (2014) have noted the lack of examination of Tier 1 instruction in multi-tier research despite the noted value of ensuring the effectiveness of Tier 1 as the foundation for early prevention. Rather, most R*t*I studies primarily explore the effectiveness of Tier 2 or 3 levels of an intervention, with examinations of the implementation and quality of Tier 1 interventions almost entirely neglected (Fien et al., 2015; Hill et al., 2012), which suggests a need for additional research on classroom level R*t*I.

A growing movement has arisen to use more R*t*I informal early assessments to identify a child at risk of a reading disability, rather than waiting until the child is older and relying solely on IQ tests for formal coding of a reading disability (Catts et al., 2015; Stuebing et al., 2002; Vaughn & Swanson, 2015). This suggests that continued investigations into identifying and validating effective reading interventions are still needed. Specifically, a need still exists for reading interventions that teachers can easily integrate into their existing reading programs to help reduce the number of students subjected to the *wait and fail* model (Al Otaiba, Wagner, et al., 2014). Supporters of the R*t*I approach have contended that if used effectively, R*t*I models employing validated reading interventions, especially at the Tier 1 stage of instruction, have the potential to assist teachers in the early identification of students at risk of reading difficulties (Slavin et al., 2011). Also, school boards that promote the use of R*t*I practices will be able to avoid this traditional practice of waiting for a student to fail before providing remedial services.

Questions still to be explored. Previous reading research has demonstrated that early literacy interventions can be effective for improving the reading skills of young readers in the short-term (e.g., Bus & van IJzendoorn, 1999; Camilli, Vargas, Ryan, & Barnett, 2010; Slavin et al., 2011). One specific area of focus of early literacy intervention efforts has been to address the phonological deficits of students identified as at risk of reading difficulties (e.g., Ehri et al., 2001; Hatcher et al., 2006; Snowling & Hulme, 2011). Such intervention studies have demonstrated that effective literacy interventions can have immediate positive effects on children's early reading skills. Al Otaiba and Fuchs (2006) also found that at least in the short-term, well-designed interventions generally assist students' reading growth.

As discussed in length in Chapter 2, the selection of studies examining the longitudinal effectiveness of reading interventions is relatively sparse in comparison to the volume of the reading research studies that have been conducted (Savage & Cloutier, 2016; Suggate, 2014). Since most follow-up studies last less than 2 years (Pfost, Hattie, Dörfler, & Artelt, 2014; Suggate, 2014), more long-term follow-up studies are needed to examine whether the gains in reading skills developed immediately after the implementation of early reading interventions can be maintained over time (Al Otaiba, Kim, Wanzek, Petscher, & Wagner, 2014; Blachman et al., 2014).

Recently, however, one notable study by Blachman et al. (2014) evaluated the reading outcomes of students who had participated in an 8-month reading intervention 10 years earlier. Blachman et al. (2014) research interest in the long-term educational impact of reading interventions was influenced by Stanovich's theoretical framework, the *Matthew effect* assertion that *the rich get richer* (Stanovich, 1986). The Matthew effect in the reading literature contends that children who receive early reading intervention and become stronger readers early on will continue to have more positive reading experiences, be exposed to more print, and consequently develop even stronger reading skills. In addition, the assumption is that through the process of *self-teaching* (this model is discussed in greater detail later in this chapter), those benefiting from early literacy skills intervention (i.e., phonics, word reading interventions) will experience its positive effects extending non-linearly to later reading domains and outcomes (i.e., reading accuracy and comprehension). In contrast, young children who have a weaker literacy foundation will read less in the future, and hence their reading attainment growth and wider educational outcomes will be inhibited in comparison to their stronger reading peers. As a consequence, the Matthew effect phenomena suggests that without early reading intervention for weaker students, a widening of the achievement gap will continue between initially poor and strong readers as they continue onto high school (Stanovich, 1986). At the time of their 10-year follow-up, Blachman et al. (2014) found that the intervention students who had received the 8 months of explicit reading treatment outperformed the control students on 2 of 12 reading measures (i.e., word reading measures). Although the Blachman et al. findings provided some longitudinal evidence for higher reading outcomes for those students receiving reading interventions, the results did not provide conclusive support for the Matthew effect as the study did not find any evidence of an increased achievement gap over time between the good and poor readers on later reading skills and reading fluency. This finding is consistent with a recent a meta-analytic review of the longitudinal data on Matthew effects by Pfost et al. (2014), which also found a lack of significant support for a pattern of a widening achievement gap between good and poor readers in later years. Limitations to the Blachman et al. study, such as a small sample size (*N* = 58) and a lack of conclusive findings, suggest that additional long-term research support is still required to provide a true test of the Matthew effect.

Many questions remain concerning the long-term effectiveness of early prevention intervention efforts (Coyne et al., 2004; Savage & Cloutier, 2016; Suggate, 2014). For example, one issue with follow-up studies is not knowing what happens with participants after the study ends and prior to the follow-up. For example, in some instances, as it was with the ABRA studies, the control classrooms have the option to be taught and exposed to the intervention program once the original study has finished. In other instances, schools might intervene with the 'weaker' control group children after post-test. The uncertainty of which participants may have received additional or remedial support and the type of additional support they may have received post-intervention is an issue that longitudinal researchers need to be conscious of and address (Blachman et al., 2014).

Furthermore, Coyne et al. (2004) have contended that classic intervention studies are problematic because at follow-up, they merely test for statistical differences between the experimental group and control group to determine whether the intervention produced long-term effects. Coyne et al. (2004) stress that a need exists for more R*t*I research that will examine the in-group variance of the children who respond and do not respond to well-designed interventions. Another question to consider is: Do students who show progress following an intervention program and catch-up to their peers maintain this progress? Coyne et al. (2004) addressed this inquiry, contending that much of the current focus has been on students that they refer to as *treatment resisters*. However, a more suitable term used in the literature to classify the students who do not respond to reading interventions is *non-responders* (see Al Otaiba & Fuchs, 2002; Savage, Carless, & Erten, 2009; Wanzek & Vaughn, 2008). Non-responders or low responders are 20 to 40% of the sample of students at risk of reading difficulties that show little or no growth despite receiving an intervention (Denton, 2012; Vaughn et al., 2009). Reading researchers' definitions of the non-responsiveness of struggling students vary from performance below the 10th percentile to a score below the 50th percentile on a given reading skill measure (i.e., O'Connor & Klingner, 2010; Torgesen, 2000; Vellutino, Scanlon, Zhang, & Schatschneider, 2008). Coyne et al. (2004) have concluded that more research is needed to examine the enduring effects of early intervention for students who have shown the strongest response to treatment: "that is, those at-risk students whose reading-related skills after intervention are no longer distinguishable from their peers who are not at risk" (p. 92).

Relatively few studies have examined response to intervention longitudinally, the evidence base for the long-term effects of early intervention is limited (Lam & McMaster, 2014; Suggate, 2014). Moreover, the evidence base for long-term efficacy is rare for children with early reading difficulties, and in particular, R*t*I research relevant to Canadian content continues to be sparse (McIntosh et al., 2011). Thus, despite the success of early reading intervention studies in improving the reading outcomes of many students with reading difficulties, many variables still need further examination. Specifically, a number of questions remain unanswered, such as why some students improve after exposure to an effective intervention while others progress at a much lower rate and continue to struggle with reading even after being involved in an effective intervention?

Pupil-level Variations in R*t*I

This section addresses the issues described in the preceding paragraph by examining some of the pupil-level variables that may influence a student's response to intervention.

Pupil-level R*t*I predictors. When undertaking an R*t*I study, Al Otaiba and Fuchs (2006) have contended that researchers should be mindful of potential pupil-level variables that may influence a child's response to otherwise effective early literacy interventions. Some support has been found for a number of pupil-level characteristics of non-responders to well-designed interventions. Lam and McMaster (2014) have defined a *variable* as a predictive factor of responsiveness if at pre-treatment performance levels it can assist in distinguishing between students who respond to an intervention versus those who do not. While an in depth examination of all potential pupil-level predictive variables is beyond the scope of this paper, it is nevertheless important to draw attention to a few of the key pupil-level R*t*I predictors that have been identified by researchers.

For example, Torgesen et al. (2001) found that the best overall predictors of long-term growth were resource room teacher ratings of student's attention/behaviour, general verbal ability, and prior levels of component reading skills. Reviews by Al Otaiba and Fuchs (2002) and Nelson, Benner, and Gonzalez (2003) on effectiveness studies examining learner characteristics also have reported a range of pupil-level R*t*I predictors consisting of, but not limited to, basic meta-cognitive difficulties in the phonological domain. For example, Al Otaiba and Fuchs (2002) reviewed 23 studies in a literature search spanning from 1967 to 2000 in which the influence of non-responding children's characteristics were examined. The authors found that phonological awareness deficits were the most prominent predictor of R*t*I identified in the 23 studies. Nearly all of the groups of investigators (n = 21) examined the importance of phonological awareness in children's reading development process. Of these 21 studies, most (n = 16) reported that children who were non-responsive to early literacy interventions demonstrated phonological deficits. In their review, they found phonological awareness to be a predictive factor in 16 studies, not predictive in 1, and an inconclusive factor in 4 studies. In addition, Al Otaiba and Fuchs (2002, 2006) identified six additional categories of child characteristics most commonly associated with non-responsiveness in the literature, and each of these factors varied in their predictive nature: (1) phonological memory [predictive in 4 studies, not predictive in 2, and mixed in 1], (2) rapid naming [5 predictive, 1 not predictive, 1 mixed], (3) vocabulary, verbal ability, and IQ [5 predictive, 7 not predictive, 3 inconclusive], (4) attention or behaviour problems [7 predictive, 2 not predictive, 0 mixed], (5) orthographic awareness or spelling [3 predictive, 4 not predictive, 0 mixed], and (6) home background and demographics (including socioeconomic status, parent education, and student's familiarity with English) [2 predictive, 0 not predictive, 2 mixed].

In their meta-analytic review of 30 well-designed early literacy intervention studies, Nelson, Benner, and Gonzalez (2003) identified many predictive factors similar to those identified by Al Otaiba and Fuchs (2002). From the literature from 1966 to 2000, concerning the treatment of non-responders, Nelson et al. not only examined key individual and cultural pupil-characteristics but also the magnitude of prediction of each factor. They identified seven primary factors from all of the pre-intervention learner characteristics for which an effect size for intervention effectiveness could be calculated. In order of magnitude, they ranked the seven primary learner characteristics identified as influencing the treatment responsiveness of early literacy interventions by their level of predictive power (the most to least predictive): *rapid naming* ($Z_r = 0.51$)*, problem behaviour* ($Z_r = 0.46$)*, phonological awareness* ($Z_r = 0.42$)*, alphabetic principle* ($Z_r = 0.35$)*, memory* ($Z_r = 0.31$)*, IQ* ($Z_r = 0.26$)*, and demographics* ($Z_r = 0.07$). The effect sizes for the majority of the core factors were moderately large, with the exception of the demographic factor (i.e., disability, ethnicity, grade-level status), which was not statistically distinct from zero. As well, even though intelligence (I.Q.) was a significant factor, it was modest in size when compared to the other five reading-related cognitive processes and behaviour.

A more recent review by Lam and McMaster (2014) has provided an updated examination of predictive student responsiveness factors to early literacy intervention. Their review findings were somewhat consistent with the findings of Al Otaiba and Fuch (2002) and Nelson et al. (2003), since Lam and McMaster also found somewhat similar support, with some differences, for the predictiveness of the factors identified in earlier reviews. In their synthesis of 14 studies from 2000 to 2012, they identified and classified potential predictive factors of responsiveness as consistently predictive, inconsistently predictive, or generally not predictive. Word identification, alphabetic principle, fluency, and phonemic awareness were factors identified as consistently predictive. Intellectual functioning and memory were found to be inconsistently predictive, while vocabulary, language, and demographics (race, sex, and socioeconomic status) were factors identified as generally not predictive of responsiveness.

As noted earlier, one identified pupil-level R*t*I predictor was *attention* (Al Otaiba & Fuchs, 2006), since reading and attention issues have been shown to overlap in early literacy intervention research (Deault, Savage, & Abrami, 2009; Stage, Abbott, Jenkins, & Berninger, 2003). ABRACADABRA (ABRA), a technology-based reading intervention program, has shown to be an effective tool to use with students with attention difficulties. For example, negative reading outcomes—as a response to effective reading instruction of children identified at risk of attention difficulties— have been shown to improve with exposure to ABRA. Deault et al. (2009) examined the connections between reading and attention within the context of reading instruction and found that for children who were participants in the ABRA reading intervention, inattention was not a predictive factor for phonological skills and reading achievement at the end of Grade 1. However, for the children who remained in the control condition (i.e., with regular non-ABRA language arts lessons), inattention remained a significant pupil-level predictor of reading outcomes. Deault et al. (2009) have attributed this effect to key elements of the ABRA intervention program that appear to be beneficial for children at risk of attention difficulties, such as the provision of visually-engaging material, opportunities for active engagement, peer interaction, and immediate feedback. These findings suggest that the integration of visually-engaging and interactive technology innovations into early literacy programs may have a positive influence on pupil-level variables, such as inattention, which may have traditionally placed some children at a greater risk of developing long-term reading difficulties.

Savage and Deault (2010) also have provided a valid critique of some of the findings of other meta-analyses such as those carried out by Nelson et al. (2003), commenting that the results of the meta-analyses reflect "the samples in individual studies, many of which used the then-largely unquestioned I.Q.– discrepancy model with all of its associated exclusions of children assumed to experience 'other' problems" (Savage & Deault, 2010, p. 585). In addition, Tran, Sanchez, Arellano, and Swanson (2011) have argued that the Nelson et al. (2003) synthesis was limited, since these authors failed to compare responders and non-responders at pretest to determine if pre-existing conditions predicted R*t*I outcomes.

Another issue of concern associated with meeting the inclusion criteria for meta-analyses is whether, for a given variable, sufficient analysable data were reported in the R*t*I studies. For example, of the 30 early literacy intervention studies that met the inclusion criteria for Nelson et al.'s (2003) meta-analysis, only 5 of the studies reported on the children's demographic information. Of these five studies, three reported analyzed data on the effects of a known disability on R*t*I; one study (Hurford, 1990) reported findings on the impact of grade-level or grade-retention; and only one study (Foorman, et al., 1998) included analyzed data on ethnicity as a factor in predicting R*t*I. This concern leads to the following question: Was sufficient data available to draw a conclusion concerning the impact of these three subsumed demographic variables on the effectiveness of early literacy interventions? Specifically, if researchers were interested in the impact of ethnicity on a child's R*t*I, it would be impossible to evaluate by referring to Nelson et al. (2003) meta-analysis findings.

As mentioned previously, the R*t*I predictive variables that have most commonly been examined are cognitive response variables (i.e., phonological awareness, letter-sound knowledge, short-term memory, etc.). However, if we are to obtain a comprehensive account of response to intervention, there also exists for additional research on the predictive R*t*I nature of non-cognitive pupil-level variables. For example, there exists a need to extend the existing literature on R*t*I to include research on student characteristics like ethnicity, language, or socioeconomic status that have yet to be fully explored (Al Otaiba et al., 2011; Klingner & Edwards, 2006; Vanderwood & Nam, 2007). One contributing factor for the need for this additional research may be that generally, when reporting the results of intervention studies, researchers have neglected to provide information about their students' ethnicity or socioeconomic status. Further, the use of the R*t*I model may assist in reducing the number of students whose first language is not English being referred to special education services (Graves & McConnell, 2014; McIntosh et al., 2011). As one major feature of the R*t*I model is the role of high quality learning experiences in which many students with modest experiences of English may benefit compared to the IQ discrepancy-based systems. However, these assumptions need to be explored empirically.

Thus, the value of the R*t*I method over traditional approaches is that R*t*I may be better at distinguishing between students at risk of reading difficulties that are able to benefit from short-term intervention and those who genuinely have a reading disability. In many instances, second language learners whose reading development may be months or even years behind their peers are responding to appropriate classroom-level interventions, but at a delayed rate.

While the examination of non-cognitive pupil-level factors is limited in the current R*t*I literature, the significant effects of SES, ethnicity, and sex-differences have been reported consistently in educational assessment studies using standardised test measures (e.g., Sammons & Smees, 1998; Savage & Carless, 2008; Savage, Carless, & Ferraro, 2007) and in national test data (e.g., in England and Wales, DFes, 2007). Savage and Deault (2010) have contended that based on these educational trends, it would be surprising if R*t*I research did not reflect similar findings. Further, these authors have discussed how the evidence base concerning R*t*I predictive factors shows support not only for cognitive processes (e.g., phonological skills, letter knowledge, and receptive vocabulary) but also for a number of socio-cognitive factors (e.g., home literacy and language environment). Ultimately, any discussion around intervention needs to look at what we believe works, which should be anchored in the evidence-based approach. Finally, to provide effective intervention as early as possible, the aim of R*t*I research continues to be the development of valid procedures for an early identification of students at risk of reading difficulties, and further to identify pupil-level R*t*I predictor factors that may hinder or facilitate a child's R*t*I.

Pupil-level R*t*I Theoretical Models

This section examines two theoretical hypotheses concerning how and why students may respond differently to effective reading interventions: the **inoculation** versus the **insulin** hypothesis. Borrowing primarily from (Coyne, Kame'enui, & Simmons, 2001; Coyne et al., 2004), the primary assumptions of the inoculation and insulin theories can be used to explain the degree of *within-group pupil-level variation* R*t*I patterns that often are observed. These two theoretical positions can be used to examine the within-group variations in students' short-term R*t*I and their long-term response to effective early literacy interventions.

Inoculation hypothesis. This hypothesis suggests that well-designed literacy interventions, strategically and intensively delivered within a specific window of time, should be able to remediate "the phonological and alphabetic deficits of a significant percentage of children who are initially identified as at risk for reading difficulties, making further intensive intervention at a subsequent time during reading development unnecessary" (Coyne et al., 2004, p. 91). In other words, according to the inoculation assumption, effective early literacy intervention can be a preventive measure to ward off later reading difficulties in children, since the intervention *jump starts* children's acquisition of essential skills, which facilitates the reading development process (O'Connor, 2000).

The foundation of the inoculation hypothesis draws on Share (Share, 1995) and Stanovich's (1986) hypotheses of children's early reading acquisition. Share's *self-teaching* model suggests that if children are able to develop a strong phonological and alphabetic knowledge base, they are better equipped to engage in a self-teaching mechanism. Stanovich's work examined the compounding impact of the *Matthew effect* on the reading achievement differences between children who enter school with a strong early literacy foundation versus those who do not. Share and Stanovich (1995) have suggested that an early onset of self-teaching parallels the beginning stages of learning to read. They have argued that a self-teaching mechanism will further facilitate the development of complex higher order reading skills. Specifically, Share (1999) has identified three factors that are needed to facilitate a child's early self-teaching ability: "Letter–sound knowledge, some minimal phonological sensitivity, and the ability to utilize contextual information to determine exact word pronunciations on the basis of partial decodings" (p. 97). All three of these factors assist a child's phonological recoding ability, which corresponds to the ability to translate printed words into their spoken equivalents.

Share (1995, 1999) has concluded that phonological recoding is critical for the acquisition of word-specific orthographic representations as suggested by the self-teaching hypothesis. Experimental studies have produced evidence to support the impact of phonological recoding on the storage of orthographic representations (e.g., Cunningham, 2006; Ecalle, Magnan, & Calmus, 2009). In addition, other researchers (e.g., Scheltinga, van der Leij, & Struiksma, 2010) have suggested that if a child can assemble pronunciations based on graphemes and phonemes, she/he would be equipped with a mechanism for independent learning. However, the strong claim that Share makes of item-specific learning (i.e., decoding a specific word means storing word-specific orthographic knowledge for that particular word) has not been supported (see McArthur et al., 2012; Ricketts, Bishop, Pimperton, & Nation, 2011). For example, McArthur et al. (2012) found that there was a lack of casual evidence support of early phonics training for poor readers leading to later improvements in wider word reading and spelling. Further, others add that the inoculation focus on the phonics aspect of literacy neglects to consider the complexity of the deep orthography aspect of English, suggesting that both morphology and phonics training are needed in early literacy education (Savage & Cloutier, 2016).

In general, the assumption supporting the inoculation hypothesis is that if a child is taught phonics strategies in an alphabetic system, he/she has the foundation and strategies to begin the self-teaching process and learn new words on his/her own. Supporting this R*t*I model is the belief that a strong phonics strategy is needed to learn how to read; that is, if students at risk of reading difficulties respond favourably to an intervention and cement the skill of phonics awareness, they should be able to break the code and read new words when exposed to them. The inoculation hypothesis contends that students who make adequate gains in beginning reading intervention programs will benefit from the reciprocal effects of establishing strong phonological foundations and alphabetic skills equal to those of their peers. Thus, these students should continue to make acceptable reading progress without additional support (Coyne et al., 2004). Further, empirical evidence supports the value of teaching phonics rather than just sight words or words in context (the whole language approach), which on their own are not as effective for fostering the self-teaching process in children (Share, 1995, 1999). Thus, the general idea of the inoculation theory is the following: if provided with the proper strategies, children should be able to apply them and self-teach themselves to read new words after acquiring the essential understanding of how words can be read.

Insulin hypothesis. In contrast, the *insulin hypothesis* suggests that the "positive short-term effects (i.e., the elimination of at-risk status) gained through early intervention can be maintained only with continued intensive support" (Coyne et al., 2004, p. 91). This assumption proposes that even though substantial treatment effects are evident immediately after intervention, such gains can only be maintained through additional continued reading support efforts. Similarly, some researchers have noted that early intervention for students at risk of reading difficulties should not be viewed as a one-shot inoculation against reading difficulties (Blachman et al., 2014; Shanahan & Barr, 1995). Such opponents of the inoculation hypothesis have found that early intervention does not necessarily completely prevent reading difficulties in the future, especially as some children's reading problems are more severe than others. Snow, Burns, and Griffin (1998) have contended that thinking of early phonological awareness intervention as a *one-shot inoculation* against reading difficulties for all students at risk is not realistic (p. 251); rather, they maintain that ongoing effort often is required. Contrary to the inoculation hypothesis, the insulin theory suggests that without ongoing intervention, at-risk children initially identified with early literacy difficulties with phonological and alphabetic skills remain vulnerable to developing reading difficulties.

When looking at some at-risk children or specifically children with reading difficulties (RD) (e.g., many of the children currently identified as dyslexic), Locke et al. (1997) argued that the insulin hypothesis could be viewed more as a psychopathological model. In other words, the insulin hypothesis suggests that for some children, one can provide a literacy intervention program deemed to be effective, and teach them the essential phonics skills over and over, although they will continue to not respond very well. Even though immediately after the intervention, the RD children may respond a small amount, the insulin theory maintains that they will not catch up to their typically reading peers (Coyne et al., 2004). Rather, the RD students will continue to have ongoing reading deficits, since they may have an inability to acquire the strategies being taught. So, the RD children never arrive at the self-teaching state, since the insulin hypothesis proposes that these children may have an *underlying processing deficit* that is quite different from the environmental notion suggested by the inoculation hypothesis.

Siegel (2006) has suggested that some of the common underlying cognitive difficulties of children with RD include learning sounds of letters, speech perception problems, recognizing and manipulating the basic sounds in a language. Siegel added that children with a reading disability may have brain abnormalities and that dyslexia is a neurological condition with genetic underpinnings. Also, some research has indicated that children with parents with dyslexia are at heightened risk for developing it themselves. For example, Gallagher, Frith, and Snowling (2003) and Bryne et al. (2009) have found that in instances of dyslexia, a genetic predisposition exists that places these children more at risk than children without the genetic component.

On the other hand, researchers also have found that if children, including those identified with an RD, are provided with a strong literacy environment, most will respond positively and demonstrate some growth in their reading skills, which suggests that genetics only plays a partial role (see Snowling & Hulme, 2011; Taylor, Roehrig, Hensler, Connor, & Schatschneider, 2010). For example, a recent twin study by Taylor et al. (2010) found that the quality of teaching influences gains in children's reading development. These authors found that when teacher quality was very low, the genetic variance on reading outcomes was constricted, whereas when the quality of teaching was very high, the genetic variance increased, which suggests an interplay between environment (teacher quality) and genetics. Such findings also support the impact that quality teaching can have on pupils' responses to intervention. However, the reliability of the estimates of heritability in the Taylor et al. (2010) study has been questioned by researchers who assert that the Taylor et al. twin sample was too small to provide reliable point estimates (Kovas et al., 2013).

With respect to RD students, the inoculation and insulin theories propose two different visions of why children may be struggling with acquiring the necessary skills to self-teach and subsequently learn how to read. A practical problem arises when teachers meet children who are struggling with reading: ***How do they decide the nature of at-risk children's problems?*** Is it as the inoculation hypothesis supposes that the problem lies within the RD children's environment, since they have not received the necessary teaching of fundamental skills? Or is it more in line with the insulin assumption in that the children at risk of reading difficulties may have a genetically predetermined deficit (risk) that may be hindering the effectiveness of literacy interventions, and therefore, some children will require ongoing support rather than a one-shot exposure to a literacy intervention deemed effective?

A third position: An inoculation-insulin combination. For educators, the practical questions become: How many children at risk of reading difficulties are merely dealing with environmental issues, and who will become strong readers after they receive effective literacy instruction? How many of these students at risk of reading difficulties have true deficits and need more ongoing support, rather than a one-shot exposure to an effective intervention program? Which students at risk are dealing with both environmental and true deficits with respect to their difficulties with reading? Coyne et al. (2004) have contended that the reality of ongoing R*t*I effects are more complex than can be explained by the inoculation-insulin dichotomy, since a multitude of mediating factors may influence the effectiveness of the intervention and students' responses to the intervention. These authors have summarized their position by suggesting that such a dichotomy is not an either-or phenomenon: "For some students under certain conditions, intervention may act more like an inoculation, whereas for other students under different conditions, intervention may be similar to the first shot of insulin" (Coyne et al., 2004, p. 92).

Coyne et al. (2004) suggested that not only can some of the previously mentioned student-level factors (e.g., phonological awareness deficits, attention deficits, etc.) affect the enduring effects of beginning reading interventions, but also that methodological and instructional factors can have mediating effects. For instance, methodological factors that could influence students' R*t*I may include the type of reading outcome measures used at follow-up (e.g., word identification or comprehension), the time of the follow-up (e.g., 3 months, 1 year, 2 years), and the type of follow-up analyses used (e.g., examining absolute levels of achievement or mean differences between treatment and control groups). In addition, instructional mediating factors could include the nature and quality of the intervention; the role that the focus, duration, and intensity of the intervention play; and whether the intervention instructions are consistent with the subsequent classroom instruction. In this sense, for those students who require a more interactive learning environment, an intervention that provides an engaging technological component in line with the general classroom instruction may positively influence their short- and long-term response (e.g., see Deault et al., 2009; Van Daal & Sandvik, 2013).

A fourth position: Is R*t*I of value? At the pupil-level, both the inoculation and insulin positions propose that R*t*I actually works; however, these approaches differ in how they suggest it works. In the simplest terms, the inoculation theory suggests that when appropriate R*t*I is initially well delivered, students at risk of reading difficulties will respond favourably to the one-time intervention, and ongoing support will not be necessary. On the other hand, the insulin theory suggests that R*t*I works as long as the student at risk of reading difficulties continues to receive the intervention and ongoing support. However, a third position that some researchers have advocated suggests that R*t*I does not work, and its efforts may not be necessary for the identification of students with reading difficulties.

For example, in their recent meta-analysis of R*t*I literature for children at risk for reading disabilities, Tran, Sanchez, Arellano, and Swanson (2011) have proposed an alternate position on the effectiveness of interventions and potential mediating factors. Contrary to the findings of Al Otaiba and Fuchs (2002), Tran et al. (2011) did not find that phonological awareness was a significant moderator of effect sizes on post-test reading measures when all other measures were controlled. Rather Tran et al. (2011) have argued that the best predictor of R*t*I outcomes is pre-intervention reading ability. They found that students' pretest performance fairly accurately predicted their post-test performance, reporting an overall correlation coefficient of .72 between pretest and post-test measures. These authors have contended that regardless of the length or type of intervention treatment, R*t*I conditions were not as effective in mitigating learner characteristics (e.g., identifying students at risk of reading difficulties/low responders) when compared to pre-intervention reading ability, since their synthesis found that the achievement gap (differences) between responding and low-responding children remained consistent across pre- and post-test conditions. Therefore, they concluded that a reliable prediction of students' later reading ability from pre-test scores alone is sufficient, as the authors did not find that R*t*I post-test scores provided any additional predictive value, beyond that of pretest scores, when it came to identifying students at risk (Tran et al., 2011). Rather, such findings support a third sceptical position that R*t*I may not be effective for identifying students at risk of reading difficulties any more than pretest scores on appropriate reading measures.

Along these lines, Tran et al. (2011) have suggested that the notion of R*t*I may be redundant, since students' pretest scores can account for over 30% of the explainable variance of post-test effect size outcomes between responders and low-responders. On the other hand, a counter argument can be presented that this 30% represents a minority of the total group of children, and thus, an amount of movement exists that may be explained by the introduction of R*t*I. This viewpoint suggests that the Tran et al. (2011) study does not undermine the value of R*t*I fully; rather, it merely demonstrates that some children (approx. 30%) may just not be responsive or benefit from exposure to academic interventions.

Non-Responders

As defined by Torgesen (2000), *non-responders* are children who despite participating in preventative literacy interventions fail to acquire word-reading skills within the normal range of the rest of their peers. Until relatively recently, little research existed on children who were non-responsive to early treatment interventions deemed as effective. Around the mid-1990s, groups of investigators started to design studies with the intention to identify the characteristics of non-responsive children (for a review of earlier studies see Al Otaiba & Fuchs, 2002). Since these initial investigations, scientific curiosity continues to grow around the concept of non-responders and on identifying characteristics of those children who do and do not respond to effective interventions.

Similar to the findings of Tran et al. (2011), other researchers also have found that as many as 30% of children at risk for reading difficulties may be unresponsive to generally effective Tier 1 classroom instruction and to Tier 2 early literacy interventions (e.g., Mathes, Howard, Allen, & Fuchs, 1998; Richards, Pavri, Golez, Canges, & Murphy, 2007; Wanzek & Vaughn, 2008). In addition, Al Otaiba and Fuchs (2006) have suggested that the percentage of non-responders among children with learning disabilities may be as high as 50%. In their earlier review of 23 R*t*I studies, Al Otaiba and Fuchs (2002) found that depending on the individual study and outcome measure, anywhere from 8 to 80% of the at-risk students demonstrated little or no improvement on examined reading outcome measures after exposure to an early literacy intervention.

Al Otaiba and Fuchs (2002) have acknowledged that this large discrepancy in the range of non-responders may be attributed partially to the poor quality of some of the R*t*I studies in their review, especially since 22 out of the 23 studies (e.g., including O'Connor et al. [1993] and Ehri & Robbins [1992]) failed to report any information on the quality of treatment fidelity. This situation suggests that a continuing need exists for stronger quality intervention studies that place greater emphasis on the fidelity of implementation.

In addition, Al Otaiba and Fuchs (2002) have contended that such a high degree of variability of non-responders may be a function of the following variables: (1) the criteria used to operationally define non-responders, (2) the diverse characteristics of the children in the samples, (3) the level of effectiveness of the early literacy intervention, and/or (4) the measure of treatment fidelity (for further details, see Al Otaiba & Fuchs, 2006). These authors have argued that subsequent R*t*I studies should consider how best to address these four variables so to improve the quality and consistent reporting of the R*t*I studies conducted to describe responsive and non-responsive students. Further, a number of researchers have maintained that a need exists for further quality short-term early literacy interventions at the classroom level to best identify those students who are at risk and also not responding to quality evidenced-based instruction (Al Otaiba & Fuchs, 2006; D. Fuchs & Fuchs, 2006; Stein et al., 2008).

Supporters of R*t*I research have suggested that the estimates of up to 80% of non-responders to effective literacy interventions are too high. Rather, researchers such as Vellutino and colleagues (1996; 2006) have found that closer to 80% of the students exposed to an effective early literacy intervention respond and can be identified as either responders or low-responders. Savage, Carless, and Erten (2009) have reported similar findings that two out of three children respond favourably to reading interventions. The reading interventions by Savage et al. (2009) were sustained over time, focused on phonics, and were given by well-trained school staff in small group (n = 4) contexts after school-wide screening. Denton (2012) has found that approximately 20% to 30% of students continue not to respond to Tier 1 interventions and require some form of Tier 2 supplemental support in addition to their core curriculum. Such findings indicate the need for ongoing R*t*I research to identify early intervention approaches that are effective in distinguishing between those students who are responding well versus low-responders who may need additional ongoing support.

What Still Needs To Be Explored in Responsiveness Research?

The current R*t*I focus has been the degree to which a student identified at risk for reading problems benefits from participating in a reading intervention study (Fuchs & Vaughn, 2012). Primarily, R*t*I research has focused on the within-group variation of those who do, and do not, respond well to intervention at the pupil level (e.g., Al Otaiba & Fuchs, 2006; Foorman & Wanzek, 2016). A need still exists for additional research support to identify pupil-level variables that may influence a child's R*t*I. Such research needs to extend beyond an examination of the traditional cognitive response variables (i.e., phonological awareness, letter-sound-knowledge, short-term memory, etc.) to include analyses of demographic factors (e.g., ethnicity, sex, and socioeconomic status) that have been discussed in the literature but have not been fully explored in R*t*I studies.

The primary focus of R*t*I literature on a pupil-level analysis is understandable, since the students are receiving the reading intervention. However, rarely acknowledged is the fact that teachers also are responding to the intervention. Thus, absent from the literature are investigations of the within-group variation with respect to the responses of teachers who implement a new reading intervention in their practice during an intervention study. In addition, while scholars have continued to point out that long-term follow-ups on reading studies are generally rare (see Suggate, 2014; Torgerson et al., 2006), particularly absent in the current R*t*I literature is an examination of the long-term response of *both* sets of participants—students and teachers—in the same study.

An evaluation of the role of teachers in R*t*I studies is an area of research that has yet to be fully explored. In their review analysis of 23 early literacy intervention studies, Al Otaiba and Fuchs (2002, 2006) found that the existing R*t*I literature lacked studies that had examined the literacy interventions conducted by teachers. Out of the 23 studies examined, classroom teachers ran three, and resource room teachers ran two, indicating a lack of well-designed RCT literacy intervention studies implemented by classroom teachers.

As reviewed in this chapter, much has been written on students' responses to reading interventions (e.g., Al Otaiba & Fuchs, 2006; Lam & McMaster, 2014; Pfost et al., 2014); however, less is known about how teachers change their practice short- and long-term in response to participating in an intervention study. This phenomenon is surprising, since one essential component of the whole concept of R*t*I is the teacher. That is, the teacher needs to change their practice to implement the intervention, which indicates that they are responding (in some way) to the intervention they are implementing. The Slavin et al. (2009) review highlights the value of the role of teacher change in improving student achievement outcomes, since these authors have argued that no matter what intervention is used, the greatest impact on the effect size of student achievement happened when teachers changed what they were doing. In particular, the Slavin et al. (2009) study suggested that teacher change is the most important factor when examining the effects of technology interventions on student reading outcomes.

While many studies have looked at factors that influence teacher change (Fullan, 2005; Ross & Bruce, 2007; Wagner & French, 2010; Whitworth & Chiu, 2015), they have not examined this change in an R*t*I context. Identifying teacher-level factors that may predict teachers long-term R*t*I is of value to reading intervention researchers, since greater awareness of predictive teacher-level R*t*I factors could assist researchers in designing better intervention studies. If researchers are able to design reading intervention studies that support teacher-level R*t*I predictive factors, a greater likelihood exists that a teacher's use of an effective literacy intervention practice will transfer from their involvement in a research study to implementation in their long-term literacy practice.

In summary, in the field of education, a shift has occurred towards a model of early identification and prevention of learning difficulties through the use of scientifically-validated reading intervention strategies and high-quality instruction (Pullen, Tuckwiller, Konold, Maynard, & Coyne, 2010). The R*t*I model is the dominant model to emerge for examining how children at risk are responding to quality instruction. However, some key shortcomings of the current R*t*I research still need to be addressed. For example, at the pupil-level R*t*I, researchers have not adequately explored sociological factors, long-term follow-ups, or teacher-delivered interventions. More work in all of these areas needs to be conducted before conclusions can be drawn about the role of any of these factors in students' R*t*I. Thus, this follow-up study has aimed to add to the reading intervention literature by examining these areas of R*t*I research, both at the pupil- and teacher-level where evidence of support currently is inconclusive.

Conducting studies in classroom settings where teachers are carrying out the interventions opens up a whole number of issues about scaled-up field studies and how they should be conducted in the future when examining R*t*I. In such studies, researchers have less control over the implementation of the literacy interventions, since teachers are the primary implementers, in comparison to controlled experiment conditions where the same researcher goes in and provides the intervention. So, questions of fidelity regarding how teachers are implementing the intervention approaches under investigation also need to be further explored in future R*t*I research studies. Such questions addressing teachers' roles in R*t*I are explored in the next chapter.

Chapter 4: Teacher Change and R*t*I

Overview

Chapter 4 examines traditional teacher change models and applies them to a response-to-intervention (R*t*I) framework of analysis. Four domains of influence impacting teachers' change processes (external domain, personal domain, domain of practice, and domain of consequences) are presented. Three conceptual teacher change models also are introduced. These three models are shown to be similar in that each model consists of the same four domains of teacher change; however, the models differ, since each reflects a different trajectory of the teacher change process. This chapter also reviews and identifies potential variables that may best support teacher learning, and their likelihood of long-term change and acceptance of a new innovation.

Teacher Change Literature

A substantial literature base exists on the topic of teacher change (see, for example, Clarke & Hollingsworth, 2002; Fullan, 2005; Guskey, 1986; Kaasila & Lauriala, 2010; Whitworth & Chiu, 2015). This literature includes varying interpretations of *teacher change* as training, adaptation, personal development, and/or learning (Clarke & Hollingsworth, 2002). Kaasila and Lauriala (2010) also have suggested that teacher change has been viewed as a form of growth, which assumes that teachers learn and change in professionally relevant learning situations through the process of acting and interacting. For the purpose of the present study, the term teacher change will be used broadly to encompass all of these interpretations.

Opfer, Pedder, and Lavicza (2011) have asserted that teacher change is a multidimensional process in which a multitude of factors can potentially influence the quality of teacher learning, and subsequently impact teachers' classroom practices. Desimone (2009) has suggested that in the process of implementing a new innovation into long-term practice, many variables play a role that can either stimulate or hinder. In fact, extensive research over the last 30 years in the area of teacher change has led to the identification of a myriad of variables that can potentially influence a teacher's decision to adopt newly acquired knowledge or a novel innovation into practice.

While discussing every possible teacher-level variable is beyond the scope of the present study, an effort was made to highlight the key influencing factors of teacher change that may affect how teachers respond to being part of an intervention study. To accomplish this aim, this chapter begins with a focus on four domains of influence that have been widely used by scholars proposing conceptual frameworks for studying teacher change. Albeit, the terminology used by scholars to identify this four-component structure may vary from model to model (e.g., Clarke & Hollingsworth, 2002; Desimone, 2009; Guskey, 1986). One main objective of this chapter was to demonstrate how these four domains could be used in classifying and examining the key variables of teacher change that also could influence teachers' R*t*I. The four domain designations that were used specifically for the purpose of this present study were drawn from Clarke and Hollingsworth's (2002) model of professional growth, which are categorized as follows: (1) the ***external domain***, (2) the ***personal domain***, (3) the ***domain of practice***, and (4) the ***domain of consequence***. These four analytical domains reflect the contexts in which teacher change occurs.

Research on educational change has found that many new innovations fail to be incorporated into teachers' existing practice (Hoekstra & Korthagen, 2011; Holmes, 1998). Therefore, understanding the potential variables that influence teachers' learning and likelihood of long-term change is important. Such knowledge is of value to researchers and practitioners who may wish to incorporate the best ways to support and promote teacher learning to facilitate the long-term implementation of new knowledge into practice.

Four Domains of Teacher Change

The external domain. This domain represents the various sources of stimuli by which teachers may be exposed to new knowledge. A number of experiences can lead to teacher learning, for example, most of the teacher change literature focuses on teacher learning that occurs after a teacher has attended a professional development (Pro-D) workshop or been involved in a one-day training session on a new innovation or teaching method. Researchers (such as, Desimone, 2009; Penuel, Fishman, Yamaguchi, & Gallagher, 2007) have identified a number of core professional development characteristics that they found to be essential for facilitating teacher knowledge and skills development. With respect to the literature, these authors have suggested that the key core variables that have been found to enhance the effectiveness of teacher learning include: (a) ensuring that the content focus of teacher learning integrates activities that incorporate subject matter content with practices that best reflect how students learn; (b) providing opportunities for teachers to engage in active learning as opposed to passive learning during professional development sessions, which are collaborative and provide opportunities for teachers to interact with peers; (c) ensuring the level of coherence—that the new knowledge being taught to teachers is consistent not only with teachers' knowledge and beliefs and can be embedded in teachers' work, but also is consistent with school, district, and provincial policies and reforms; (d) making sure that the span of time during which teachers are exposed to the new teaching learning activity (e.g., a one semester exposure versus an afternoon workshop) and the hours spent actively engaged with the activity are of sufficient duration, and that adequate time is provided for follow-up support.

In the teacher change literature, professional development (Pro-D) workshops have been the external source of information predominately studied (Diamond & Powell, 2011; Kennedy, 2016; Maskit, 2011). However, as Borko (2004) has asserted, new teacher learning does not exclusively occur as a result of teachers attending Pro-D seminars or half-day workshops; rather, teacher learning can occur in many different aspects of practice, including their classroom and school communities. In addition, although Clarke and Hollingsworth (2002) have distinguished the external domain as outside of the teacher's personal world, the present study suggests that new knowledge from external sources, other than Pro-D workshops, can be brought into a teacher's personal world of practice. For example, teachers can obtain new knowledge when they take part in an intervention study by implementing a new program or innovation into their curriculum. Teacher change and learning as a result of teachers' involvement in an intervention study is one area that has yet to be fully examined, which suggests that further research is needed to better understand teacher learning and teacher change within different contexts outside the Pro-D realm.

The present study has aimed to address the current lack of studies of teacher change outside of the Pro-D context by investigating factors that influence teacher change in an ***intervention experiment context***. The researcher specifically focused on teachers' long-term response (change) to being actively involved in the implementation of an experimental practice in an intervention study, since currently, this topic has not been addressed in the literature.

The personal domain. This domain encompasses variables related to teachers' knowledge, beliefs, and attitudes. As previously mentioned, researchers have found that if new teacher learning experiences are consistent with teachers' existing knowledge and beliefs, then a greater likelihood exists that this coherence will lead to teachers willingness to experiment with a new innovation (Clarke & Hollingsworth, 2002; Ertmer, Ottenbreit-Leftwich, Sadik, Sendurur, & Sendurur, 2012). For example, Opfer, Pedder, and Lavicza (2011) have argued that teachers' past experiences and existing pedagogical beliefs determine not only the instructional decisions they make but also may influence what they are willing to learn. Opfer et al. (2011) found a moderate relationship between

teachers' orientation to learning that encompasses both beliefs and practices and teachers' self-reported change. The literature examining ***attitude towards change*** has found that the association and dissonance of attitudes with action also can impact the level of change observed (Coleman, Gibson, Cotten, Howell-Moroney, & Stringer, 2015; Zhang & Han, 2008). For example, Zhang and Han (2008) have discussed how even though attitude is not an action, attitude influences the direction and drive of people's reaction to change. Furthermore, Wagner and French (2010) have found that teachers who demonstrated an ***intrinsic interest*** and higher levels of ***motivation*** to seek out and embrace opportunities for change were more likely to adopt a new innovation into their practice.

When focusing on teachers' use of technology in the classroom, Judson (2006) and Hermans et al. (2008) have found that teachers who held more traditional pedagogical beliefs were less likely to integrate the use of computers into their lessons, whereas teachers with more constructivist beliefs were found to implement more student-centered (higher-level) use of technology. Mueller, Wood, Willoughby, Ross, and Specht (2008) and Cullen and Green (2011) have found that teacher attitudes were a strong predictor for their motivation to use technology and how frequently they integrated technology in their practice. More recently, Coleman et al. (2015) also found that teacher attitudes and anxiety regarding computers impacted how they integrated technology-based innovations into their classrooms.

Bandura's work on ***self-efficacy*** often is explored when attempting to develop an understanding of teachers' technology use. Bandura (1997) has stated that self-efficacy is what a person believes about her/himself and her/his ability to persevere in a situation or accomplish a task. In the literature, Mueller et al. (2008) have found that teachers with low self-efficacy in their own teaching abilities and knowledge of computers were less willing to take on new computer-based innovations, whereas Wheatley (2005) has found that teachers' low self-efficacy could be a motivating factor for change. Wheatley's findings suggest that teachers' self-doubt in their existing abilities may lead to reflection, which in turn may motivate them to learn about new innovations so as to improve (change) their teaching skills and practice.

Clarke and Hollingsworth (2002) have concluded that change is evident in the personal domain as teachers increase the value that they attach to a new teaching strategy. These authors have argued that this increased attachment (change) may be a reflection of new pedagogical knowledge for teachers. Garet, Porter, Desimone, Birman, and Yoon (2001) and Lawless and Pellegrino (2007) have suggested that learning opportunities that encourage teachers to reflect on their existing beliefs and knowledge, and to share their reflections with colleagues may facilitate teachers' willingness to embrace new innovations. Thus, researchers need to ask themselves: *What does it take to change teachers' attitudes and beliefs when it comes to adopting new innovations reflecting educational evidence-based practices*? Since such findings suggest that if the goal is to successfully influence teacher change over the long-term, researchers need to consider all of the variables that might influence teachers' personal domain (i.e., pedagogical beliefs, existing knowledge, self-efficacy, and motivation).

The domain of practice. Teachers' classroom practices and professional experimentation, such as experimenting with a new teaching strategy, reside in the domain of practice. Frank, Zhao, Penuel, Ellefson, and Porter (2011) have suggested that for new innovations to be adopted, teachers need sufficient time to experiment and explore (what the authors refer to as *fiddle*) to develop the necessary knowledge of the ways by which they could effectively modify and improve their practice with the new innovation they just acquired. Again, this supports the value of the duration of exposure to new knowledge being a critical variable of how effective new teacher training may be (Desimone, 2009). Many Pro-D workshops are a one-shot training session on a new teaching technique or program that may be very effective if utilized, but without teachers getting an opportunity to experiment with the innovation shortly after being trained, researchers have found that the likelihood that the innovation eventually will be implemented is unlikely. Desimone (2009) has suggested that no exact *tipping point* exists for the amount of experimentation time that teachers need to acquire the new knowledge. However, Desimone also cites research that shows support for at least 20 hours of contact time, spread over a semester, and during which teachers receive some follow-up support and feedback on their implementation efforts. Further, Ertmer and Ottenbreit-Leftwich (2010) have asserted that researchers can support teachers so that their personal experiences using an innovation are successful. For example, Mueller et al. (2008) have noted that teachers—who initially may be hesitant about integrating technology into their lessons—may, after a number of positive experiences using computers with their class, change their original beliefs and confidence in the potential of technology as an instructional tool.

The domain of consequence. The domain of consequence involves the influence of ***salient outcomes*** on teachers' use of a new innovation in their practice. Clarke and Hollingsworth (2002) have stated that the significance of the designation *salient outcomes* lies in the need to acknowledge that individuals (teachers) value and consequently attend to different things they consider of value. These authors have suggested that change in the domain of consequence is firmly tied to teachers' existing value systems and to the inferences they draw from their efforts to implement teaching practices in the classroom. For example, when implementing a more constructivist, interactive approach, one teacher may view the increased student-student dialogue as a favourable consequence, whereas another teacher may construe the same outcome of increased verbal engagement among students as a sign of loss of control, and thus interpret the use of the new strategy as a failure and subsequently discontinue its use. Barnes (2005) has found that teachers' decisions to either maintain or alter their current curriculum was influenced by a factor he termed ***flagging student interest***. Teachers would often initiate a change to their practice that would be tailored to their students' needs so to increase their students' interest in a subject matter. Barnes noted that it was the students' enthusiasm for the implemented changes to the curriculum that provided the impetus for teachers to continue with the change process and to adapt the changes to their practice long-term. This finding suggests that teachers must value the outcomes of using the new innovation with their students so to continue the long-term use of the innovation. Other examples of possible outcomes that teachers may value are improvements in student learning outcomes, increased student interactions, student-centered activities, teacher-centered activities, classroom management, and/or addressing provincially mandated learning outcomes (e.g., Desimone, 2009; Wagner & French, 2010).

Applying the domains of teacher change to research findings. When examining teacher change in the context of integrating technology into the classroom, researchers have aimed at identifying the variables that discriminate between teachers who do and do not integrate computer-based information and communication technologies (ICT) into their practice. For example, Muller, Wood, Willoughby, Ross, and Specht (2008) surveyed a random sample of teachers (elementary $n = 185$ and secondary $n = 204$) to provide a comprehensive summary of teacher-level variables that would best discriminate between teachers who integrate ICT use in the classroom and those who do not. Using a discriminant function analysis (DFA), Mueller et al. (2008) identified seven variables for elementary teachers and six for secondary teachers (accounting for 74% and 68% of the variance, respectively) that discriminated between high and low integrators. The variables included positive teaching experiences with computers (*domain of practice*); beliefs supporting the use of computers as an instructional tool, motivation, and teaching efficacy (*personal domain*); and training and support (*external domain*). A review by Wozney, Venkatesh, and Abrami (2006) also found support for how teachers' own experiences and perceived competence in using technology (*personal domain*) influences their knowledge and attitudes (*personal domain*) regarding even considering employing technical innovations to their standard teaching practices.

Ertmer (1999; 2012) has distinguished between two sets of barriers—first-order and second-order—that impact the uptake and implementation of technology-based resources into teaching practice. First-order barriers, or external barriers, refer to obstacles that are generally external to teachers, factors such as lack of access to computer and Internet resources, time (*domain of practice*), training, and support (*external domain*). Second-order barriers, or internal barriers, are intrinsic to teachers, such as teaching attitudes and beliefs; beliefs in the value of technology to the teaching/learning process; teachers' knowledge, skills, and confidence about how children learn and in their own ability to integrate technology into their teaching practice (*personal domain*). In his studies, Ertmer (1999; 2012) found support for second-order barriers, especially teachers' own beliefs and attitudes regarding the relevance of technology to students' learning as having the greatest impact on the success of technology integration into teachers' existing practice.

Whereas in their analysis of 48 empirical studies (from 1995 to 2006), Hew and Brush (2007) found that the three most commonly cited barriers influencing teachers' integration of new technology innovations were: 1) access to resources, 2) teachers' knowledge and skills, and 3) teachers' attitudes and beliefs. While other researchers, such as Guskey (1986) and Thomas, Barab, and Tuzun (2009), have also found that teachers are more likely to implement an innovation that has been used by a colleague in their school and has demonstrated some success (*domain of practice*). The first-hand account from a fellow teacher (*domain of practice*) and seeing for themselves how the tool is implemented and how children in their school respond to the resource *(domain of consequences*) can influence a teacher's interest (*personal domain*) regarding trying the innovation themselves, especially if they feel supported and confident enough to use the resource themselves. Rogers (2003) has referred to this factor as the *trialability* of the innovation (*domain of practice*), which is the extent to which the teachers can try out the innovation with support, or watch the resource being successfully being used by earlier adopters. Specifically, when it comes to the introduction of a new ICT resources, Granger, Morbey, Lotherington, Owston, and Wideman (2002) have suggested that the quality of training and support for assisting teachers to transfer their newly acquired knowledge of a new resource (*external domain*) to actual classroom implementation is essential. Pyle (2011) has suggested that while *best practices* and *effectiveness data* that have empirical support may convince researchers of the value of adopting a new innovation, the actual impact on classroom practice and student learning trajectories are more likely to validate the practical success of a new program for teachers (*domain of consequences*). Even though a number of potential teacher-level factors have been identified as affecting teachers' use of ICT, teacher-level factors determining the quality of the implementation of computer-based intervention programs have been examined less commonly in detail (see Mueller, Wood, Willoughby, et al., 2008; Wood, Specht, Willoughby, & Mueller, 2008).

Thomas, Barab, and Tuzun (2009) have found that teachers were more likely to implement an innovation into their practice if the following four factors were present. First, the innovation aligns well with teacher's pre-existing curricular goals and pedagogical beliefs (*personal domain*). For instance, a goodness of fit exists between the technical resource and the teacher's current practice and beliefs regarding appropriate classroom activities, which confirms that teachers are more motivated to implement a technical resource if they can see how it fits with what they are already doing in their classroom. Second, the flexibility and the customizability of the innovation (*domain of practice*) is important, since this aids in facilitating the teachers' ability to create this *fit* for themselves. Third, if teachers felt that their students actually enjoyed the implementation (*domain of consequences*) of the technology and working with the innovation, they were more likely to implement it into their practice. Thomas et al. (2009) found that a teacher's level of implementation of a new educational innovation benefited from their students' enthusiasm and the ensuing pressure they put on them to continue using the innovation, which suggests that students' enthusiastic response to an innovation plays a role in influencing a teacher's motivation for long-term implementation (R*t*I). On the other hand, if the innovation was not received well by students, it surely would have fallen flat with the teachers in turn. Thomas et al. (2009) have contended that this factor "highlights the notion of fun and centralizes the role of the students in the implementation of innovations designed to serve them" (p. 146). Finally, regarding the adoption of new innovations, teachers stressed that appropriate support (*external domain*) with the implementation of the innovation also was essential. The value of Thomas, Barab, and Tuzun's (2009) findings is that they illustrate how the four domains of teacher change influence could be used to identify and classify practical factors when examining why teachers may or may not adopt a new innovation into their practice. However, these findings were based on a qualitative case study examining the implementation levels of only three teachers. Given the small sample, further studies with larger sample sizes need to be carried out before such findings should be considered generalizable.

In conclusion, the potential teacher-change factors discussed in the previous overview do not represent a conclusive list; rather, they are a sample of some of the key correlational relationships that researchers have proposed influence teacher change. With respect to R*t*I, the value of the identification of teacher-level factors impacting R*t*I research is that knowledge of such factors can be used to bring greater efficiency to the processes of implementing long-term pedagogical changes. Understanding ways to improve teachers' learning and ways to provide effective training and support is critical to better understand factors that may influence the success or failure of the adoption of effective innovations. Powell, Diamond, Burchinal, and Koehler (2010) have highlighted that the basic premise of studies on literacy-focused Pro-D interventions is that they can lead to improvements in teacher quality that in turn can lead to greater gains in the learning outcomes of children. Desimone (2009) has suggested that an effort to improve the quality of teacher learning is one of the keys to improving the quality of instruction students are receiving in school. For example, when conducting intervention studies, it may be valuable for researchers to apply teacher change models to better examine which potential variables influence teachers' adoption of the intervention strategy.

Next, the value of using theoretical models to understand how teacher-level factors may impact teachers' implementation of new innovations, as well as their response to being part of an intervention study, are explored. The four domains outlined in the present section are used in the next section to examine models of teacher change, thus, moving from a description of the four contexts in which change occurs to the relationships (trajectory) between these domains. However, before reviewing the three teacher change *trajectory models* that the present study examined, a review is presented of how teacher change has been measured in intervention studies. Generally, teacher change measurement models are based on what motivates teachers to change their practice and ultimately incorporate a new innovation into their existing literacy practice. By applying an R*t*I lens, an examination is carried out as to how researchers of technology intervention studies could apply these models to measure teachers' R*t*I and so better understand how best to support and facilitate teachers' responses to change.

Measuring Teacher Change in Intervention Studies

In effectiveness research studies, rigorous investigations of factors that have influenced teachers' implementation of interventions and how well these interventions have been implemented are rarely reported in detail, since they are time-consuming and more expensive to conduct (Savage et al., 2013). Thus, a need exists for more research in this area to understand better not only what factors influence a teacher's willingness to try new innovations and how well these innovations are implemented, but also what factors have a bearing on a teacher's subsequent successful long-term implementation of these innovations into their practice.

Savage et al. (2013) have pointed out that such investigations are particularly rare with respect to teachers' *fidelity to treatment implementation* of computer-based interventions. For example, although a recent systematic review by Campuzano, Dynarski, Agodini, Rall, and Pendeleton (2009) found no significant effects of technology on literacy attainment, this study failed to provide any assessment of teachers' levels of treatment integrity. Without a detailed examination of the fidelity of implementation, a thorough evaluation of the effectiveness of technology-based reading interventions is limited. Not examining the variation in teachers' delivery of interventions is of concern, since researchers have found that the quality of the implementation of non-technology interventions (e.g., Davidson et al., 2009; Stein et al., 2008) and interventions involving technology (e.g., Chambers et al., 2008; Savage et al., 2013; Savage et al., 2010; Wolgemuth et al., 2013) can significantly affect student-level outcomes. Furthermore, data on teachers' quality of implementation is of value, since it can be combined with teachers' self-reported data of potential R*t*I factors that have been identified as influencing teachers learning, change, and adoption of new innovations.

Measuring fidelity of implementation. One method for examining teachers' levels of implementation of technology-based programs is the use of formal models (Mueller, Wood, & Willoughby, 2008). For example, Wozney, Venkatesh, and Abrami (2006) use expectancy theory in their model to explain how a teacher's decision to use a new technical innovation is impacted by three motivational categories: (1) the perceived value of using technology, (2) the perceived expectancy of the success of the innovation on desired outcomes, and (3) how high the costs of using the technical innovation are perceived to be. That is to say, based on the model of Wozney et al. (2006), a new technical innovation is more likely to be integrated into a teacher's existing practice if the perceived value of the innovation and the expectation of success are high, and combined their perceived benefits are not outweighed by the perceived associated costs of implementation.

Sandholtz, Ringstaff, and Dwyer (1997) have provided another model of instructional change that has five stages of technology implementation: ***entry, adoption, adaptation, appropriation, and invention***. Based on their Apple Classroom of Tomorrow (ACOT) research, Sandholtz et al. found that teachers' use of technology progresses through these distinct stages of development as their comfort level and use of technology becomes more integrated into their teaching practice. Sandholtz, Ringstaff, and Dwyer's model also suggests that a change occurs in a teacher's beliefs about effective teaching practices and the value of technology as part of their pedagogy as teachers move through these stages. For example, Matzen and Edmunds (2007) have pointed out that with respect to the use of technology, initially a more traditional teacher-centered approach to instruction often is reinforced (e.g., in the *entry* and *adoption* stages of implementation), whereas as teachers' comfort level with technology increases, they integrate more student-centered learning opportunities into their lessons (e.g., in the *adaptation* stage).

A study by Savage et al. (2010) empirically examined the variations in teachers' utilization of technology by creating a measurement tool that incorporated distinct levels of Sandholtz et al. (1997) *technology integration model* (entry, adoption, and adaptation). Teachers' use of technology at the *entry stage* is characterised by teachers grappling with technical problems, making time-consuming mistakes, which result in a limited actual use of technology in their lessons and often frustration and high levels of discontinuation. Teachers at the *adoption stage* begin to apply technology in a systematic manner, although with minimum experimentation. At this stage, the method of teaching is more teacher-centered, and opportunities for other types of learning—such as collaborative or experiential—are limited. Teachers at the *adaptation stage* appear more self-confident about their ability to effectively use technology. At this stage, teachers tend to experiment more with different ways to integrate technology into their lessons, often with greater connectivity between all forms of learning, which helps to transform their classroom teaching and learning. In their examination of teachers' use of an innovative reading technology (ABRACADABRA), Savage et al. (2010) found evidence for each of these three stages of technology implementation. They found the *Adaptive* group demonstrated significantly higher performance than the control group in phonological blending ($t = 3.68$, $p < 0.01$), letter-sound knowledge ($t = 2.15$, $p < 0.05$), and phoneme segmentation fluency ($t = 1.69$, $p < 0.05$). Also, their study found that the *Adoptive* group differed significantly from the control group in phonological blending ($t = 2.23$, $p < 0.05$). These findings show that students who were in the classrooms taught by teachers rated as integrating reading technology at the *Adaptation* level had significantly greater gains in reading.

As mentioned previously, the data obtained from the implementation fidelity measures has provided valuable information about teachers' quality of implementation. However, when researchers are able to analyze such data in conjunction with teachers' self-reported data of potential R*t*I influencing factors, a more comprehensive picture may emerge regarding how best to facilitate teacher change.

Examining Teacher Change Models

Designing learning opportunities for teachers to explore new innovations requires researchers to consider several variables, including teachers' knowledge, beliefs, and attitudes about their own learning process. One way to examine teacher change is to follow-up with teachers and have them reflect on their beliefs and experiences concerning their learning processes so to better understand what factors played a role in influencing them to incorporate (or not incorporate) newly acquired knowledge into their existing curriculums (Opfer et al., 2011). Using theoretical models to assist teachers in reflecting on their learning processes can be an effective way to initiate dialogue with teachers, and have teachers share their thoughts and unique experiences (Clarke & Hollingsworth, 2002; Creswell, 2009; Gibbs, 2003).

As discussed earlier, four analytic domains of teacher change can be viewed as being universal for describing the four main contexts in which teacher change occurs (i.e., the ***external domain, personal domain***, ***domain of practice***, and the ***domain of consequence***) (Clarke & Hollingsworth, 2002). However, differences exist in how these four components relate to each other. Thus, this chapter next provides a brief overview of three conceptual models representing the different *trajectories* of teacher change. In these three trajectory models, the terminology used to discuss the four conceptual domains of teacher change may not be identical, although the four domain components are still considered to be analogous.

Understanding the possible conceptual trajectories of teacher change is of value because in spite of the universality of these four domains of teacher change (Desimone, 2009), teachers' learning experiences are unique, and variations will occur from teacher to teacher regarding how they interpret the impact of each domain on their own process of change. For example, teachers' views may differ on the order of the influence of the four domains, where they attribute change occurring first, which variables are impacting change within each of the four domains, and on how (or even if) they view the change in one domain influencing change in another. However, by referring to the different theoretical models of possible trajectories of teacher change, researchers who are examining the factors that influence teacher-level R*t*I can more effectively explore the variations in teachers' responses.

It also should be noted that even though the following models of professional growth examine *teacher change* in the context of teachers taking on a new innovation, for the purpose of the present study, this notion of teacher change is reframed. The teacher-level R*t*I that this study examined did not focus on whether teachers chose to take on a new and novel technological innovation, but rather on whether *teachers continued to use an intervention* deemed effective (e.g., ABRA) after their participation in the intervention study concluded. Therefore, the objective of such an examination was to investigate whether being part of an intervention study in itself played a role in a teacher choosing to change their practice long-term. Such an examination is of benefit, since an investigation of teachers' responses to being part of a long-term classroom level (Tier 1) intervention study and the possible variables that influence teacher change during such a process has yet to be adequately explored in the R*t*I literature (Hill et al., 2012). Thus, by examining the teacher change trajectory models in an R*t*I context, a teacher-level R*t*I analysis will be valuable for identifying potential factors that may facilitate a teacher's long-term use of an evidence-based practice deemed to be effective.

Model A: Implicit model of the purpose of teacher professional development. Fullan (2005) has argued that to be of interest, any new innovations that staff development programs want teachers to employ should offer practical suggestions that can be efficiently used by teachers to enhance the desired learning outcomes for their students. With this end goal in mind, an implicit *model of the purpose of teachers' professional development* suggests that many Pro-D programs are based on the hypothesis that teacher professional change results from changing the beliefs and attitudes of the teachers first. This proposed implicit trajectory model reflects the assumptions of the purpose of most traditional professional sessions: that is, if (1) teachers experience effective professional development (*external domain*), then (2) the professional development increases teachers' knowledge and skills, and/or changes in teachers' attitudes and perceptions are made (*personal domain*), then (3) this will subsequently lead to changes in classroom practices and behaviours, since teachers will use their new knowledge and skills to develop the content of their instruction and/or improve their approach to pedagogy (*domain of practice*), which in turn will lead to (4) the ultimate goal of growth in student outcomes, since the instructional changes fostered in the domain of practice will increase student learning (*domain of consequences*) (Clarke & Hollingsworth, 2002; Desimone, 2009) (see Figure 1, an implicit model of the purpose of teacher professional development).

When applying the assumptions of this simplified model to how teachers may implement a new innovation into their practice, the assumed trajectory of teacher change is quite straightforward. For example, this model predicts that if teachers obtain adequate training and information on, for example, a new technological innovation (*external domain*), this new information may change their opinion about the value of technology in the classroom and how they think about delivering a particular language arts lesson (*personal domain*). The model suggests that such changes in teachers' attitudes and perceptions about the usefulness of technology will lead to specific changes in their classroom behaviours and practices (i.e., an increased use of technology) (*domain of practice*), which, in turn, will lead to improved student learning (*domain of consequence*).

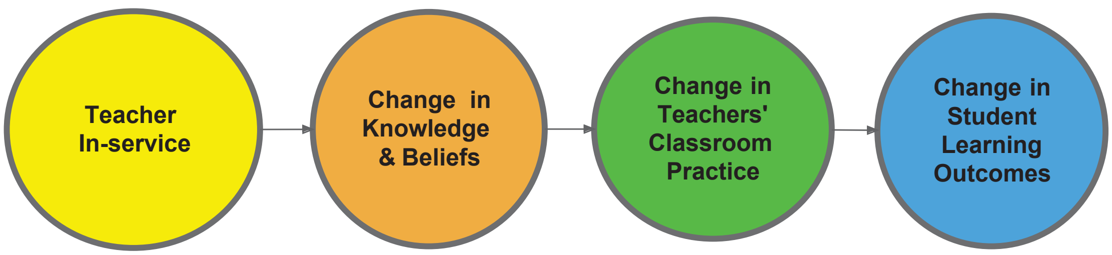


Figure 1. Model A: An implicit model of the purpose of teacher professional development (Clarke & Hollingsorth, 2002, p. 949).

Critics of this model have questioned its simplistic nature and projected linear trajectory. For example, Fullan (1982, 2005) has argued that the causal sequence of events, on which many professional teacher in-service programs are based, are misleading, since they fail to consider the complex process of teacher change. Guskey (2002) also has suggested that teachers' change in values rarely occurs after an introduction to a professional development session, so the assumptions of Model A are misleading. More recently, Desimone (2009), who favours the proposed sequential order of the domains of change in Model A, has argued that the flow of influence between the domains should not be viewed as unidirectional— rather, the links between the domains are reciprocal in nature. For example, Desimone's model suggests that change observed in the end goal (improved student learning outcomes) will lead to reciprocal changes through the other three domains (see Figure 2).

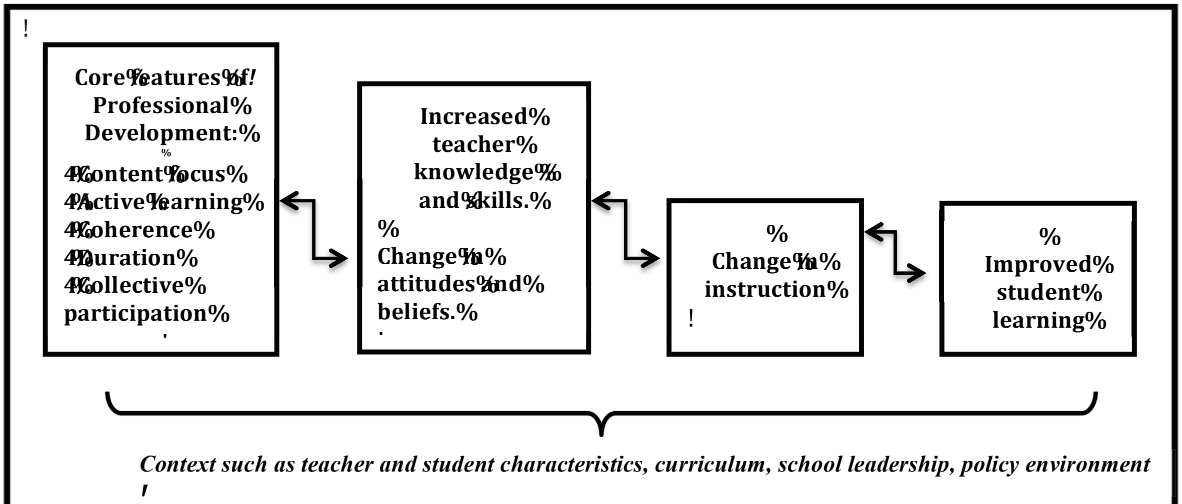


Figure 2. Desimone (2009) proposed core conceptual framework for studying the effects of professional development on teachers and students.

Further, Desimone’s (2009) suggested model includes *context* as an important mediator and moderator of change in all four domains, which is a factor neglected in Model A. Desimone has argued that context—such as teacher and student characteristics, curriculum, classroom environment, school leadership, and policy conditions at multiple levels—can influence the degree of change observed in all four domains of change. Thus, Desimone’s updated model represents more interactive and reciprocal relations, although the relationships between the critical domains of professional development and teacher learning continue to remain linear in nature.

Other researchers like Guskey (1986, 2000) and Clarke and Hollingsworth (2002) have questioned the order of the three key outcomes of the *model of the purpose of teacher professional development*—change in teachers beliefs and attitudes, change in the classroom practices of teachers, and change in the learning outcomes of students. Specifically, these authors have questioned the proposed temporal sequence in which changes in these three domains of teacher change occur. Subsequently, Guskey (1986, 2000) and Clarke and Hollingsworth (2002) have provided alternative trajectory models of teacher change, which are reviewed in the next section.

Model B: Model of the process of teacher change. Guskey (1986) has proposed an alternative teacher change model that challenges the assumptions and trajectory sequence of the *model of the purpose of teacher's professional development* (Model A) and Desimone's (2009) updated conceptual model. *Guskey's model of the process of teacher change* predicts that changes in teachers' beliefs and attitudes (*personal domain*) are likely to occur after teachers observe improved changes in student learning outcomes (*domain of consequence*), but not before (see Figure 3). This model suggests that teachers are more likely to embrace a new strategy or proposed curriculum change (*external domain*) after they have field-tested it in their classrooms (*domain of practice*) and have personally seen positive gains in student attainment. Thus, teachers need evidence that a new innovation or practice is effective before significant changes in their attitudes and beliefs will occur. Thus, Guskey has suggested that it is not teachers' involvement in a professional development experience per se, but rather the experience of the successful implementation of a new innovation that ultimately leads to changes in their personal domain.

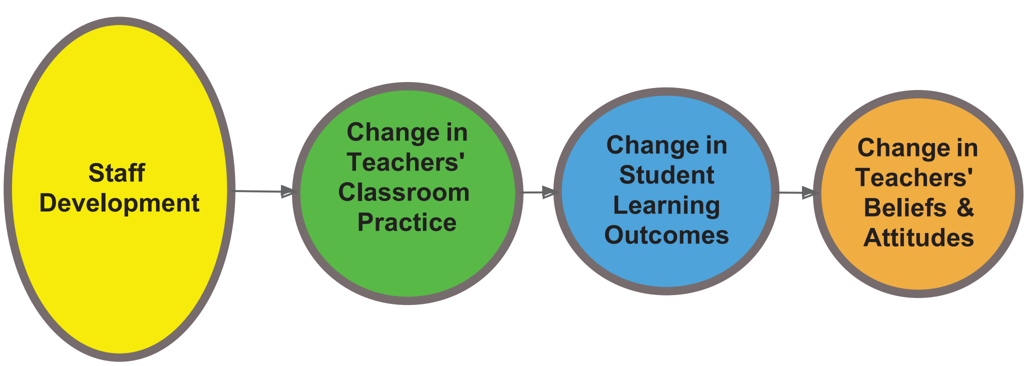


Figure 3. Model B: Guskey's model of the process of teacher change (Guskey, 1986, p.7).

However, as with Model A, Guskey's *trajectory model of teacher change* (Model B) also has been criticized for its representation of teacher change as a linear process (Clarke & Hollingsworth, 2002; Opfer et al., 2011). The authors have contended that a linear representation of teacher change neglects to consider the complex nature of the teacher learning process. Further, these authors have argued that teacher change models need to acknowledge that multiple pathways of influence exist between the domains, which they contend is a better representation of the individuality of teacher growth. However, Opfer et al. (2011) have suggested that Guskey's model fails to illustrate that the relationships between the domains of change often are reciprocal with changes in one domain being contingent on changes in another. Moreover, Opfer et al. have argued that the process of teacher learning and change does not have to result from a linear process that begins with a professional development experience, but rather that the process of teacher change can be initiated at any point in the change process (i.e., change in students, change in beliefs, or change in practice). It should be noted that Guskey (1986) has acknowledged the reciprocal nature of the outcomes from professional development sessions, although he has asserted that considering the *order* of outcomes is necessary and has value.

In summary, while Desimone (2009) has contended that change in beliefs leads to change in practice that leads to change in students, Guskey (1986) has asserted that change in practice leads to a change in students that leads to change in beliefs, which clearly illustrates that an apparent disagreement exists between these models concerning the order in which the change sequence of the domains occurs. Although research has shown support for a reciprocal relationship between attitude and action (Zhang & Han, 2008), the difference in the trajectory of influence of the domains depicted in Model A and Model B has fostered a debate between researchers with respect to whether attitude leads to action or whether action leads to a change in attitude. By employing these two linear models (Model A and Model B), it would be of interest to examine what teachers report as experiencing first—a change in beliefs or an observed change in student outcomes—before they decide to incorporate a new innovation into their practice long-term. Further, an examination of teachers' responses to change could also explore the ways in which teachers provide a fuller picture of the complex nature of teacher change. Through teachers' reflections on their teacher change process, we may see models emerging more reflective of Clarke and Hollingsworth's (2002) nonlinear *interconnected model of professional growth* (Model C).

Model C: Interconnected model of professional growth. Unlike Model A and B, Clarke and Hollingsworth's (2002) *interconnected model of teacher professional growth* allows for multiple growth pathways between the four domains that are neither linear or unidirectional, which is a more complex representation that takes into account the complex nature of teachers' professional growth and reasons for change.

As previously mentioned, the four domains of this model are analogous (but not identical) to the four domains identified by Guskey (1986) and Desimone (2009). These domains include the **personal domain** (teacher knowledge, beliefs, and attitudes), the **domain of practice** (professional experimentation), the **domain of consequence** (salient outcomes), and the **external domain** (sources of information, stimulus, or support). Clarke and Hollingsworth's (2002) interconnected model of teacher professional growth expands on Guskey's (1986) linear model, since it provides a more comprehensive picture of teacher change. Clarke and Hollingsworth have contended that their nonlinear teacher change process model more accurately represents change that is continuously occurring through the mediating processes of *reflection* and *enaction*. The term *enaction* is used in the model to differentiate between the "translation of a belief or a pedagogical model into action" and the simple deed of *acting*, since the authors have contended that acting occurs in the domain of practice, in which each action represents the enactment of something a teacher has knowledge of, believes in, or has experienced (p. 950). Thus, Clarke and Hollingsworth (2002) have suggested that it is through the processes of reflection and enaction that change in one domain leads to change in another (see Figure 4).

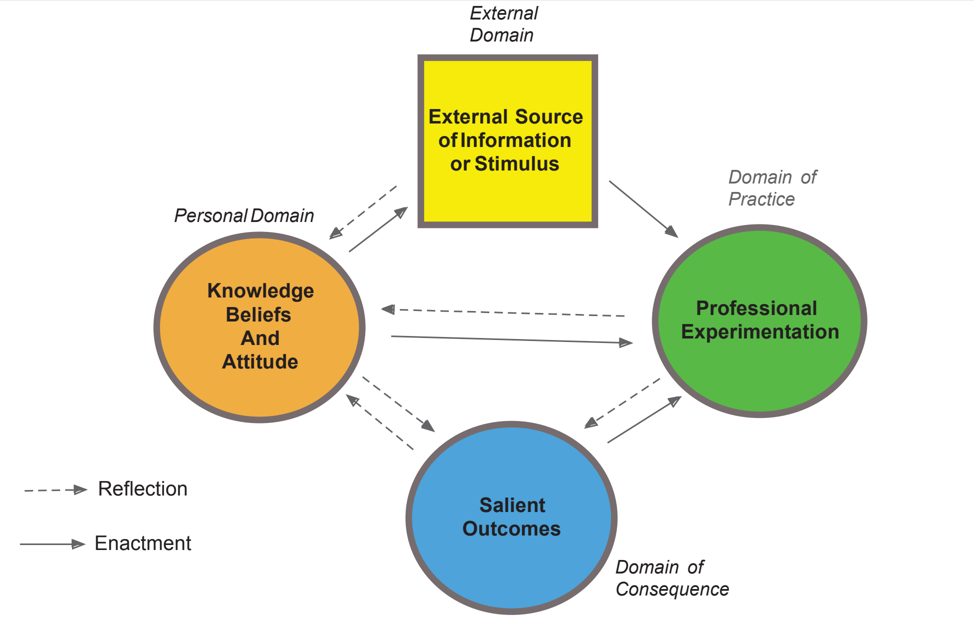


Figure 4. Model C: Clarke & Hollingsworth's (2002) interconnected model of proessional growth (p. 951).

Using the computer-based ABRA study as an example (see Savage et al., 2013), some of the key variables of the interconnected model of professional growth are described further. Beginning with the personal domain, possible ABRA teacher-level R*t*I mediating factors within this domain include: ABRA teachers' knowledge, beliefs, and attitudes about using technology in the classroom; teachers' thoughts on effectiveness of ABRA as a literacy tool; attitudes towards using computer activities rather than paper and pen or working with actual books; experience and comfort level with computers; value of technology. Clarke and Hollingsworth (2002) found that "change in teacher beliefs and attitudes were evident in the increasing value that the teacher attached to the new teaching strategies that represented in themselves new pedagogical knowledge for that teacher"

(p. 953). Finding change in this domain may be the key to identifying those teachers that not merely adopted ABRA but also adapted and integrated ABRA into their literacy practice and continued to use ABRA after the conclusion of the original study.

The *external domain* has many sources of information, stimulus, and support for new knowledge that may facilitate and/or hinder teachers' access and adoption of new innovations into their practice. As with the previous linear models, this domain encompasses professional development opportunities and factors that contribute to the effectiveness of those PD sessions (e.g., the duration of exposure to the new innovation, the interactive nature and hands-on experience of the training sessions, the relevance of the content to teachers' existing practice, and the level of ongoing support teachers receive during the early stages of implementation). With respect to teachers involvement in the ABRA intervention study, some possible teacher-level R*t*I mediating factors contributing to teacher change in this domain could include how teachers felt about the ABRA training workshops and the in-class support they received with lesson planning, the implementation of ABRA, and external help in trouble-shooting during the early stages of implementation.

The *domain of practice* involves the teachers' professional experimentation with the ABRA tool. For both the ABRA and control classroom teachers, this experimentation occurred during the training sessions. ABRA teachers also were able to work with and explore the ABRA program during the 3 months they were implementing the tool. Perhaps, teachers who continued to use ABRA after their involvement in the intervention study ended may have done so because they made a connection between their personal beliefs about the value of technology and their domain of practice, since they were able to effectively implement ABRA into their ELA lessons without too many difficulties and witnessed the benefits of using this technical innovation in their class. Other teachers who encountered obstacles (i.e., lack of computer time, poor Internet access, etc.) when trying to implement effective ELA lessons with ABRA may have discontinued using ABRA after the study ended (Wood et al., 2008).

For many North American school districts, many external barriers, such as access to computers and technical support, have improved immensely over the last decade (Ertmer et al., 2012). However, school contextual factors, such as the level of support provided (i.e., by principles, district literacy coordinators, technical support, etc.) could still have played a role for the ABRA teachers by ensuring that they had access to working computers and efficient Internet service during their ELA lessons. Thus, these factors could still be critical variables in the domain of practice, and they have had an impact on the ABRA teachers' level of implementation.

The *domain of consequence* involves the influence of *salient outcomes* on teachers' R*t*I. Clarke and Hollingsworth (2002) have suggested that the significance of the designation salient outcomes lies in the need to acknowledge that individuals (teachers) value and consequently attend to different things they consider of value. With respect to this domain, possible factors that ABRA teachers may reveal that they consider as salient outcomes that impacted their use of ABRA could include: students' response to ABRA; students' engagement and motivation during ABRA-ELA lessons; teacher's control over class during ABRA lessons; and shown improvement in students' development of literacy skills. Stein, McRobbie, and Ginns (1999) have contended that before teachers fully embrace new technology innovations, they need to experience the value of the resources through changes in student learning. Clarke and Hollingsworth (2002) have stated that change in the domain of consequence is firmly tied to a teacher's existing value system and to the inferences that the teacher draws from the practices of the classroom, which suggests that ABRA teachers must value the outcomes of using the ABRA tool with their students to be motivated to continue its long-term use. Further, experimentation and reflection may have motivated some teachers to move from *adopting* ABRA into their lessons plans early on (i.e., using just ABRA phonics activities to reinforce phonics lessons) during the intervention phase to an *adaptation* level (i.e., integrating higher skill activities, such as comprehension activities into ABRA lessons, or including cross-curriculum extension activities that were linked to the ABRA lesson) as they became more comfortable with the tool and witnessed positive outcomes. In turn, this comfort level with the technology and the positive student outcomes may have motivated some teachers to continue to use ABRA after the end of their involvement with the Pan-Canadian study.

These hypothesized examples reflect a few of the many possible factors that teachers may identify as influencing their acceptance and long-term adoption of a technical innovation into their regular practice. Further, as Chen (2008) and Savage, Erten, et al. (2010) have contended, a need still exists for additional research on possible factors (e.g., teacher-related and contextual) that may influence not only teachers' long-term adoption of a new innovation but also how effectively that educational innovation is implemented.

Application of Teacher Change Models to ABRA Teachers R*t*I Analysis

At the teacher level of analysis, the objective was to examine R*t*I data on ABRA teachers' ongoing use of the ABRA tool. Direct and open-ended questions were used to inquire whether teachers responded favourably to being part of the ABRA study, and subsequently incorporated ABRA into their ELA lessons and continued to use the tool a number of years later. The central question probed in the follow-up concerned the continued use of the intervention (e.g., *Do you continue to use ABRA in your ELA lessons? If so, in what capacity? If not, could you elaborate on why not?*). In addition, data collected from the teachers' responses to the follow-up questions were used to explore the different predictions of the three teacher change models:

1. *Model of the purpose of teacher professional development (**Model A**)*
2. *Guskey's model of the process of teacher change (**Model B**)*
3. *The interconnected model of teacher professional growth (**Model C**)*

Specifically, questions were asked to find out when during the process of implementing a new resource do teachers tend to begin believing in the value of the resource. For example, did the ABRA teachers believe in the value of the new resource *before* implementing it (Model A), or did they believe in the value of the new resource *after* seeing its benefits on the desired salient outcomes (Model B)?

In conclusion, studies identifying factors and establishing models of teacher change have been based, for the most part, on teachers' involvement in traditional professional development workshops. These earlier teacher change studies have provided a strong basis for developing hypotheses about which key variables influence the effectiveness of professional development experiences on teachers (e.g., Penuel et al., 2007). However, little of the current teacher change literature comes from teachers' involvement in experiment studies examining the effectiveness of classroom level (Tier 1) intervention. In the original ABRA study, ABRA was introduced as a supplement to teachers' exisiting Tier 1 classoom-level early literacy instruction (Savage et al., 2013). Thus, the present study aimed to examine whether similar findings would arise when looking at teachers not involved in a professional development workshop but rather in a longer supported intervention study.

Early literacy interventions that are effective are those that are employed by the teacher. So if an innovation is not being used, it cannot be effective. To gain value from intervention studies examining the long-term effectiveness of technology-based literacy innovations, it is necessary to understand the process of teachers' professional growth (Clarke & Hollingsworth, 2002).

Specifically, a need exists to identify what factors and conditions best facilitate and support teachers' incentives to modify their existing practice to integrate a new innovation. Understanding the factors that influence teacher change equips researchers with valuable knowledge that they can use when they are creating and testing the effectiveness of an early literacy innovation delivered by classroom teachers. Specifically, researchers can use this knowledge to support factors that would best facilitate teachers' change in their literacy practice. The goal of researchers should be to develop not only an effective literacy tool but also to encourage and support teachers' long-term use of a resource demonstrated to be effective.

Understanding how teacher-level R*t*I variables interacted and fell under these four domains assisted in the collection and analysis of the teacher data for the present study. The teacher change models were used as analytical tools for the categorization of the teacher data and hypothesized trajectories (Clarke & Hollingsworth, 2002). They were used as interrogatory tools to frame the theoretical and practical questions the teachers were asked concerning their experiences of being part of the ABRA study and their current use of the ABRA tool. Therefore, by quantitative and qualitative examinations of teacher responses, another goal of this follow-up study was to identify key teacher-level factors that may have been mediating teachers' long-term adoptions of a technological innovation after their involvement in an intervention study in which the innovation was proven to be effective.

Chapter 5: Research Objectives for the ABRA R*t*I Follow-Up Study

Chapter Overview

Chapter 5 provides the background and objectives for this dissertation. It begins with an overview of the original ABRACADABRA 2007–2009 Pan-Canadian Cluster RCT study. Then, it outlines the purpose of the present follow-up study, and highlights the research objectives and potential research questions that guided this dissertation.

ABRACADABRA the Tool

ABRACADABRA (hereafter referred to as ABRA) is an acronym that stands for **A** **B**alanced **R**eading **A**pproach for **Ca**nadians **D**esigned to **A**chieve **B**est **R**esults for **A**ll. ABRA is a freely accessible web-based literacy resource designed for elementary classrooms. The ABRA software is made up of four different modules: letters and sounds, reading, comprehension, and writing/spelling. ABRA has over 30 distinct types of levelled-activities and 17 stories designed for children learning to read. It also has modules for parents and teachers that contain additional resources, activity suggestion ideas, and assessment and communication tools. The modular format of ABRA was developed to enable the maximal flexible use of the software in different classrooms contexts. Thus, this flexibility enables ABRA to be tailored to meet the specific needs of individual children or used for whole-class instruction (Savage et al., 2010).

ABRA as a literacy tool was conceptualized and constructed based on best evidence findings on effective reading interventions for phonics and letter skills, reading fluency, and reading comprehension (Hipps, Abrami, Savage, Cerna, & Jorgensen, 2005; Savage, Abrami, Hipps, et al., 2009). Further, ABRA's design emphasizes a balanced literacy approach (e.g., Pressley, 2002), since the foundations for ABRA activities are drawn from reviews of effective reading practices (e.g., National Reading Panel, 2000; Rosenshine & Meister, 1994). Moreover, ABRA's literacy content, interactive interface component, and instructional design, help it to meet the key criteria in Bishop and Santoro's (2006) framework for evaluating the quality of educational software for literacy. ABRA is available free online at *http://abralite.concordia.ca.* The *ABRA lite* version contains all of the early literacy activities and stories as well as the instructional, professional development, and parent modules that can be used in the classroom or at home. A more complete version of ABRA is available (see *http://grover.concordia.ca/abracadabra/*) that must be downloaded and stored on a school board server. This complete version of ABRA also is free, and in addition to containing everything in the lite version, it also contains an assessment and a communication module for teachers (Savage et al., 2010).

Since the development of the ABRA tool, a growing body of research carried out by researchers affiliated with the Centre for the Study of Learning and Performance (CSLP, Montreal Canada) has explored the effectiveness of this tool. Studies examining the impact of ABRA on students' reading outcome measures have provided evidence for some of the equivocal issues around intervention effectiveness for ICT (see Abrami, Savage, Wade, Hipps, & Lopez, 2008; Comaskey et al., 2009; Deault et al., 2009; Di Stasio, Savage, & Abrami, 2012; Hipps et al., 2005; Savage, Abrami, Hipps, et al., 2009). The development and evaluation of ABRA has been an ongoing cyclical process to which the Pan-Canadian 2007–2009 RCT study carried out by Savage et al. (2013), the replication field study by Piquette et al. (2014), along with a recent meta-analysis summarizing the existing research on the effects of ABRA by Abrami et al. (2015), have all provided scientific support for the effectiveness of this early literacy tool in the hands of classroom teachers. The present study further adds to the current early literacy literature base by applying an R*t*I approach of analysis to examine the long-term effectiveness of this technology-based intervention tool.

Original ABRACADABRA 2007–2009 RCT Study

The Pan-Canadian ABRACADABRA (ABRA) randomized control trial (RCT) study arose out of the need for more detailed investigations of the implementation issues related to technology-based interventions, since currently, very few well-designed RCT studies have been effectively implemented (see Savage et al., 2013). The Pan-Canadian ABRA study was a cluster RCT design, since the participating classrooms, rather than individual students, were randomly assigned to the experimental conditions. This study used a pretest–post-test experimental intervention design to examine the effectiveness of ABRA. Compared to earlier research projects that examined the effectiveness of ABRA for literacy skills development (e.g., Comaskey, et al., 2009; Savage, Abrami, Hipps, et al., 2009), the Pan-Canadian ABRA RCT intervention was a large-scale study, which included 74 participating classrooms and more than 1,000 students from across the three provinces of Alberta, Ontario, and Quebec. Thus, this study was the largest non-US study of ICT and the largest true experiment (RCT) reading intervention reported to date in Canada (Savage et al., 2013).

Further, the focus of many of the earlier ABRA studies was on whether technology could be efficacious. However, ABRA is designed as a tool for teachers, and thus, the Pan-Canadian study was valuable for testing the ecological validity of the earlier studies' findings by placing the tool into the hands of well-trained classroom teachers to measure ABRA's effectiveness in the classroom as a Tier 1 level of intervention. Therefore, ABRA was introduced as a supplement to the intervention teachers' exisiting Tier 1 classoom-level literacy programs. As discussed in earlier chapters, extensive research into literacy acquisition has provided a wealth of information regarding effective means to help children experience early reading success (e.g., Justice, 2006; Richards et al., 2007). Thus, Savage et al. (2013) set out to examine the existing challenge of how to teach early literacy, and also to understand better how teachers can practically deliver various aspects of suggested effective interventions that are responsive to the unique learning needs of their students. Specifically, the 2007–2009 ABRA study had two main objectives. The first was to learn more about how teachers use technology (ABRA) in their English Language Arts (ELA) lessons, and the second was to examine if the ABRA literacy program while in the hands of the classroom teacher could produce a qualitatively significant impact on students' literacy development.

Closely following the CONSORT criteria for implementing and reporting high quality RCT studies, the 2007–2009 Pan-Canadian study had well-trained and supported teachers deliver the ABRA intervention to their students for 20 hours over 1 full semester. Post-test findings found that the ABRA intervention classroom showed significant advantages over controls in phonological blending ability (d = .42), letter-sound knowledge (d = .19), and phoneme segmentation fluency (d = .28). Further, secondary analyses exploring the effects of different levels of program implementation found that with the highest quality of implementation, advantages were evident at post-tests in phonological blending (d = .92), sight word reading (d = .43), letter-sound knowledge (d = .38), and story retell fluency (d = .22) (Savage et al., 2013). These findings suggest that at least in the short-term, ABRA in the hands of classroom teachers can be an effective literacy resource for developing key skills associated with early reading.

However, what still needs to be explored is whether the observed short-term differences are maintained over the long-term. Specifically, a closer examination of pupil R*t*I is needed to identify for which at-risk students the ABRA program was an effective Tier 1 level of intervention, and for which at-risk students additional literacy support is still required. Although great efforts were taken in the original ABRA study to try to ensure that teachers received comprehensive training and support on the use of ABRA in the classroom, a follow-up on teachers' perspective on their training, support, and additional influencing factors that impacted their use of ABRA was not conducted. Thus, a richer picture of the type of support teachers require to adopt ABRA into their long-term literacy practice also still needed.

Purpose Statement

The original 2007–2009 Pan-Canadian ABRA study was a large-scale study that provided a very rich database that has not been fully explored. The present study adds to the existing database with a collection of follow-up data and a more thorough analysis at both the pupil and teacher levels. The current study has two overarching goals: (1) to assess students' reading skills up to 4 years after their involvement in the ABRA intervention study and (2) to examine whether teachers are using ABRA up to 4 years later, in what capacity, and what factors led to this present use.

Further, an original contribution of this dissertation is that it reframes how researchers examine response-to-intervention (R*t*I). For example, this study addresses the current neglect in the literature to examine students' R*t*I to a classroom level, Tier 1, stage of intervention, given that most R*t*I studies explore student-level responses at Tier 2 and 3 stages of intervention at which students already identified as at risk are receiving supplemental support (either in a small group or one-on-one). Further, this dissertation addresses the value of simultaneously examining two levels of R*t*I—both the pupil **and** the teacher-level of response to a Tier 1 stage of intervention. An observation of both levels of response is of value whenever teachers receive training on a new innovation to implement a classroom level intervention, because ***two grains of intervention*** are simultaneously taking place. Thus, R*t*I (*change*) is occurring not only at the student level but also at the teacher level when teachers are encouraged to use a new technical innovation with their students. Accordingly, researchers conducting a study that involves the implementation of a new innovation at the classroom level (e.g., an integration of a web-based literacy program into an existing language arts program) need to start following-up on the R*t*I at both student and teacher levels to better understand possible factors that may influence the successful implementation of an intervention approach deemed to be effective. Therefore, another overarching purpose of this dissertation is to examine these two levels of R*t*I, by examining both students' and teachers' long-term responses to being involved in an ABRA literacy intervention experiment. For this reason, the definition of R*t*I in this study is interpreted as exploring not only how students respond to being involved in a Tier 1 literacy intervention study, but also how teachers respond to being involved in a Tier 1 literacy intervention study.

Research Questions

Pupil-level research questions. One goal of applying a R*t*I approach to teachers' literacy practice is to improve student reading outcomes for all students by providing them with high-quality instruction that is evidence-based (Preston et al., 2016). The delivery of well-implemented ABRA ELA lessons has empirical support documenting its effectiveness in the growth of phonological and word level skill development (Savage et al., 2010). A key objective of the present follow-up study at the pupil level is to quantitatively examine whether the short-term reading gains obtained by students, who received ABRA as part of their ***Tier 1*** instruction, were maintained 4 years later. Thus, this study adds to the existing ***effectiveness literature*** by exploring interactions between the literacy intervention program ABRACADABRA and children's initial reading levels.

Researchers such as Al Otaiba and Fuchs (2006) and Kamps et al. (2008) contend that at risk readers are a diverse group who differ in background knowledge, language ability, response to intervention, and subsequent levels of achievement. At the pupil level of analysis, the present study also empirically adds to the existing evidence-based literature on potential mediating factors that can influence the effectiveness of reading intervention, both over the long- and short-term. This addition contributes to the current movement in the R*t*I literature that claims there is a need for more early literacy interventions at the classroom level to identify those students who are at risk and who also are not responding to quality evidence-based instruction (Al Otaiba & Fuchs, 2006; D. Fuchs & Fuchs, 2006). Thus, another objective of this study is to examine the incidence of students who do and do not respond (non-responders) to effective Tier 1 intervention after having been exposed to 20 hours of a research-based literacy intervention tool. The population of non-responders is of interest because students who do not respond to the intervention are most likely those who may require special education services (Denton, 2012; Lam & McMaster, 2014). In the traditional R*t*I literature, the focus has been primarily on the influence of cognitive variables (i.e., phonological awareness, letter sound knowledge, short-term memory, IQ) as pupil-level predictors of R*t*I, and what needs to be added is an examination of a range of demographic variables (e.g., ethnicity, language, and sex) that have been discussed in the literature, but not been fully explored in R*t*I studies. Since these demographic variables have yet to be adequately examined, the analysis of demographic variables as R*t*I predictors will be more exploratory with the following research questions guiding this analysis:

Question 1. *Are there identifiable pupil and contextual demographic variables of at-risk students who had a greater response to Tier 1 intervention at T2?*

a) *Do demographic variables predict student-level RtI?*
b) *Do demographic variables predict student-level RtI above and beyond the cognitive variables (i.e., literacy ability at pretest, such as phonological awareness)?*

The guiding questions above, will not only explore possible main effects of demographic variables on student outcomes but will also provide an examination of possible interaction effects between the demographic variables of interest and the level of ABRA implementation students received. There continues to be a lack of sufficient evidence for why some students respond to interventions deemed effective while others do not (Al Otaiba, Wagner, et al., 2014; Lam & McMaster, 2014; Preston et al., 2016). Therefore, the need to look at interactions effects between the demographic variables of interest and exposure to an effective reading intervention is of merit.

Furthermore, long-term follow-up studies are generally rare, independent of whether these studies are technology intervention studies or other RCT studies (Torgerson et al., 2006). The present study aims to identify non-cognitive pupil-level mediating factors that may influence the long-term effects of evidence-based early literacy interventions for students at risk of experiencing reading difficulties. Specifically, this was done by collecting data on demographic pupil-level factors that have been discussed in the literature but have not been fully explored in RTI studies, for example: sex; ethnicity; socio-economic status (SES). Data on some of these variables were collected at the time of the original 2007—2009 study but the demographic variables were never examined as R*t*I predictors. One aim of collecting follow-up data on these variables was to examine the relationship between sociological factors and students' reading risk before and after the ABRA intervention. Further, these mediating pupil-level measures are analyzed and correlated with how well teachers implemented the intervention under investigation. Therefore examining not only possible main effects of demographic variables on student reading outcomes but also exploring possible interaction effects between the demographic variables of interest and the level of ABRA implementation students received.

In addition, at the pupil-level, the ***inoculation theoretical*** model was directly tested in the follow-up of students who received the ABRA intervention. The inoculation model is fitting for ABRA data, since phonics development has a strong focus in this web-based literacy program. The phonic-level activities are perhaps one of the most developed parts of the ABRA tool, and the primary area of ABRA where teachers tend to go to and use (as shown in earlier findings, e.g., Savage et al., 2013; Savage, Abrami, Piquette-Tomei, Wood, & Deleveaux, 2009; Savage et al., 2010). Also, studies on ABRA have found that phonics skill development has had significant effects, since the students who received an ABRA intervention were better at phonological awareness (i.e., blending words) and letter-sound knowledge in comparison to a control group (Savage, Abrami, Piquette-Tomei, et al., 2009).

According to inoculation theory, students who receive an effective Tier 1 literacy intervention that focuses on phonics development (within a balanced literacy program) should gain the necessary skills to continue to learn how to read new words on their own. Therefore, if inoculation theory is correct, the ABRA students that appeared to show a marked improvement in their phonics development following their exposure to the ABRA intervention should have been able to continue (4 years later) to use their classroom literacy experiences to become typical readers of all words and catch-up to their non-at-risk peers. Students identified at risk have baseline reading composite scores in the bottom 30% on three key literacy measures—Letter-Sound Knowledge (LSK), FRY Sight Words (FRY), and Phonological Blending Words (BW). However, after running ABRA some of these at-risk students moved into the *average* reading range, with reading scores above the 30th percentile on the BW reading measure. A goal of this present follow-up study was to determine whether these students remained average readers. Inoculation theory predicts that at-risk students who showed progress following an intervention program and caught-up to their peers will be able to maintain this progress in the long-term, whereas the insulin model asserts that a one-shot exposure to an effective intervention will not guarantee long-term success, and an at-risk child will need ongoing support. In addition, Coyne's (2004) study on effective reading interventions has found a combination result—insulin for some and inoculation for others.

Specifically, the pupil-level research questions that this study used to determine whether the T3 follow-up data collected supports the inoculation hypothesis are the following:

Question 2. *Does ABRA (the intervention) provide protection from long-term reading development problems at TIME 3 for students identified at risk at TIME 1 baseline?*

a) *Do students that responded to the intervention at TIME 2 immediate post-testing maintain the gains they made 4 years later at TIME 3 follow-up (i.e., Are they still average readers?)?*

b) If not, can we predict who will respond and who will not at TIME 3?

c) *Can we identify for whom inoculation works and for whom it does not at TIME 3?*

A need also exists to examine the influence of teachers on student R*t*I outcomes. Thus, the present study addresses Fuchs and Fuchs' (2006) critique that future R*t*I research must examine the role of teachers in the process of delivering effective reading interventions, specifically how a teacher's level of implementation may influence their students' R*t*I outcome. ABRA was never meant to be a plug and play tool; therefore, in the Pan-Canadian RCT study, implementation fidelity measures (IFM) were added to monitor how the teachers implemented the ABRA tool (see Savage et al., 2013), which illustrates that the designers of the original study understood that, in the classroom, technological tools such as ABRA are only effective with pedagogical support. IFMs enable an examination of the effects of an ABRA intervention on teachers in the sense that researchers can observe to what degree and how effectively teachers integrate ABRA into their literacy practice. For example, Savage et al. (2013) reported that of the 37 ABRA intervention teachers in their study, 7 were *entry*-level teachers, 20 were *adoption*-level teachers, 7 were *adaptation*-level teachers, and 3 were *differentiated adaptation*-level teachers. In addition, Savage et al. (2013) found that their analysis of the variation in the observed quality of teaching showed that ABRA effects on student outcomes were partly mediated by the quality of the teachers' implementation of the technology. For instance, a subset of excellent technology teachers who adapted their word-level phonics teaching to ABRA had a greater impact on phonological blending (d = .92), sight word reading (d = .43), letter-sound knowledge (d = .38), and story retell fluency (d = .22).

Teachers' responses to their ABRA training, demonstrated by their level of ABRA integration into their ELA lessons, were the factors that the present study considered when measuring the ABRA teachers' short-term R*t*I. In this sense, such an analysis is an extension of the notion of R*t*I at the teacher level. Further, the use and examination of the IFMs is essential to gain a more accurate picture of students' R*t*I. Without fidelity data, it is difficult to know whether a student's non-responsiveness to a Tier 1 intervention is a consequence of improper treatment implementation or a result of student characteristics. For example, are students responding to a well-designed intervention or did they receive a poorly implemented intervention, which may suggest that they are still dealing with an environmental deficit? The prediction is that those students who received a higher quality delivery of an ABRA literacy intervention will demonstrate greater short- and long-term reading gains. Specifically, the student follow-up data collected and the original IFM data collected on how teachers implemented the ABRA intervention were used to explore the following key questions about how a teacher's implementation of an ABRA intervention can influence student outcomes:

Question 3. *Does a teacher's level of implementation of ABRA ELA lessons during a Tier 1 intervention stage impact the influence of pupil-level non-cognitive variables on students' learning outcomes?*

a) *Can a teacher's level of implementation of ABRA ELA lessons during a Tier 1 intervention stage influence student-level outcomes at immediate post-testing (TIME 2)?*

b) *Can a teacher's level of implementation influence the short- and long-term reading ability gains of at-risk students (both the responders and non-responders to an intervention)?*

The guiding questions above will explore both possible main effects of demographic variables on student outcomes as well as provide an examination of possible interaction effects between teacher's level of implementation of ABRA ELA lessons and the demographic variables of interest.

Identifying students at risk. To address the pupil-level questions above, students who could be at risk for reading problems at TIME 1 baseline will first need to be identified. In the R*t*I literature, students identified as at risk have commonly been those who are performing below the 30^{th} percentile, (Al Otaiba & Fuchs, 2006; Simmons et al., 2008; Van der Kooy-Hofland, Van der Kooy, Bus, Van IJzendoorn, & Bonsel, 2012).

The present study created a *composite variable* (*z*-score) by combining the raw scores from multiple literacy measures (i.e., letter-sound knowledge [LSK], phonological awareness [PS], sight word reading [FRY words]). The factorability of these three variables was first examined, with all three correlating at least .3 with each other, suggesting sound factorability. The Kaiser-Meyer-Olkin measure of sampling adequacy indicated that the strength of the relationship among variables was high (KMO = .74), as it was above the recommended value of .6. The Bartlett's test of sphericity was significant (χ^2 (3, N = 1181) = 1685.81, p < .001). These findings suggest that our literacy variables are highly intercorrelated, so combining them to create a *z-score* produced a more reliable latent variable, since this measure is clustered across three scores, which removes a certain amount of measurement error (Leech, Barrett, & Morgan, 2015; Tabachnick & Fidell, 2013).

Using this new composite variable, the lowest performing 30% could be identified. It should be noted, that the TIME 1 listening comprehension (LC) variable was not included in the creation of the TIME 1 latent literacy skills variable, but rather LC was used as an IQ measure of students' oral vocabulary.

At each stage of the pupil-level assessment – TIME 1 baseline pretest (T1), TIME 2 immediate post-testing after intervention (T2), and TIME 3 delayed post-testing at follow-up (T3) – two distinct student groupings were created: (1) the bottom 30% identified as the at-risk group and (2) the remainder 70% identified as average (not at-risk) readers. This process created the *binary categories* that were used in the pupil-level analyses. The T3 analysis used two groups of binary predictors: for T1—at risk/not at risk; and for T2 —responders/non-responders. This analysis examined whether the categorization at T1 or T2 could predict the categorization at T3. Using these binary predictors, the *logistic regression* analyses that examined the pupil-level questions were run.

Teacher-level research questions. Response-to-intervention (R*t*I) is a process that hopefully leads to certain desired outcomes (Al Otaiba et al., 2010). At the pupil-level of analysis, the R*t*I outcome that is measured is often students' post-test achievement scores. However, as discussed previously, often in the literature, an examination of teachers' R*t*I outcomes—how they respond to being part of an intervention study—has been neglected. One of the important R*t*I outcomes at the teacher-level is reflected in the following question: Are teachers still using the new resource they were trained on, and utilized, during the intervention study?

Part A. Examining teacher-level RtI outcomes. At the teacher-level of analysis, one objective was to examine teachers' responses to being part of the ABRA intervention study. As discussed in Chapter 4, the teacher change models that the present study examined are concerned with what motivates teachers to change their practice to take-on or not take-on something new. However, the situation examined by this follow-up study differs slightly in comparison to the situations commonly discussed in the teacher change literature. Since the ABRA teachers had already decided to be part of the ABRA intervention study, they had changed their practice (in varying degrees, and at least for the duration of the intervention phase of the original study) by taking on a new literacy innovation. Therefore, this R*t*I follow-up study has modified the teacher change question. The R*t*I teacher-level response that it explored does not focus on whether the teachers chose to take-on something new or not, but whether the ***teachers continued to use a technological intervention*** that was deemed effective for improving students' literacy outcomes—the ABRA literacy program. Thus, one of the primary objectives at the teacher-level of analysis was to explore teacher-level R*t*I outcomes by applying the four domains of teacher change influence: the personal domain, external domain, domain of practice, and domain of consequence. For example, for the purposes of this study, the following particulars about the four domains of interest are assumed: 1) ***External Domain***. The measure for the external domain is uniform and constant in the equation, since the ABRA teachers were to have been provided with similar training and support on the ABRA tool. 2) ***Personal Domain***. This is an unknown measure in the equation because currently, we do not have information on the ABRA teachers' personal (internal) domain, although we can measure their external actions—how they implemented ABRA. 3) ***Domain of Practice***. The domain of practice is somewhat uniform, but it also can be tied to values, which may have been reflected in the level of implementation of the ABRA tool. 4) ***Domain of Consequence.*** The measure of this domain was student outcomes.

According to the four-component theory model, the prediction is that those teachers who implement ABRA well will carry on using it in the future. The rationale for this prediction is the assumption of high value. For example, if teachers are making the effort to change everything they are doing to integrate the ABRA tool into their long-term practice, they are either obedient or really believe that this tool is going to work. Based on the assumption of high value, the prediction is that those teachers who demonstrated a higher level of implementation (i.e., adaption-level of implementation according to an IFM rating) during the intervention phase of the study will be more likely to carry on using the intervention at the time of the follow-up, compared to their colleagues who incorporated ABRA into their practice at a basic-level.

A comparison of the teachers' follow-up use of the ABRA tool and how well they implemented ABRA during their time in the study was undertaken in the present study with an aim to determine whether a relationship exists between teachers using ABRA at the time of follow-up and their level of implementation of ABRA in 2007–2009 intervention. The questions guiding the investigation to address teacher-level R*t*I outcomes were the following:

Question 4. *Have teachers integrated ABRA (technology) into their practice? If so, in what capacity are they currently using ABRA, and what are the key factors influencing their decision? If not, what key factors are influencing their decision not to use the ABRA program?*

a) *Does a teacher's level of implementation of the ABRA ELA lessons during the Tier 1 intervention stage influence a teacher's decision to continue using the ABRA intervention after the conclusion of the study?*

b) *In relation to treatment integrity levels, with a particular focus on adaption, does a pattern exist regarding teachers still using ABRA 4 years later?*

Part B. Retrospective questions exploring how teacher change may have occurred. The present study also carried out an exploratory examination of teacher change models. This was a retrospective exploration, tentative in nature, since teachers were asked to reflect on what they did 4 years ago. The teachers' sense of the causality of their own change process was investigated through their reflections. Specifically, teachers' responses to factors influencing their decision to adopt or not adopt the ABRA tool into their long-term practice were analyzed in relation to the four domains of teacher change influence discussed in Chapter 4.

The focus of the teacher follow-up questions was to explore whether, and to what extent, teachers had integrated ABRA into their early literacy practice, to identify the factors that may have influenced their continued use of ABRA, and to examine whether being part of the actual intervention played or did not play a role in teachers changing their literacy practice long-term. The follow-up questions were created with the domains of influence of the teacher change models in mind, with an objective to identify and better understand potential factors that may influence teachers' R*t*I while implementing a new innovation during a Tier 1 level intervention study. By examining the quantitative and qualitative dimensions of teacher responses, another goal of this follow-up study was to identify the key teacher-level factors that may influence teachers' long-term adoption of a technological innovation proven to be effective. The exploratory retrospective questions explored in this study were the following:

> ***Question 5.*** *Are there key factors within each of the four domains of teacher change that can best predict whether a teacher will respond favourably to being part of a technology-based intervention study, and subsequently take-on a new resource into their long-term practice?*
>
> ***Question 6.*** *Is there an identifiable trajectory order of the four teacher change domains (i.e., external domain, domain of practice, domain of experience, and domain of consequence) that predicts whether teachers will integrate a technological innovation into their long-term practice?*

To address these questions teachers were asked to reflect on their involvement and training during the original 2007–2009 ABRA study. The prediction was that if the external domain was particularly strong, this could lead to teacher change. Specifically, the researcher was interested in determining whether teacher change is a result of changing teachers' beliefs and attitudes first, or whether making a change in student behaviour/learning first is what may have had a greater influence on teachers' decisions to continue using the ABRA tool. For example, through the interview process, the initial focus of the teacher change model comparison was to identify whether the teachers' linear trajectory process of change was closer to: the **Model (A)** Training in ABRA ➔ Your attitudes, beliefs, and values ➔ Change in practice (implementing ABRA) ➔ Students learning to read (response to ABRA); or the **Model (B**) Training in ABRA ➔ Change in practice (implementing ABRA) ➔ Students learning to read (response to ABRA) ➔ Change in attitude and values about the use of ABRA (technology in ELA lessons); or whether teachers' reflections on their decision-making process to continue using (or not) ABRA in their practice represents a change model that is closer to a more interactive, nonlinear process like that of **Model (C).**

Through a careful analysis of teachers' responses to the follow-up questionnaire and interview questions, the objective of the present study was to identify factors that may predict a teacher's greater likelihood to have a more favourable R*t*I? Although some components of the teacher R*t*I analysis were retrospective and tentative in nature, this study was worth doing to better understand the model of teacher change. Furthermore, the findings from such an examination may have direct practical implications in terms of identifying the factors that may contribute to the most efficient professional development focus for teacher change.

Chapter 6: Research Design and Methodology

Overview of Research Design and Methods

Chapter 6 provides a detailed description of the research design and methods of this follow-up study. This chapter contains CONSORT diagrams illustrating the flow of the student and teacher samples from the original study to the follow-up. It includes a descriptive review of the data collection timeline, process, and methods used. The chapter ends by restating the key research questions guiding this dissertation study.

Research Design

This was a quantitative follow-up study with a nested qualitative component that examined the long-term responses of both students and teachers who had participated in a computer-based reading intervention study up to 4 years earlier. At the pupil-level, a standardized reading assessment measure was used to collect quantitative data on students' reading ability and skill level 4 years after their involvement in the 2007–2009 ABRA Pan-Canadian intervention study. Logistic regression analyses, using Statistical Package for Social Sciences (SPSS) software, were run to investigate the pupil-level data.

At the teacher-level, a mixed-methods approach of data collection, referred to as *questerview* (Adamson, Gooberman-Hill, Woolhead & Donovan, 2004), was employed to collect both quantitative and qualitative data on teachers' past and current use of the computer-based reading intervention tool ABRACADABRA. The questerview data collection technique involved using completed follow-up quantitative survey questionnaires as guides to gather additional in-depth qualitative data during the teacher interviews. The teacher-level quantitative data were examined using standard parametric tests in SPSS. A *deductive* thematic analysis was the guiding analytical approach used to examine the teachers' qualitative responses to the open-ended survey and interview questions (Braun & Clarke, 2012).

Participants

Original 2007–2009 samples. The present study focuses on two separate follow-up population samples from the 2007–2009 ABRA Pan-Canadian study: the student-level sample and the teacher-level sample.

Original pupil sample. This was a follow-up study of the original participants of the Pan-Canadian 2007–2009 study (see Savage et al., 2013). The final RCT study sample included 1,067 consenting students from 74 classrooms: 316 kindergarten students (n = 154 ABRA, n = 162 Control), 616 first graders (n = 352 ABRA, n = 264 Control), and 135 second graders (n = 43 ABRA, n = 92 Control). The sample was comprised of 543 girls and 524 boys. The random allocation of classrooms resulted in 284 girls and 265 boys in the experimental group (total n = 549), and 259 girls and 259 boys in the control group (total n = 518). The participating students at pretest averaged M = 73.69 months old (SD = 10.11 months), with a range of 67 months.

In addition to the original RCT sample, seven additional classrooms participated in the intervention phase of the study from 2007 to 2009. Complete data was collected on the consenting students (n = 114) from these additional classrooms. However, these seven classrooms did not meet the random assignment requirement to be part of the final RCT analysis (for details see Savage et al., 2013). Six of the seven classrooms used the ABRA intervention during the phase of the study. This additional subsample consisted of 51 girls and 63 boys. At time of pretest, these additional students averaged M = 77.10 months old (SD = 7.73 months), with a range of 34 months. The aim of this dissertation was to collect follow-up data on as many of the original consenting students (n = 1181) as possible who took part in the 2007–2009 ABRA intervention study.

Original teacher sample. Efforts were made to follow-up with as many of the classroom teachers as possible who consented to be part of the original ABRA RCT study. From 2007–2009, 81 classrooms participated in the data collection phase of the original study. Seventy-four classrooms represented the final RCT study sample, while an additional seven participating classrooms were excluded from the final analysis because they did not meet the RCT allocation criteria.

From the original 81 classrooms, 19 of the teachers taught two different participating classrooms during the study phase. These teachers were classified as *duplicate teachers* in the ABRA database. These duplicate teachers either participated in both years of the study (n = 8), or in one research year, they taught both a control class and an ABRA class (n = 11). Teaching two different classes in the same academic year was possible for these teachers, since they taught two different classes of kindergarten students or they taught the English language arts program to two different groups of students. The removal of the 19 duplications from the teacher dataset left a sample pool of 62 participating teachers that could be followed up. The 62 original study teachers taught at one of 24 non-selective urban, suburban, or rural elementary schools from one of the three participating provinces—Alberta (n = 15), Ontario (n = 16), or Quebec (n = 31).

Recruitment procedures. This section details the steps the researcher took in obtaining the participants for this follow-up study.

Originally consented 2007–2009 participants and confidentiality. In the original study, appropriate steps were taken to protect the confidentiality of the participants by assigning each a numerical ID code. For example, all the paper records for the original study were kept at in locked files, within a locked RA room, with a database number (i.e., Jonny Smith is coded 001, and all assessment measures conducted with Jonny have the code 001 instead of Jonny's name).

Kept separate from the dataset used for the present analysis is a data file with corresponding identifying information for all the students, teachers, and school codes in the original study. This information was used to identify the students and which schools they attended up to 4 years earlier when they participated in the original ABRA study.

Obtaining consent and ethical considerations. This present study followed up with the students and teachers from the original ABRACADABRA study. In full accordance with the ethical APA guidelines governing research, the principal researcher took appropriate steps to obtain consent from all participants involved, including school board administrators, principles, classroom teachers, parents, and students. These steps included: (1) submitting to the Research Ethics Board of McGill University a detailed copy of the proposal for this study along with copies of all the instruments to be used, so to acquire permission to proceed with the data collection; (2) applying to the Ethics and Research Committees of each of the seven school boards (i.e., English Montreal and Sir Wilfrid Laurier school boards in Quebec; the Foothills school division in Alberta; and the Thames Valley, Waterloo Region, and Avon Maitland school boards in Ontario) that were involved in the original 2007–2009 ABRA Pan-Canadian study, (3) contacting individual schools for proper authorization to collect data from the teachers, students, and parents of the students involved in the original ABRA study, and (4) obtaining informed consent from original participants to be involved in this follow-up study.

The principal researcher contacted all school professionals to discuss the goals and parameters of the project and to request their assent for involvement. Current administrators of the original 2007–2009 participating schools received a letter (Appendix A) outlining the purpose and nature of this investigation. The letters received by the principals, teachers, and parents included an assurance of confidentiality and a statement emphasizing that all participants have the right to withdraw from the research project at any point during the course of the study.

After administration consent was obtained, the principle received a list of all of the students who were originally part of the ABRA study. This researcher, with support from school staff, then determined which students were still in attendance at the school and which classrooms they were in. Classroom teachers were contacted directly to obtain their consent. Current classroom teachers were provided with letters explaining the objectives of this follow-up study, what their expected involvement would entail (i.e., the distribution and collection of parent consent forms and providing class-time to administer the student reading assessments), and a projected timeline when data collection in their classrooms would occur. Only **verbal** consent was required from current classroom teachers, since no data was collected specifically on them or their current classroom practices. Current classroom teachers assisted with the distribution of the parental consent forms.

Next, information about participation in the follow-up research was communicated to parents and legal guardians. Working with the current classroom teachers, parental consent forms (Appendix B) were sent out to every 2007–2009 participating student still attending the participating ABRA schools. Given that the subject population involved young children, obtaining parental permission was needed before proceeding with the data collection.

The parental consent form letters outlined the required reading tasks the children would be asked to complete and included an allocated time frame for these tasks. An additional background information checklist in the form of a parent questionnaire (Appendix C) also was attached to the informed consent letters. Parents were asked to provide demographic, home language, and literacy environment information; and to identify whether their children had any documented disabilities.

This information was obtained to ascertain the characteristics of the student sample and identify those students who currently may be receiving additional literacy support.

Approximately 2 weeks after the consent forms were distributed, the researcher made arrangements with the current classroom teachers to determine when and where the follow-up reading assessments could be administered. At TIME 3 follow-up, the student participants were in grades three to six. Therefore, as mentioned previously, informed written consent was re-sought and obtained directly from the parents and legal guardians of the participating children. Before testing, children with parental consent were orally informed about the study (see Appendix D for the script read to the children to obtain verbal assent). After the study had been orally explained, the students were required to give verbal assent before participating in the study.

Concurrently, attempts were made to contact as many of the original ABRA study participating teachers as possible to collect the teacher-level R*t*I follow-up data. The accessible original teachers all received consent form letters (Appendix E). The teacher consent forms outlined the objectives and value of the study, and provided details about the data the researcher wished to collect from them. Follow-up questionnaires for the teachers (Appendix F) to fill out were attached to the consent forms. On the consent forms, teachers also were asked whether they would be willing to be interviewed regarding their experiences being in a literacy intervention study. Teachers were informed that the in-person interviews would be audio recorded, transcribed, and coded for research purposes. Of those teachers who provided consent, the researcher confirmed the times when follow-up interviews could be conducted.

Consenting students and teachers were free to refuse or discontinue participation in the follow-up study at any time. All data collection and use was conducted in accordance with the ethical standards of the American Psychological Association. All original data are securely stored and will be destroyed after 5 years of being published.

Compensation for participation in this study. Participants in this study were all volunteers and did not receive any monetary compensation for their participation. As a minimum gesture of gratitude, participating students got to choose a *thank-you for participating treat* (a sucker, little chocolates, or stickers) while they wrote their assessment, and they also could select another treat when they handed in their completed reading measure. Prior to distributing the student thank-you treats, the researcher always consulted with both school administrators and classroom teachers for permission to do so. Seeking permission was necessary, since the researcher did not wish to go against school policies regarding treats and also wanted to avoid triggering any potential food allergies that participants may have had.

The teachers who participated in the interview sessions also received a small gesture of appreciation for their time. When setting up the teacher interviews, the researcher would enquire about each teacher's drink preference from a local coffee shop in the school's vicinity. The researcher would then bring the teacher her choice of beverage to the interview session. After the interview session was completed, each teacher also received a handwritten thank-you note with a coffee card enclosed.

Follow-up R*t*I samples. This section provides details about the student sample and teacher sample for this present follow-up study.

Pupil T3 follow-up sample. The goal was to collect T3 data on as many of the original consented students (n = 1181) as possible who were involved in the 2007–2009 ABRA Pan-Canadian study, not just those students included in the final RCT published paper. After school board approval had been obtained, the current administrators of the 24 elementary schools involved in the original study were all approached to be part of this follow-up study. Two principals from Ontario did not provide the required consent to access the students (*n* = 207) at their schools. Another 294 students were not accessible, since they were no longer attending one of the 22 remaining ABRA intervention schools. Thus, a pool of 680 students from the original ABRA study was available for possible recruitment.

After administration approval was obtained, parental consent forms were sent home. From the sample pool of 680 potential participants, 488 consent forms (72%) were returned. The pupil-level attrition rate was 4% as a result of 21 students being absent on the days of testing. In the end, 467 of the original 1,181 participants (40%) took part at T3 (*n* = 249 ABRA, *n* = 218 Control) (see Figure 5).

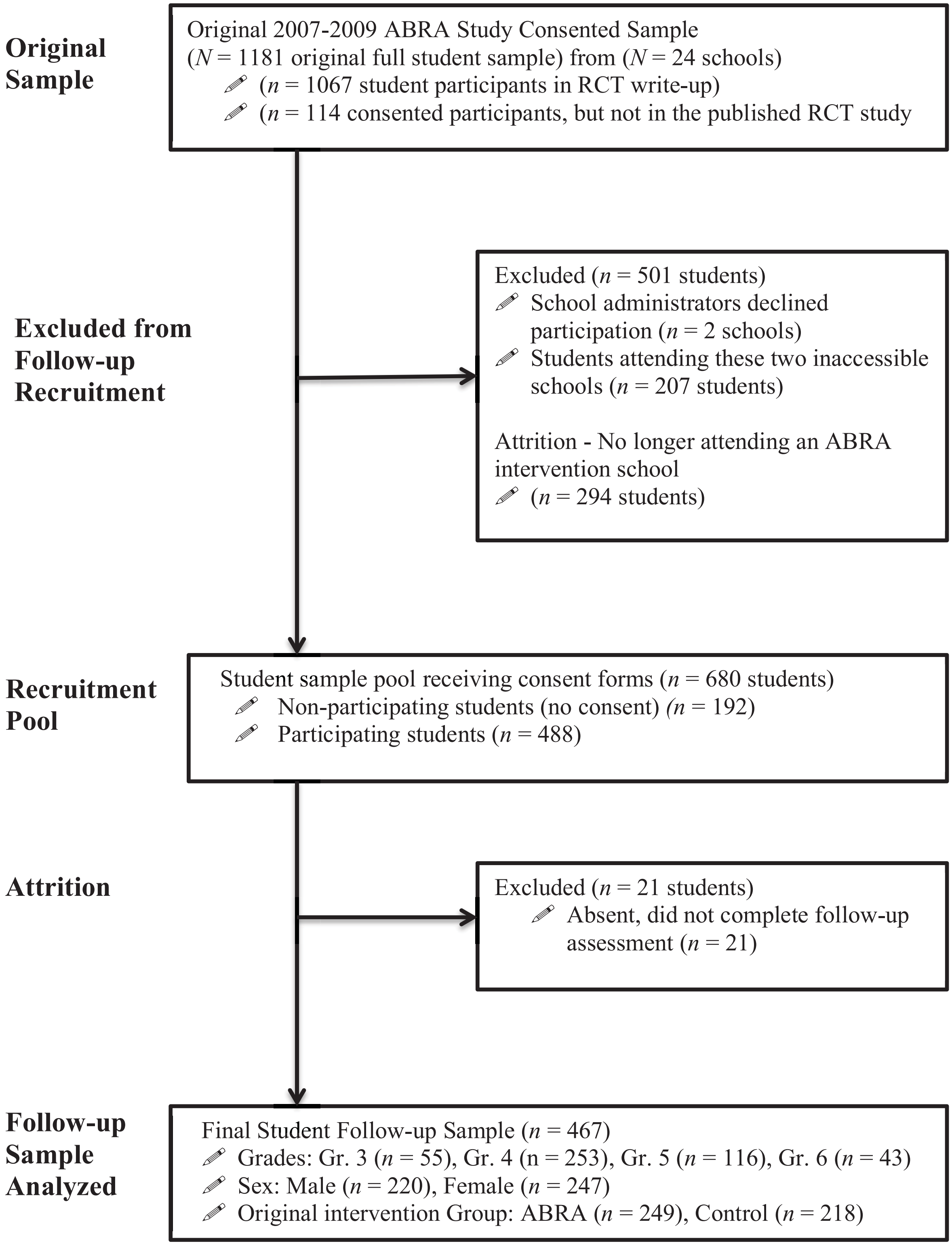


Figure 5. CONSORT flow diagram of the follow-up pupil sample.

Teacher follow-up sample. The teacher contact information provided at the time of the original study (i.e., the school of employment and school email addresses) was used to locate the teacher participants. As with gathering the student data, approval from the seven school boards involved in the original ABRA Pan-Canadian study during 2007–2009 was first obtained. Then, the current administrators from the 24 original participating schools were approached to gain permission for the researcher to contact the teachers and collect the follow-up data. As previously noted, two Ontario school principals declined permission to allow their staff and students to participate in the follow-up study, resulting in a lack of access to 12 teachers. From the remaining 50 potential teacher participants, a number had moved away from the school district ($n = 13$) and others had retired ($n = 4$). As a result, at time of the follow-up, 33 teachers were available in the sample pool of potential teacher participants ($n = 11$ Alberta teachers, $n = 3$ Ontario teachers, $n = 19$ Quebec teachers). Each of these 33 teachers received a consent form letter outlining the details and value of this study. Follow-up questionnaires for the teachers to fill in were attached to the consent forms (Appendix F). Teachers received hard copies of these documents in their school mailboxes, and electronic copies also were sent to their school email addresses. On their consent forms, teachers also were asked if they would be willing to be interviewed to discuss their experiences of participating in a literacy intervention study.

Eleven teachers declined consent and did not take part in the follow-up study. Twenty-two of the original 62 teacher participants (35%) agreed to participate in some capacity with this study ($n = 3$ questionnaire only portion of the questerviews, $n = 19$ questionnaire and interview portions of the questerviews) (see Figure 6).

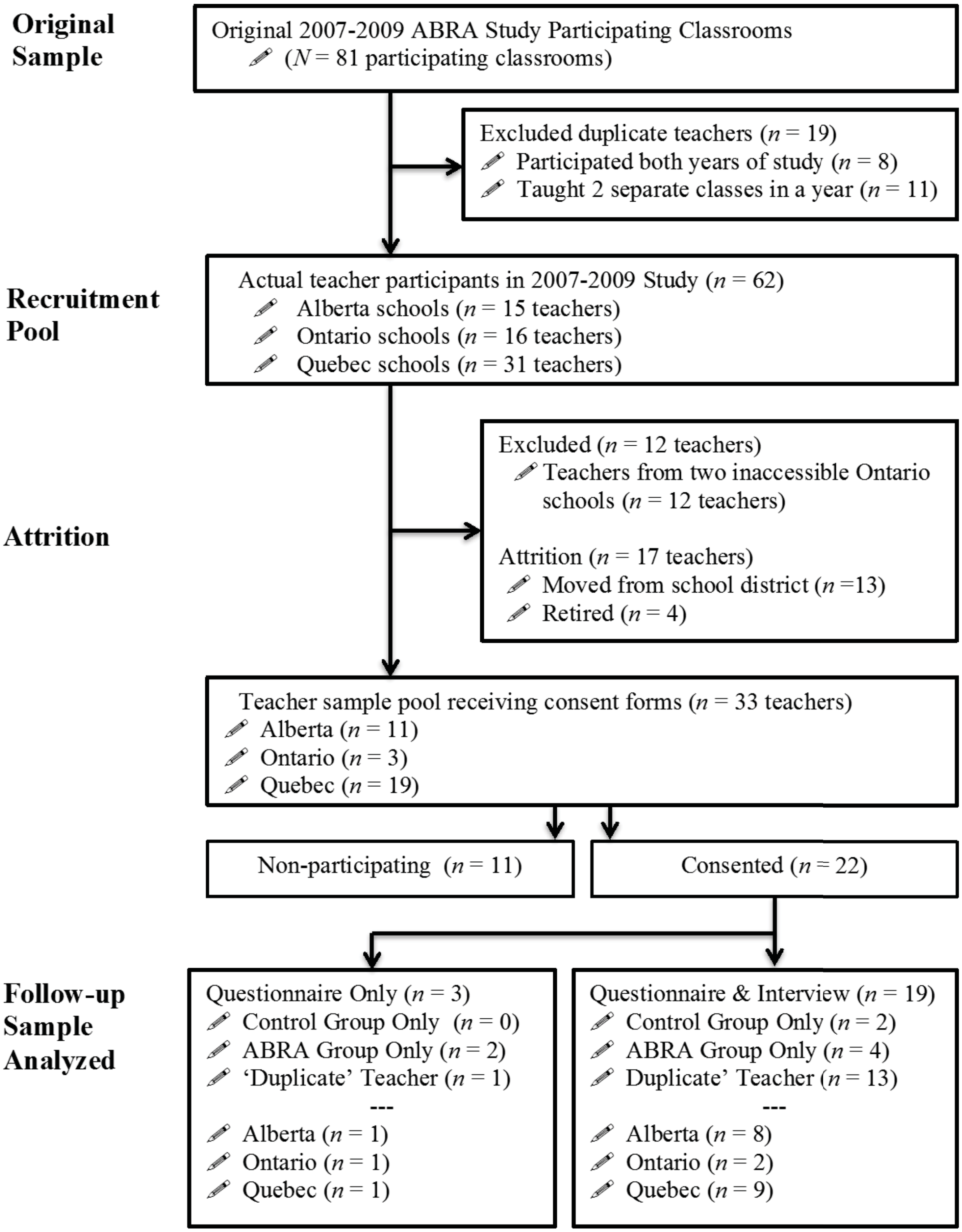


Figure 6. CONSORT flow diagram of follow-up teacher sample.

Data Collection

During this follow-up study, the researcher was immersed in the data collection process over a period of several months that spanned two school years (see Table 1). The formal data collection for this study began in March 2012 immediately after obtaining ethics approval from McGill University and passing the research proposal process. By mid-February 2012, each of the six school boards from the three participating provinces had received a request to conduct this follow-up study. By mid-March, the school boards from Alberta (n =1) and Quebec (n = 2) and two out of three of the Ontario school boards had provided consent to contact their schools. The sixth remaining school board from Ontario provided consent by mid-May

Table 1

Data Collection Timeline from January to November 2012

Province	**Obtaining Ethics & Consent**	**Obtaining Consent**	**Follow-up Data Collection**	
	Contacted and worked with school boards to obtain consent for follow-up study.	Contacted schools and re-obtained parental consent of previously consented students. Obtained teacher consent to participate in completing follow-up questionnaires and taking part in follow-up (tape-recorded) interviews.	**Students:** Group-administrated assessment of students' reading skills (60–90 min/class).	**Teachers:** Distribution and collection of teacher follow-up questionnaires. Conducted teacher interviews (60–90 min/teacher).
Alberta	February–March	March–April	March–April	
Quebec	February–March	March–June	March–June	
Ontario	February–May	September–October	October–November	

The researcher collected all of the student and teacher data on her own from one province at a time. The data collection began with the province of Alberta, since its school district gave almost immediate approval. Before going to Alberta, the researcher used email correspondence to arrange meeting times with the Alberta school administrators (n = 7). After arriving in Alberta, the researcher first met with the school administrators to go over the study. School by school, the researcher addressed any questions the school administrators had about the follow-up study, and began the process of distributing the consent forms to the students and teachers of the original study.

Approximately 2 weeks after the initial distribution of the parental consent forms, the researcher revisited the participating schools to collect the returned consent forms and begin the data collection. Student data collection, and the time and place of the student GRADE testing was arranged with the classroom teachers. Whenever possible, efforts were made to carry out the testing during the students' regular language art periods.

Of those Alberta teachers who provided consent, the researcher confirmed and arranged times for the follow-up audio-recording interviews to be conducted. The in-person interviews were arranged around the student testing times. The teacher interviews were carried out before or after school, at lunch, or during a teacher's spare (prep time) blocks.

Similar data collection processes were then undertaken in the Quebec schools (n = 14), between April 2012 and June 2012. The researcher obtained consent late in the spring of 2012 from one of the three Ontario school boards. Thus, the Ontario data collection was not carried out until the following school year from October 2012 to November 2012. Despite contact efforts, the researcher was only successful in collecting data from one of the three original Ontario schools.

Pupil Data Collection Process and Measures

With the current classroom teachers' permission, arrangements were made to collect the student-level data. Based on the number of consented students, arrangements were made on a class-to-class basis for how best to administer the group-administered literacy assessments with the least amount of disruption to the teacher's and children's regular classroom schedule.

For example, when consent was received for more than half the students in a classroom, in most cases, the classroom teachers preferred that the researcher take the whole class (consented as well as non-consented students) for the testing period. This made the classroom teacher's planning for the day easier and often allowed her/him a break from his/her students if she/he wished to get something else accomplished during the 90 minutes the researcher needed with the students. As the GRADE is a literacy task, the teachers generally viewed the activity as extra literacy practice for their students and that is perhaps why in most classes with more than 50% of student participation, the GRADE was often administered to the whole class. In the cases where the GRADE was administered to all students in the class, the completed reading measure booklets for those students for whom we did not receive consent were labeled as non-consented. The booklets from the non-consented students were shredded prior to analysis and the data collected was never used.

In situations where the number of consented participants was fewer than half the class, the classroom teacher decided where T3 follow-up testing would be carried out. For example, in most instances, another room in the school was found (i.e., the resource room, library, or empty classroom) where T3 testing could be carried out.

Pupil-level measures. In the original ABRA study, a battery of instruments were selected to assess the specific literacy skills targeted by the intervention tool under investigation. As ABRA aids alphabetic (e.g., word-level skills related to decoding, letter knowledge, and word reading), reading fluency, spoken language, and reading comprehension growth, tasks were selected to examine these aspects of students' general cognitive, reading, and listening development before and after exposure to the ABRA intervention (see Savage et al., 2013).

However, for the present follow-up study, only one standardized measure was used to collect data on students' reading abilities at T3. As a comprehensive standardized reading measure, the Group Reading Assessment and Diagnostic Evaluation (GRADE)—also used in the original study—was the primary tool used to assess students overall reading ability at T3. The practicality of the GRADE measure (for example, its comprehensive nature, capacity for group administration, previous experience with, ease of administration, and minimal time requirements) made it the most suitable assessment tool to use.

The GRADE. At the pupil-level, quantitative data was collected to assess students' reading abilities at Time 3 follow-up (T3). This data was collected through the administration of the GRADE, a standardized measure of reading competency. It is a norm-referenced, developmentally based assessment of reading skills. The GRADE is designed to measure reading skills at a variety of levels and provides grade-appropriate versions to analyze students' reading strengths and weaknesses, identify students who require remedial or enrichment programs, and assist in monitoring students' progress throughout the academic grades (Williams, 2001).

The GRADE includes subsets for measuring a range of essential and complementary reading skills (e.g., listening comprehension, word reading, vocabulary, and reading comprehension). Further, the GRADE can provide a precise measure of change by isolating different aspects of reading skills. At T3, students were assessed on word reading (Grade 3 only), vocabulary, listening comprehension, and sentence and passage comprehension.

Word reading. The word reading subset was administered only to students being assessed at Level 3 of the GRADE. This subset assessed children's recognition of sight words and their ability to decode regularly spelled words. Words that are defined as *decodable* are those that can be pronounced by regular phonics rules and follow conventional spelling patterns. *Sight words* are not pronounceable using regular phonics rules, so students must learn to recognize them by sight. For both decodable words and sight words, high-frequency words were selected for the design of the test so that the target words would reflect familiar grade-appropriate vocabulary. For this task, students identified the target word from a list that included four or five distracter words. The distracter words were real words that were visually or phonologically similar to the target word (although, at times, some unrelated words were included). To successfully accomplish this task, students needed to have developed word recognition skills (such as sound-symbol correspondences) to discriminate the target words from the distracters.

Vocabulary. This subtest of GRADE was administered to all children from grades three to six. The vocabulary subtest assessed students' knowledge and understanding of words. Students independently completed this task after the test administer had modeled two examples. This task required students to silently read short phrases or sentences with a bolded target word followed by five possible definitions. The students' task was to select the correct meaning of the target word, which could be a noun, verb, adjective, adverb, or preposition. Target words included phonetically irregular words, which are regular words for which the pronunciation can be derived by applying common grapheme-to-phoneme rules. Again, the target vocabulary words were common and grade-appropriate sight words or decodeable words.

Listening comprehension. This subtest of the GRADE was administered to all children from grades three to six. This subset was used to assess students' understanding of spoken (oral) language. Children were read sentences and then asked to select a picture from four choices that best illustrated the meaning of each sentence. In this subtest, students responded to vocabulary, grammar, and inferential questions as the test items increased in linguistic complexity, which required them to attend to subtleties in the verbal cues. This subtest was of value, since reading is a receptive language skill.

Sentence comprehension. This subtest of the GRADE was administered to all children from grades three to six. Sentence comprehension is a cloze procedure whereby students read a sentence that has a missing word and then select one of four words that best fits the context. Item difficulty varied based on the type of sentence structure presented (e.g., simple, compound, or complex) and the part of speech that was absent from the text (e.g., noun, verb, adverb, adjective, or preposition). To successfully complete this task, students had to be able to comprehend the sentence as a whole or complete thought by extracting meaningful information from each sentence based on the context, vocabulary, and syntax.

Passage comprehension. This subtest of the GRADE was administered to all children from grades three to six. In this subtest, students read short passages and then tried to select the correct multiple-choice response. In this task, the test passages were examples of synthetic texts that had been constructed with controlled vocabulary, length, and level of difficulty appropriate to the children's grade-level. To answer the passage comprehension questions, students had to rely on information provided in the text and not background knowledge, since students should not have encountered these passages before. Thus, this subtest assessed valuable metacognitive strategies associated with the educational objectives highlighted by Bloom's taxonomy of level of learning in reading unfamiliar text. This subtest also identified developmental differences between students by examining whether they were paying attention to decoding and comprehending while they read. In Level 3, this subtest measures students' questioning, clarifying, and summarizing ability. In Levels 4 to 6, the passage comprehension questions also incorporate the metacognitive strategy of predicting, in addition to the three previously mentioned abilities.

For Level 3, the raw scores on the sentence comprehension and passage comprehension subtests were combined to yield a comprehension composite score that represented students reading comprehension skills level. Further, the reading comprehension tests measure the metacognitive skills of previewing, predicting, clarifying, and summarizing (Leech et al., 2015), which make them a good measure of the comprehension skills taught in ABRA.

Administrating the measure took between 60–90 minutes, completed in a whole-class/group format. The test administrator read aloud the instructions for each subtest and guided students through an example of each subtest item. As the examiner read the instructions and examples, the students marked their answers individually in a student response booklet. Raw scores from each of the subtests can be converted into stanines, standard scores, percentile ranks, normal curve equivalencies, and grade equivalencies (Williams, 2001). At T3, the raw scores from listening comprehension and vocabulary subtests were calculated in stanines (standard scores ranging from 1 to 9 ($M = 5$, $SD = 2$), while standard scores were obtained for a sentence comprehension composite, passage comprehension composite, and total reading score.

Reviews of the GRADE (Fugate, 2003; McBride, Ysseldyke, Milone, & Stickney, 2010; Waterman, 2003) have concluded that it is a well-designed, reliable, and valid measure of early reading ability. Williams (2001) also has pointed out that the content of the GRADE test items have been analyzed carefully to prevent item bias in terms of sex, ethnicity, special education needs, and geographic region. Further, the GRADE can assist in the evaluation of literacy programs through the documentation of individual students' progress and the assessment of reading competencies of students at different grade levels, since the GRADE is designed to identify low-, middle- and high-performing students in each grade level of schools (Fugate, 2003). The GRADE also provides guidelines for out-of-level assessment if a student is identified to have exceptional reading skills (either too low or too high for the current grade level) (Fugate, 2003). The instruments in the GRADE are reported to have strong internal consistency (*r* = .95 - .99), which indicates a high degree of homogeneity among the test items in the GRADE subtests. Depending on the grade version of the test, the test-retest reliability is also moderate to high (*r* = .80 - .96) (Williams, 2001).

Parent questionnaires. Responses to the *parent questionnaires* provided additional pupil-level data. The follow-up parent questionnaires were attached to the parent consent forms (Appendices B and C). The updated parent questionnaires followed a similar framework as those from the original study. The parent questionnaire was redistributed to reconfirm and update general student demographic information (i.e., age, sex, ethnicity, and language(s) spoken) and to also provide information on potential non-cognitive pupil-level variables that may influence a student's response to intervention.

Pupil reading ability checklist. After parent consent forms and parent questionnaires were returned, current classroom teachers were asked to provide general information on students' reading abilities by indicating which students: (1) were developing as typical readers, (2) may be at risk of experiencing long-term academic difficulties as a result of current reading difficulties being exhibited, and/or (3) have been identified with specific learning and/or behavioural difficulties that may be impacting their current reading outcomes. This information was tracked by using a checklist of consented students. Classroom teachers were asked to check-off which of the three formerly mentioned categories best represented each participating student. At the time of analysis, this information was compared and triangulated with parents' questionnaire answers to confirm the identification of students who may have learning and/or behavioural difficulties that could be impacting their reading outcome scores.

Teacher Data Collection Process and Methods

Along with a description of this study and a consent form for their participation, the teachers were asked to complete a follow-up questionnaire and provide feedback about their original experience being part of the Pan-Canadian study and their current use of ABRA in their ELA program. Twenty-two teachers agreed to participate in the study in some capacity.

Questerviews. Data for the teacher analysis was derived in part from survey analysis using teacher questionnaires as the primary data collection instrument. In addition, participants were asked to review and elaborate on their questionnaire responses by taking part in a 60–90 minute semi-structured interview. This data collection technique has been coined *questerview* (Adamson, Gooberman-Hill, Woolhead, & Donovan, 2004). The authors use the term questerview to describe the data gathering technique of combining the use of self-completion questionnaires and questions during in-depth interviews. The inclusion of the teacher interviews enabled the researcher to delve deeper into survey responses that may have otherwise have been limited in depth and breadth (Kvale, 1996).

All of the consenting teachers (n = 22) agreed to complete the survey portion of the questerviews. Nineteen of the 22 consented teachers also agreed to review their responses to their questionnaires by participating in an in-person interview session that was audio recorded. The 19 teachers who agreed to the interview portion of the questerviews also completed a teacher change model-creation exercise at the end of their interview sessions.

Participants were audio-recorded while they reviewed their responses to the follow-up teacher questionnaire. Teachers were encouraged to discuss their responses to the survey items in more depth by providing clarifying and supporting detail. The inclusion of the survey questions as guides in the qualitative interviews facilitated a direct link between the two sources of data

(Adamson et al., 2004). Specifically, the qualitative data assisted in validating the survey responses and helped to explain and enhance the quantitative data from the teacher questionnaires.

Questionnaire portion of the questerview method. Every teacher in the follow-up sample received a hard copy of the 7-page questionnaire to complete (Appendix F). For teachers who preferred to do the questionnaire online and submit it electronically, an online version of the questionnaire was also available through *SurveyMonkey* (www.surveymonkey.com). One of the purposes of the questionnaire was to better understand teachers' use of technology in the classroom (in general) and ABRA (specifically) from the time of their involvement with the ABRA intervention study. The wording of the questions on the teacher questionnaire reflected the targeted respondents (i.e., the teachers involved in the ABRA study) and the technology context of the ABRA tool. The questionnaire was broken down into four specific areas of focus (1) general demographic background, (2) ABRA study participation background, (3) ABRA training and support, and (4) present use of ABRA. The question formats varied and included both open-ended and closed questions.

In addition, Likert-type scale questions also were incorporated into the survey to quantitatively examine potential positive and negative influencing factors. For example, participants were asked to report on how important potential factors were positively or negatively influencing their decision to incorporate a new teaching innovation into their teaching practice (see Appendix F, questions C-7 and D-9). These factors were assessed using a 5-point Likert-type scale anchored by 0—*Not important at all*, to 4—*Very important*.

Interview portion of the questerview method. The teacher interview data were collected using the teacher survey questionnaire as a guide for the interviews. Teachers reviewed their responses of their completed questionnaires while the researcher encouraged further details by asking additional open-ended questions (Appendix G). This original data were used to gather evidence and insights from the teacher participants. Teachers were asked to reflect on their experiences of being part of an intervention study and how their experiences influenced their current literacy practice (which may or may not include the ABRA tool). It was paramount that the evidence and insights obtained from the teacher interviews were as clear as possible. Wherever possible, the interviewees were asked to elaborate on their descriptions to clarify earlier understandings. As a researcher and the interviewer, I was self-aware that my primary role was to facilitate the dialogue so to capture the authentic voice and viewpoints of the interviewees (Gall, Gall, & Borg, 2007).

The primary objective of the interview portion of the questerviews was to get the teacher's story from their own perspective and provide greater detail and depth to their questionnaire responses. Each interview lasted approximately 60–90 minutes. The teacher interviews were digitally audio-recorded and transcribed verbatim (Kvale, 1996). The recording of the interviews was for transcription and quality assurance purposes only.

Questerview interview questions. In addition to using the teacher questionnaire as a guide for the interviews, open-ended questions also were incorporated into the interview process. For example, efforts were made by this researcher to encourage the teachers to describe their literacy practices in more depth, provide examples, and elaborate further whenever possible to assist them in articulating their own experience of using ABRA during the intervention phase and afterwards. The researcher made a conscious effort to use language that would encourage further dialogue and assist the teachers in confirming that an accurate picture of their story had been told; for example, the researcher used phrases such as "could you give me an example," "could you describe that," "am I getting this right," and so on. To provide the most accurate representation of the teacher's experiences of being part of the ABRA study, the collaboration between the researcher and the interviewee during the interview process was of great importance, since the end goal was to get as accurate as possible account of the teacher's perception of the factors that influenced her decision to incorporate (or not) ABRA into her reading program.

Neutrality. I was mindful to suspend judgment and control for researcher bias (Yin, 2015). The goal of this process was to encourage the teachers to honestly open up about their experiences using the ABRA program during and after their time in the intervention study. The participants were aware of this researcher's involvement in the original ABRA study, and I did not want this knowledge to influence their responses. At the onset of every interview, I made it clear that there were no right or wrong answers and that the participants should not feel that they needed to provide only positive comments about their involvement in the study; rather, I hoped that the teachers would express their candid views of what worked for them and what issues or concerns they may have had during their involvement in the study or with the ABRA resource. I stressed that I had no vested interest in the ABRA resource itself, and my only intention was to gather data in order to get a true picture of what factors influence teachers' use of a new resource.

Transcribing questerview interviews. The primary researcher transcribed each of the audio-recorded interviews. As mentioned, the in-person interviews that were audio-recorded were initially transcribed verbatim. All spoken words, including pauses, cut-offs in speech, and laughter were documented (Kvale, 1996). Most transition sounds or filler words (e.g., *hmmmm, ah-ha*) made by both the interviewer and interviewee also were documented in the initial transcriptions (Edwards & Lampert, 1993). However, it should be noted, that after the participating teachers had an opportunity to review their transcribed responses, the transcriptions were cleaned up. That is, the researcher consciously took steps to clean up the extracted data quotes by removing most of the filler words (e.g., *like, um*, *yup*) during the coding process. Also most of the *yups* and *yeahs* were noted as *yes*. These slight modifications to the transcribed data did not alter the actual content or meaning of the extracted quotes.

Member checking. A process of member checking was used to ensure that an accurate portrayal of the teachers' narrative was presented (Coyne et al., 2004; Creswell, 2014). All of the interview participants had an opportunity to review their transcribed interviews after they were completed. Via email, each participant received a copy of her transcription. Teachers were asked to check over the transcriptions for accuracy. Teachers also were encouraged to feel free to make any amendments to their original responses if they felt their transcribed responses were incomplete or no longer held true. Sixteen of the 19 interview participants (84%) replied back to confirm they had reviewed their transcripts. Most of the responding teachers reported that the transcripts fairly reflected their interview comments and no further changes were necessary (n = 11). A few of the teachers (n = 3 teachers) pointed out minor grammatical errors, while another two teachers suggested some content changes to their comments. For instance, one teacher after reviewing and reflecting on her responses wished to change how she initially rated the value of some of the potential factors of influence on her decision to use or not use a new innovation. Another teacher further clarified some of her comments about how she valued the use of iPad apps and touch technology over desktop programs for young students, and she also provided the names of some of the apps to which she was referring, but which had not been clearly noted in the original transcripts. The researcher made the requested changes to the transcripts, and the five modified transcripts were re-sent for a final member check. Each of these five respondents confirmed that no further changes were necessary.

Model building of the teacher change process. After the interview portion of the questerview method was completed, each teacher who was interviewed was asked to engage in one final activity. This activity involved the teachers identifying which linear teacher change model best represented their own teacher change process. Teachers were presented two different vignettes, one describing **Model A** (Change in External Domain ➔ Change in Personal Domain➔ Change in Domain of Practice ➔ Change in Domain of Consequences) and one describing **Model B** (Change in External Domain ➔ Change in Domain of Practice ➔ Change in Domain of Consequences ➔ Change in Personal Domain) (see Appendix H for an example of the vignettes used). A *vignette* is a method used to collect data by describing hypothetical situations and asking research participants direct questions to gain their perceptions (Adamson et al., 2004). These vignettes were useful for highlighting not only the differences between the two linear models of interest, but also for initiating a discussion to gain a richer perspective on the teachers' personal experiences regarding their own teacher change process (Adamson et al., 2004).

Selecting their models from linear Model A or Model B. After hearing the descriptions of the two teacher change models, teachers were asked to select which model they felt best represented their own decision-making process of incorporating a new innovation into their literacy practice. Of specific interest was to find out when during the stage of implementing a new resource did the teachers tend to report believing in the value of the new resource. Did they report believing in the value of the new resource *before* implementing it (Model A) or did they believe in the value of the new resource *after* seeing its benefits to desired salient outcomes (Model B)?

To assist in this process, the interviewer provided a premade illustration of the two models of interest (Appendix I). It should be noted that the decision to use the visual representation occurred only after the completion of the first teacher interview. On immediate reflection of the first interview, the researcher realized that a visual representation would not only aid in the process of describing the two linear teacher change models but also assist the teachers in selecting which model best fit their own process of integrating a new resource into their practice. Thus, only 18 of the 19 teachers who participated in the model building activity were shown the premade illustration during their interview sessions.

It should be noted that **Model C**, the cyclical interconnected model of teacher professional growth, was never verbally described or visually displayed to any of the participants. Prior to conducting the interviews, the researcher and her supervisor concluded that if shown Model C, most teachers would automatically favour the cyclical model without giving the two linear models a consideration. However, when teachers were encouraged to create their own teacher change process models, many of them elaborated beyond the two suggested linear trajectory models by illustrating various degrees of interconnectedness between some of the domains of their created personal teacher change models.

Creation of their own process of change model. After selecting either Model A or B, the teachers were then asked to create their own personal process of change model. The interviewer stressed that even though she would like them to try to incorporate the four domains of change (which were pre-written on post-it sticky notes), their personal process models did not have to resemble Model A or B. For many of the teachers, the interviewer also kept track of potential key variables that she would hear each teacher mention as influencing her current use of ABRA. For each teacher, the interviewer would write down their mentioned influencing variables on post-it notes and review the noted variables with the participant after their interview was completed. It was left up to each teacher to incorporate (or not) the noted variables into her final diagram.

Using the supplies provided by the interviewer (blank paper; different colour post-it sticky notes; different colour pens, markers, highlighters, and pencils), the teachers were encouraged to add any additional descriptions of variables they deemed influential in their decision-making process of incorporating a new resource into their teaching practice. Teachers were asked to think about how their process of change moved from one domain to the next and to illustrate this process. Teachers also were asked to consider from where their decision-making process commonly initiated (from which domain), and to illustrate the direction (flow) of their change process. As the teachers were creating their teacher change diagrams, the interviewer encouraged them to share aloud their thought process. Many of the participants complied and engaged in a step-by-step monologue regarding their process of integrating a new resource into their practice. Other teachers created their diagrams in silence, and provided a summary description of their completed model afterwards.

Triangulation. Multiple sources of data (questionnaire, interview, transcripts, model making, and e-mail correspondence) enabled triangulation (Miles & Huberman, 1994). Specifically, the triangulation of data through an examination of the teachers' survey and interview responses and their discussions during their teacher change model building process enabled the researcher to check the accuracy of their responses. If any inaccuracies arose, the researcher emailed the teacher for further clarification. Out of the 19 interviews, the researcher found only one glaring discrepancy that caused a teacher to be moved from one *Current Use of ABRA* group to another. For example, the teacher moved—due to her questionnaire comments—was placed initially in the *ABRA is part of my repertoire group,* but due to her responses during the interview, she was transferred to the *I am no longer using ABRA* group. A subsequent follow-up confirmation with the teacher confirmed that her transfer into the *I am no longer using ABRA* group was a more accurate portrayal of her current ABRA use.

Chapter 7: Pupil-Level Data Analysis and Results

Overview

This chapter provides the findings from the student-level analysis. *Statistical Package for Social Sciences (SPSS)* software was used to conduct the preliminary data analyses and primary statistical analyses. Using SPSS version 20, a number of standard parametric tests were run to examine the data, and address the key research questions, of this study.

It should be noted that even though hierarchical linear modeling (HLM) was employed in the original 2013 paper (Savage et al., 2013) it was not an appropriate method of analysis for this follow-up study. In the 2013 paper the unit of analysis was at the classroom level while this follow-up study focuses on pupil-level predictions. Furthermore, the T3 pupil sample is more modest in size resulting in very small *n* at the classroom level that precludes the use of HLM. Rather, factor analyses of individual differences were run on the pupil-level variables of interest (ethnicity, SES, and sex-differences) and the effects of intervention treatment on students' literacy skills (both in the short- and long-term). Logistic regression analyses were run to address the three key pupil-level dissertation questions. The use of logistic regression analysis was appropriate, since the aim of each research question was to predict reading outcomes from a set of predictor variables. Further, the dependent variable for each question was categorical rather than continuous (Leech et al., 2015; Tabachnick & Fidell, 2013).

Preliminary Data Analyses

Data screening. The first two pupil-level questions focused on the original full sample data at TIME 1 pretest (T1) and Time 2 immediate post-intervention (T2):

Question 1. *Are there identifiable pupil and contextual demographic variables of at-risk students who had a greater response to a Tier 1 (classroom-level) intervention at T2?*

Question 2. *Does a teacher's level of implementation of ABRA ELA lessons during a Tier 1 intervention stage impact the influence of pupil-level variables on student learning at T2? In addition to possible main effects, is there an interaction effect between the level of ABRA implementation students received and the demographic variables of interest at T2?*

This is a follow-up study of the Savage et al. (2013) 2007–2009 Pan-Canadian ABRACADABRA (ABRA) study. Thus, a preliminary screening and clean up of the full dataset (N = 1181) containing the T1 and T2 samples had been conducted in a thorough way already for the Savage et al. (2013) paper. Student-level missing data was one of the primary issues addressed with the original dataset. The missing data represented less than 5% of the data across all variables and were missing completely at random (MCAR) according to Little's MCAR test ($p > .05$). Since the missing data were MCAR, regression-based imputation procedures with reading pretest variables serving as predictors were run to replace the missing values for the full dataset at T1 and T2 (for further details, see Savage et al., 2013).

Question 3. *Does the ABRA program provide protection from long-term reading development problems for students at risk at T3?*

This question focuses on the T3 follow-up student data. With respect to the T3 data, 488 out of the original 1,181 students returned their consent forms to participate in the follow-up study. Of these, 467 students completed the T3 reading assessment measure. These 467 students (39.5% of the original participants) made up the final T3 follow-up sample. From this group, only one student did not return a completed parent questionnaire, so for this one participant, information about ethnicity and mother's education was not obtained; however, data on the student's sex and reading ability at T3 was available.

As the pupil follow-up sample obtained for this study is entirely dependent upon but also much smaller than that of the original ABRA RCT study (Savage et al., 2013), it should be noted that this smaller sample size will have implications to the power of the effects analyzed in the pupil results section. Savage et al. (2013) was sufficiently powered to detect a medium effect size in primary outcome measures of word reading and phonological awareness.

Identifying students at risk. In the original sample (N = 1,181), the lowest performing students (n = 362) were identified as those whose T1 baseline reading composite scores were in the bottom 30% on the three key literacy measures of Letter-Sound Knowledge (LSK), FRY Sight Words (FRY), and Phonological Blending Words (BW). To identify the at-risk group at T1, a latent variable based on a regression factor score of the three key reading measures (LSK, FRY, and BW) was created, which represented a composite T1 reading score measure (z-score). It needs to be noted that this z-score is only a representative of the sample and that a student could be a typically developing reader, but may have been identified as at risk if he/she was in the bottom 30% of this sample. The at-risk cut-off line in reading research has been a bit arbitrary, ranging from 20–35% based on the measure(s) used (Rathvon, 2004). For the present study, the cut-off line for the lowest performing 30% was used, since this appears to be more of the norm in the recently published response-to-intervention (R*t*I) and reading intervention literature (see for example, Al Otaiba & Fuchs, 2006; Coyne et al., 2004; Torgesen, 2000; Van der Kooy-Hofland, Bus, & Roskos, 2012).

Binary logistic regression analyses were run for the students identified at risk at T1 (n = 362). Logistic regressions were run to see if predictive variables could be identified that could distinguish those who responded to reading intervention support from those who did not. The dependent variable for this analysis was the students' percentile scores on the blending words (BW) reading measure at the time of T2 post-testing. If a student's percentile scores were below the 30th percentile at T2, she/he was identified as a non-responder or poor reader, and those scoring above the 30th percentile were considered good readers, since they were performing in the average range (Coyne et al., 2004; Torgesen, 2000; Van der Kooy-Hofland, Bus, et al., 2012).

Addressing assumptions. Since the T3 sample data was nearly complete (99.90%), a computation of missing data was not necessary. Rather, the focus of the preliminary data analysis was to address the required assumptions and conditions for running logistic regressions. Data were screened for normal distribution. Most of the variables of interest were within the acceptable limits of normality for skewness and kurtosis, less than plus or minus one (< +/- 1.0) (Leech et al., 2015). With the exception of the skewness statistic value (-1.64) and kurtosis value (3.69) for the Listening Comprehension raw score at T1, and the kurtosis value (-1.02) for the FRY sight words raw score at T1 being outside the recommended guidelines of normality. In spite of these exceptions, data transformation was not required, since logistic regression is quite robust with respect to departures from normality assumptions. Therefore, even a skewness or kurtosis of more than +1.0 or less than -1.0 does not seem to affect the results of a logistic regression, since no distributional assumptions exist for this statistical analysis (Leech et al., 2015).

However, multicollinearity between independent variables is an issue of concern when running logistic regressions, since correlation values above .50 between predictor variables can be a potential source for misleading and confusing results (Leech et al., 2015; Tabachnick & Fidell, 2013). Therefore, correlations were run between the predictor variables of interest (ethnicity, SES, sex, level of intervention, listening comprehension, and T1 reading score) to examine their degree of multicollinearity. The preliminary analysis showed that the correlation values between the independent variables ranged from -.16 to .25. Since high correlations (above .50) were not present within the data, the issue of multicollinearity was not a concern.

TIME 3 follow-up pupil sample characteristics. Data were collected in the Spring and Fall of 2012, with the original student participants now in grades three to six. The T3 sample consisted of 55 grade three students (*n* = 36 ABRA, *n* = 19 Control), 253 fourth graders (*n* = 153 ABRA, *n* = 100 Control), 116 fifth graders (*n* = 54 ABRA, *n* = 62 Control), and 43 sixth graders (*n* = 22 ABRA, *n* = 21 Control). The final student sample was comprised of 247 girls (*n* = 134 ABRA, *n* = 113 Control) and 220 boys (*n* = 115 ABRA, *n* = 105 Control). Participating students age ranged from 7.7 to 12.5 years of age (*M* = 10.33 years; *SD* = 9.35 months) at T3.

Original and T3 follow-up sample comparison. With any follow-up study, the problem of *experimental mortality* as a threat to internal validity is a concern, especially when participation attrition may possibly lead to crucial differences between the groups being compared (Creswell, 2014). To determine whether the experimental mortality of the T3 sample should be of concern, preliminary data analyses were run to compare the original 2007–2009 ABRA sample (*N* = 1181) with the 2012 T3 follow-up sample. Specifically, these analyses compared the sample of participants who consented to be part of the follow-up study (*n* = 467) versus those who did not take part in the follow-up (*n* = 714).

T3 reading measure comparisons. Independent-samples *t*-tests were conducted to verify whether the T3 sample was comparable to the remaining original sample on four key pretest literacy measures—Letter-Sound Knowledge (LSK), FRY sight word reading (FRY), Phonological Blending Words (BW), and Listening Comprehension (LC). These analyses presented in Table 2.

The results indicated that mean differences existed on the pretest literacy measure in favour of the T3 sample versus the sample of students not participating in the follow-up study. The T3 sample showed a slight advantage on Fry words, phonological blending, and listening comprehension, although these mean differences were not significant. The results indicated only a significant advantage for the T3 sample on letter-sound knowledge (LSK), two-tailed $t(1179) = 2.17$, $p = .031$. However, an effect often can be significant in a large sample, but modest in size, as indicated in Table 2 where the effect size value ($d = .13$) of the LSK measure suggests a low practical significance (Cohen, 1988). Since the significant finding was only on one measure with such a small effect size, the LSK difference finding was not considered a major methodological issue, indicating that overall the T1 literacy skills of the T3 sample were comparable.

Table 2

Comparing Pretest Literacy Means for T3 Sample versus Sample Not Part of T3

	From Original Sample (N = 1181)		Mean Difference	Cohen's *D*	*t*	*df*
	Not Part of T3 (n = 714)	T3 (n = 467)				
LSK	16.48 (8.53)	17.56 (8.11)	-1.078	.129	- 2.17*	1179
FRY	6.80 (7.14)	6.80 (7.11)	-.004	0.00	- 0.01	1179
BW	6.84 (4.28)	7.29 (4.48)	-.452	.103	-1.74	1179
LC	14.26 (2.73)	14.48 (2.58)	-.218	.083	-1.37	1179

Note. * = $p \leq .05$. Standard Deviations appear in parentheses below means.

Key: LSK = Letter Sound Knowledge measure; FRY = FRY Sight Words measure; BW = Phonological Blending subtest measure from the comprehensive test of phonological processing (CTOPP); LC = Listening Comprehension subtest from the group reading assessment and diagnostic evaluation (GRADE) measure.

Demographic variables comparisons. Next, Pearson chi-square tests were performed, which focused on the demographic categorical variables of interest—male or female status (sex), mother's education as the proxy for SES (SES), and ethnicity. The chi-square tests were run to determine whether the proportion of the participants for the key demographic variables was similar to the T1 sample who were not part of the T3 sample. The results of these analyses are reported in Table 3.

Table 3

Results of the Chi-Square Tests Comparing the Proportion of Participants at T3 versus those Not Part of the T3 Follow-up by Key Demographic Variables of Interest

(A) Sex

		From Original Sample (N = 1181)	
		[a]Not Part of T3 (n = 714)	T3 (n = 467)
Sex	Male	367 (51.4%)	220 (47.1%)
	Female	347 (48.6%)	247 (52.9%)

(B) Ethnicity

		From Original Sample (n = 812)	
		Not Part of T3 (n = 346)	T3 (n = 466)
Ethnicity	White	234 (67.6%)	343 (73.6%)
	Asian	47 (13.6%)	44 (9.4%)
	Other	65 (18.8%)	79 (17.0%)

(C) SES – Mothers Education

		From Original Sample (n = 908)	
		Not Part of T3 (n = 442)	T3 (n = 466)
SES	No Post Secondary	113 (25.6%)	111 (23.8%)
	College/ Technical	180 (40.7%)	229 (49.1%)
	University	149 (33.7%)	126 (27.0%)

Note. Numbers in parentheses indicate column percentages. [a]*Not Part of T3* refers to the students who did not participate in this follow-up study.

It should be noted that the total *N*-values of the three demographic variables varied. This variation reflects the number of responses received during the 2007–2009 ABRA study on the parent questionnaires (see Appendix C) that addressed these three demographic variables—male and female status (sex) (N = 1,181), mother's education proxy for SES (SES) (n = 908), and ethnicity (n = 812). The three sub-groups within the ethnicity variable were created in a matter that would best fit this study's data in obtaining large enough sample sizes at T3. The ethnicity *Asian* subgroup included participants who identified as one of the following -- Chinese, South Asian, Southeast Asian, Arab, Korean, Japanese, West Asian and Filipino (Statistics Canada, 2011); while the ethnicity *Other* subcategory included students who identified either Aboriginal, Latin American, or Black.

The results of the chi-square tests showed no significant difference in proportions of the T3 sample and those who did not participate in the follow-up on the basis of their sex, χ^2 (1, N = 1181) = 2.08, p = .15). No significant difference was found in the proportions of the T3 and T1 samples by ethnicity, χ^2 (2, N = 812) = 4.41, p = .11). However, some significant difference was found for the T1 sample and the T3 sample by SES, χ^2 (2, N = 908) = 7.18, $p < .05$ (See Table 3C).

In Table 3, the consented parent data illustrates that a trend exists for more college/technically educated parents to be part of the T3 sample (40.7% of the T1 sample not part of the T3 sample compared to 49.1% that were part of the T3 sample). Whereas an opposite trend exists for the number of participants with university-educated parents, since less of them consented to be part of this study (33.7% of the T1 sample were not part of T3 sample compared to 27.0% that were part of the follow-up).

T3 Follow-up Reading Assessments

The GRADE. At T3 all participating students completed a developmentally appropriate, grade-based reading assessment called the *Group Reading Assessment and Diagnostic Evaluation* (GRADE) (Williams, 2001). The GRADE has been shown to be a reliable and valid measure of early reading ability in young students (Fugate, 2003; McBride et al., 2010; Waterman, 2003). Students at T3 completed each of the reading sub-sections relevant to their grade level. The grade three students had five sub-sections to complete (Word Reading, Listening Comprehension, Vocabulary, Sentence Comprehension, and Passage Comprehension), and the older students (grades four to six) only had to do the four comprehension sub-sections, and were not required to complete a word level sub-test.

Descriptive analyses were run using the normative standardized scores of the GRADE to determine whether the T3 participants were a representative sample on reading ability. Standard scores for the students' vocabulary, reading comprehension, and total test results on the GRADE were calculated. The standard scores for each of these subscales have a mean of 100 and a standard deviation of 15. Table 4 compares the average means and the standard deviations for the whole T3 sample —the sample of students at T3 who were not at risk of reading difficulties at T1, and the sample of students at T3 who were at risk at T1.

Table 4

Sample's GRADE Standard Score Means and Standard Deviations

T3[a] GRADE Reading Measure Subtests	Full T3 Group (Group *N* = 467)		Not At Risk at T1[b] (Sub-Group *n* = 330)		At Risk at T1 (Sub-Group *n* = 137)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Vocabulary Composite[c]	102.78 (*n* = 55)	16.97	106.92 (*n* = 37)	13.65	94.28 (*n* = 18)	20.16
Vocabulary[d]	103.29 (*n* = 412)	11.73	104.76 (*n* = 293)	11.21	99.66 (*n* = 119)	12.23
Comprehension Composite[e]	98.47	12.94	101.28	12.30	91.70	11.94
Total Test[f]	99.23	12.82	102.19	11.99	92.10	11.95

Note. [a]T3 refers to the T3 follow-up sample. [b]T1 refers to T1 baseline sample. [c]Vocabulary composite standard score was for Grade 3 students only. This composite score was a combined score of the word reading and vocabulary subscales; [d]Vocabulary standard score was for Grades 4, 5, and 6 students. This standard score was based on the vocabulary subtest scores; the [e]comprehension composite standard score was for the full sample. This standard score was a combined score of the sentence and passage comprehension subscales; the [f]total test standard score was for the full sample. This standard score was a combined score of the vocabulary, sentence, and passage comprehension subscales.

As is shown in Table 4, the group means for each subtest was within one *z*-score above or below the normative mean of 100 for each of the three comparison groups. Based on the group averages for the standardized reading scores—for example, with a group mean of 99.23 (SD = 12.82) on the total test—the full T3 sample reading scores are representative of a typical sample of reading ability. At T3, when the at-risk sub-group (n = 137) at T1 was compared to their peers who were not at risk of reading difficulties at T1 (n = 330), the standardized means of the at-risk group were consistently lower on every T3 reading sub-test. One-way between subjects ANOVAs were conducted to confirm that the differences in the reading scores of these two groups were significant at T3; for example, as was shown by the results of the GRADE total test scores, F (1, 465) = 68.66, p < .001) (see Table 5).

Table 5

ANOVA and Cohen d Results Comparing the T3 Follow-up Reading Assessment Scores of the Students Identified Not At Risk versus At Risk at T1

T3[a] GRADE Reading Measure Subtests	ANOVA difference of T3 reading score means of students identified not at risk vs students identified at risk at T1			Cohen's *d*
	df	*F*	*p*	*D*
Vocabulary Composite[b]	1	7.53	.01	.693
Vocabulary[c]	1	16.66	.00	.435
Comprehension Composite[d]	1	59.77	.00	.790
Total Test[e]	1	68.66	.00	.843

Note. [a]T3 refers to the T3 follow-up sample. The [b]vocabulary composite standard score was for Grade 3 students only. The [c]vocabulary standard score was for Grades four, five, and six students. The [d]comprehension composite standard score was for the full sample. This standard score was a combined score of the sentence and passage comprehension subscales; the [e]total test standard score was for the full sample. This standard score was a combined score of the vocabulary, sentence, and passage comprehension subscales.

In addition, effect sizes (*d*) were calculated by dividing the differences in the means of these two groups by the pooled standard deviations (Tabachnick & Fidell, 2013). Cohen's effect size value ($d = .84$) for the total test scores suggested a high practical significance. These results illustrate that the students, as a group, identified more at risk at T1, and continued to be poorer readers at T3 when compared to their peers.

Main Data Analyses: Pupil-Level

Binary logistic regression analyses were carried out using SPSS Statistics (version 20) to address the three proposed pupil-level questions. This analysis was appropriate because the dependent variables for each pupil-level question were discrete (0, 1) and not continuous. Participants could only fall into two possible categories—at risk/not responding (0) or not at risk/responding (1)—based on whether they had reached the educational outcome variable threshold of achieving above the 30th percentile cut-off. The aim of these analyses was to explore how well category membership from the independent variables selected could be predicted.

Prior to running the regressions, descriptive analyses were run to compare the proportions of students by the three demographic variables of interest. Comparisons were made between the students who were identified at risk of reading difficulties versus not at risk at T1, and the proportion of students who responded to classroom intervention versus those who did not at T2 (see Table 6). Table 6 is further broken down by the level of intervention students received (Control = 0, Poor Implementation = 1, and Good Implementation classrooms = 2).

Table 6

Proportion of Students by Demographic Variables of Interest (Sex, Ethnicity, and SES) at T1 by At Risk or Not At Risk and Compared to Non-Responders and Responders at T2 and Subdivided by Level of Intervention Received

	Entire Sample						**Entire Sample**					
Predictor Variables	At Risk[a] at T1			Not at Risk[b] at T1			Nonresponders[c] at T2			Responders[d] at T2		
Sex (*N* = 1181)							**Sex (*N* = 1181)**					
	Intervention Implementation Level[e]						**Intervention Implementation Level**					
	0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
Female (%)	19	6	22	24	2	25	26	6	16	22	3	25
Male (%)	23	6	24	23	4	22	26	7	19	22	5	23
Subtotal (*n*)	362			819			117			1064		
Ethnicity (*N* = 812)							**Ethnicity (*N* = 812)**					
	Intervention Implementation Level						**Intervention Implementation Level**					
	0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
White (%)	25	4	31	36	3	36	31	4	22	33	4	35
Asian (%)	7	1	9	3	.3	6	15	0	7.5	3	.5	6.5
Other (%)	11	1	11	8	.7	7	13	1.5	6	8	1	9
Subtotal (*n*)	225			587			68			744		
SES (Mother's Ed.) (*N* = 908)							**SES (Mother's Ed.) (*N* = 908)**					
	Intervention Implementation Level						**Intervention Implementation Level**					
	0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
No postsecondary (%)	14	2	17	8	1	12	26	1	16	9	1	13
College/technical (%)	19	3	23	22	2	21	17	2	16	22	2	22
University degree (%)	10	3	9	16	3	15	12	5	5	15	2	14
Subtotal (*n*)	271			637			83			825		

Note. [a]At risk = Students at pretest who scored in the bottom 30% of the T1 reading measures; [b]Not at risk = Students at pretest who scored in the top 70% of the T1 reading measures; [c]Non-responders = Students who were at risk of reading difficulties at T2 (scored ≤ 30th percentile on the Blending words measure); [d]Responders = Students who were not at risk at T2 (scored > 30th percentile on the Blending words measure); [e]Intervention Implementation Level = the level of ABRA intervention the students received (0 = Control group, 1 = Poor Implementing Classrooms, and 2 = Good Implementing Classrooms).

Applying the same classification data used in the original Savage et al. (2013) paper, a *control classroom* was a classroom in which the students did not receive the ABRA intervention. A *poor implementation* classroom was one in which the teachers, delivering the ABRA intervention to the students, received an overall IFM score of one, entry level. In the poor implementing classrooms, evidence of teacher planning of ABRA ELA lessons was lacking, and very little teacher instructional guidance was occurring. Typically, ABRA exposure was through unstructured lessons during their free time when they were able to play and choose their own activities.

A *good implementation* classroom was a classroom in which the teacher delivering the ABRA intervention to the students received an overall IFM score of 2 (Adoption Level), 3 (Adaption Level), or 4 (Differentiated Adaptation Level). In the good implementing classrooms, clear evidence of basic teacher planning and delivery of structured ABRA lessons was present. Students received guidance for navigating through the program, and direction on what activities to explore. The ABRA activities to which the students were exposed were aligned with the early literacy skills they were being taught during their ELA lessons.

From this preliminary analysis of the whole sample, a few patterns emerged. When looking at the proportions of students by ethnicity, an increased risk seems to exist for pupils of the Asian community—17% identified at risk with respect to the T1 reading measure, and 22.5% identified as not responding to a classroom intervention at T2. In contrast, the proportions of pupils of white (60% to 57%) or other (23% to 20.5%) ethnic origins marginally decreased in the at risk/non-responding groups at T2.

The proportion of students from homes with a lower SES proxy (mothers with no postsecondary education) appears to be at an increased risk of having reading difficulties at T2. Pupils whose mothers had no postsecondary education moved from 33% of the T1 sample identified as at risk to 43% of the sample of non-responders at T2. In contrast, the proportion of T1 students at risk whose mothers had a college/technical education decreased at T2 (from 45% to 35%), and the proportion of students at risk whose mothers had university degrees stayed the same between T1 and T2.

Identifying students at risk of reading difficulties at T1. To identify the T1 sample at risk of reading difficulties, a *latent variable* was created representing a composite pretest reading score measure (*z*-score). This latent variable was a regression factor score created by combining the raw scores from multiple T1 literacy measures—letter-sound knowledge (LSK), blending words (BW), and sight word reading (FRY). In our samples T1 results, these literacy variables were highly correlated, with KMO = .74 and a significant Bartlett's test of sphericity (χ^2 (3, N = 1181) = 1685.81, $p < .001$); therefore, combining the LSK, BW and FRY variables to create a *z-score* produced a latent variable that was more reliable, since the measure is clustered across the three variable scores. Thus, this clustering removed a certain amount of measurement error (Tabachnick & Fidell, 2013).

From the original sample (N = 1,181), the lowest performing students (n = 362) were identified as those whose T1 reading composite scores were in the bottom 30% on the three key literacy measures (LSK, BW, and FRY subtests) (Al Otaiba & Fuchs, 2006). Binary logistic regression analyses were then run for the students identified at risk at T1 (n = 362). These logistic regressions were run to determine whether predictive variables that could distinguish those students who responded to reading intervention support from those who did not could be identified.

Identifying responders to intervention. The dependent variable for this analysis was the students' percentile scores on the blending words (BW) reading measure at T2 post-testing. If the students' percentile scores were below the 30^{th} percentile at T2, they were identified as non-responders or poor readers, and those scoring above the 30th percentile were considered good readers, since they were performing in the average range (Coyne, Kame'enui, Simmons, & Harn, 2004; Torgesen, 2000).

Results for the Pupil-Level Research Questions

Results for pupil-level question 1. *Are there identifiable pupil-level variables of at risk students who had a greater response to Tier 1 (classroom-level) intervention at T2 post-testing?*

To establish whether the pupil demographic variables (ethnicity, SES, and sex-differences) have any predictive validity beyond the pupil TIME 1 baseline meta-cognitive measures (reading skills and listening comprehension), a series of *stepwise* logistic regressions were run. A stepwise method of analysis was used because the researcher wished to test general models of influence in order to examine the potential of pupil-level factors influencing reading outcomes (Tabachnick & Fidell, 2013). The variables of interest were entered stepwise with binary reading measure scores as the dependent variable.

The dependent variable for this analysis measured whether a student responded to a classroom intervention by scoring above the 30th percentile on the blending words task. *Yes* equals 1 (the reference category) and *No* equals 0. *No* indicates that the student scored below the 30th percentile and thus did not respond to a classroom intervention.

The T1 meta-cognitive variables were entered in the first step of analysis. At the second step, each demographic variable of interest was entered, and then the ABRA intervention condition in Step 3. Each demographic variable was broken down into subgroups, with one subgroup from each the three main demographic being identified as the reference category for analysis (see Appendix J). For the Sex variable, the subgroups were *Male* (as the reference category) and *Female*. The three subcategories within the Ethnicity variable were *White* (as the reference category), *Asian* (which included participants who identified as one of the following -- Chinese, South Asian, Southeast Asian, Arab, Korean, Japanese, West Asian and Filipino), and *Other* (which included students who identified as Aboriginal, Latin American, Black or Other). The three subcategories making up the proxy for SES variable included the students with mothers with *No postsecondary* education (as the reference category), *College* (which included students with mothers with postsecondary technical training education), and *University (*for students with mothers with a Bachelor's degree or higher from a university*)*.

As noted previously, when describing the variables in Table 6, one of the key R*t*I influencing factors of interest was the level of exposure the students had to the ABRA intervention during their time in the 2007–2009 ABRA RCT study. In Table 7, and in subsequent pupil-level regression result tables, IFM (Con) refers to those students who were part of the control group condition and did not receive the ABRA intervention; IFM (Poor) refers to those students who were part of an intervention classroom where ABRA was poorly implemented, and the implementing teacher received an overall IFM score of 1; IFM (Good) refers to students who were part of an intervention classroom where the teacher implementing ABRA received an IFM score of 2, 3, or 4 (for additional details about the IFM criteria, see Appendix K, ABRA Implementation Rubric). The proportion of the whole sample (N = 1,181) represented in each of the ABRA conditions is as follows: 45% of the students were part of the control group classrooms, 8% of the students were part of classrooms where ABRA was poorly implemented, and 47% of the whole sample were part of classrooms where ABRA was implemented well.

To test the influence of step of entry in order to explore the unique variance explained, the researcher reran the analysis by rearranging the stepwise order of the demographic variables and the level of the intervention variables.

Table 7

Logistic Regression Results for the Whole Sample at T1

Dependent Variable (Reading 2) = Poor vs Good readers on T2 reading skills measure (below or above 30th percentile on blending words). Poor (non-responder) = 0, Good (responder) = 1

Predictor Variables		B	Wald	*P*	Odds Ratio Exp (B)	Nagelkerke R^2	H-L fit
A. Sex (*n* = 1181)							
Step 1	Reading1 (At risk)	-1.957	77.690	.000	.141	16.1%	.882
	LC1	.012	.121	.727	1.012		
	Constant	2.988	29.387	.000	19.850		
Step 2	Sex (F)	.025	.014	.904	1.025	16.1%	.654
	Constant	2.981	28.855	.000	19.700		
Step 3	IFM (Con)	--	7.154	.028	--	17.3%	.909
	IFM (Poor)	-.005	.000	.988	.995		
	IFM (Good)	.570	6.551	.010	1.768		
	Constant	2.621	20.863	.000	13.756		
Step 2	IFM (Con)	--	7.166	.028	--	17.3%	.968
	IFM (Poor)	-.005	.000	.988	.995		
	IFM (Good)	.570	6.560	.010	1.768		
	Constant	2.624	21.123	.000	13.785		
Step 3	Sex (F)	.007	.001	.971	1.007	17.3%	.909
	Constant	2.621	20.863	.000	13.756		
Step 4	Interaction					17.4%	.771
Sex * IFM (Con)		--	.275	.871	--		
Sex * IFM (Poor)		.101	.023	.878	1.107		
Sex * IFM (Good)		.233	.275	.600	1.262		
Constant		2.675	20.864	.000	14.513		
B. Ethnicity (*n* = 812)							
Step 1	Reading1 (At risk)	-1.841	42.839	.000	.159	14.3%	.686
	LC1	.030	.469	.494	1.030		
	Constant	2.811	17.781	.000	16.635		
Step 2	Ethnic (W)	--	3.886	.143	--	15.2%	.251
	Ethnic (A)	-.697	3.860	.049	.498		
	Ethnic (O)	-.126	.136	.713	.882		
	Constant	3.189	20.154	.000	24.267		
Step 3	IFM (Con)	--	8.410	.015	--	17.5%	.908
	IFM (Poor)	.398	.448	.503	1.475		
	IFM (Good)	.830	8.402	.004	2.293		
	Constant	2.557	12.037	.001	12.894		
Step 2	IFM (Con)		8.131	.017		16.5%	.830
	IFM (Poor)	.398	.474	.491	1.489		
	IFM (Good)	.807	8.116	.004	2.241		
	Constant	2.228	10.165	.001	9.283		

.

Dependent Variable (Reading 2) = Poor vs Good readers on T2 reading skills measure (below or above 30th percentile on blending words). Poor (non-responder) = 0, Good (responder) = 1

Predictor Variables		B	Wald	*P*	Odds Ratio Exp (B)	Nagelkerke R^2	H-L fit
Step 3	Ethnic (W)		4.237	120		17.5%	.908
	Ethnic (A)	-.731	4.163	.041	.481		
	Ethnic (O)	-.097	.080	.777	.907		
	Constant	2.557	12.037	.001	12.894		
Step 4	Interaction					18.6%	.456
Ethnic (W) * IFM (Con)		--	2.070	.723	--		
Ethnic (A) * IFM (Poor)		20.784	.000	.999	1.062E9		
Ethnic (A) * IFM (Good)		1.048	1.963	.161	2.851		
Ethnic (O) * IFM (Poor)		-.013	.000	.993	.987		
Ethnic (O) * IFM (Good)		.455	.376	.540	1.576		
Constant		2.405	10.303	.001	11.079		
C. SES (*n* = 908)							
Step 1	Reading 1 (At-risk)	-1.886	52.621	.000	.152	14.7%	.737
	LC1	.007	.031	.861	1.007		
	Constant	3.091	23.333	.000	22.005		
Step 2	No Post Sec.	--	8.544	.014	--	16.5%	.877
	College	.754	7.434	.006	2.125		
	University	.676	4.471	.034	1.966		
	Constant	2.690	16.837	.001	14.725		
Step 3	IFM (Con)	--	7.097	.029	--	18.1%	.707
	IFM (Poor)	.086	.035	.852	1.090		
	IFM (Good)	.682	6.907	.009	1.978		
	Constant	2.248	10.890	.001	9.466		
Step 2	IFM (Con)	--	6.432	.040	--	16.1%	.936
	IFM (Poor)	.136	.089	.765	1.146		
	IFM (Good)	.646	6.338	.012	1.908		
	Constant	2.695	16.592	.000	14.806		
Step 3	No Post Sec.	--	9.285	.010	--	18.1%	.707
	College	.784	7.888	.005	2.190		
	University	.738	5.181	.023	2.093		
	Constant	2.248	10.890	.001	9.466		
Step 4	Interaction					19.2%	.122
No Post Sec * IFM (Con)		--	4.621	.328	--		
College * IFM (Poor)		-1.499	1.181	.277	.223		
University * IFM (Poor)		-2.023	2.421	.120	.132		
College * IFM (Good)		-.788	1.877	.171	.455		
University * IFM (Good)		-.035	.002	.962	.966		
Constant		1.907	7.341	.007	6.735		

Note. Odds Ratio Exp(B) explains the changes in the dependent variable (DV) due to changes in the independent variable (IV). Nagelkerke R Squared refers to the proportion of variance in the DV that is explained by the covariates. H-L fit is the Hosmer-Lemeshow goodness of fit measure.

As illustrated in Table 7, when the demographic variables were entered in Step 2, the ABRA condition was entered in Step 3; and when the ABRA condition was entered in Step 2, the demographic variables of interest were entered in Step 3. The reasoning for examining both models was to confirm whether the demographic variables of interest produced a unique effect when they were entered in the model after the intervention effect was first considered.

Separate stepwise regressions, rather than simultaneous group regressions, were run for each of the three demographic variables (ethnicity, SES, and sex-differences), since the researcher was interested in each variables individual rather than group influence on predicting the probability of a student being or not being at risk for reading difficulties. Since little research has been reported in R*t*I literature on pupil-level demographic factors, placing the demographic factors in Step 2 of the analyses allowed an examination of whether the demographic variables of interest could predict above and beyond the pupil-level cognitive variables.

Before looking at the T1 at-risk group of readers (bottom 30%), this researcher began the analyses by first examining the overall effect of the whole sample data (N = 1,181). As mentioned previously, by controlling for the meta-cognitive measures in the first step, the goal was to see if each of the demographic variables of interest would add to the prediction of whether students would be identified as responders and have stronger reading skills at T2.

The first section of Table 7, Steps 1 and 2 (demographic variable), summarizes the whole sample findings with respect to Question 1. For this analysis, it did not matter whether the students received regular classroom teaching or the ABRA intervention, since it was the influence of sex-differences, ethnicity, and mother's education as the proxy of SES (SES) on students' reading skill development that was of interest.

Step 1 for each of the three demographic variable equations shows that student baseline early reading skills was a significant predictive variable of growth in blending word attainment at T2, while listening skills did not predict growth in blending skills across the whole sample.

Step 2 (demographic variable) under (A) Sex shows that no significant sex differences were found at T2 (p = .904). With respect to (B) Ethnicity, the results of the regression analysis indicated that for the whole sample, ethnicity had an effect. A 1% change in the variance of the dependent variable can be explained by the influence of ethnicity (Nagelkerke's R^2 increases from 14.3% to 15.2%). With the whole sample, the odds ratio of being in the responders group at T2 decreased by 0.50 for the Ethnic Asian variable with a 95% confidence interval of [0.25, 1.00].

In Table 7, Step 2 (demographic variable) under (C) mother's education as a proxy of SES indicates that mother's education also improved the prediction of blending scores at T2 (Nagelkerke's R^2 change = 1.8%). The odds ratio for the College coefficient is 2.13 with a 95% confidence interval between 1.24 and 3.65, whereas for the University variable, the likelihood of being in the responders group and not in the non-responders group increased by 1.05 and 3.68 times. This result suggests that the odds of having stronger reading skills at T2 are increased for the students with college and university educated mothers, since this group is twice as likely to respond to classroom intervention compared to students whose mothers had no post-secondary education.

When examining the overall regression models in Table 7, the Nagelkerke R^2 statistic indicated that 17 to 19% of the dependent variable, being in the responders group at T2, is attributable to the predictors in the model.

Results for pupil-level question 2. *Does a teacher's level of implementation of ABRA ELA lessons during the Tier 1 intervention stage impact the influence of pupil-level variables on student learning at T2? In addition to possible main effects, is there an interaction effect between the level of ABRA implementation students received and the demographic variables of interest at T2?*

Whole sample. Steps 3 (IFM) and 4 of Table 7 address this second pupil-level question. First, the complete dataset was used to determine whether ABRA had an overall effect on the whole sample. The Step 3 results, for each of the demographic variable regression models run with the whole sample, indicated that the ABRA intervention had a main effect at T2 as reported in Savage et al. (2013). For example, Step 3 under the (B) Ethnicity model shows that the chances of being in the responders group and not in the non-responders group increased by 2.29 times according to the odds ratio for the IFM (Good) variable, with a 95% confidence interval of 1.31 and 4.02 times. Overall, the odds ratios for the IFM (Good) coefficients for all three of the demographic variable models in Table 7 indicated that students from classrooms where the intervention was well implemented were on average twice as likely to have stronger blending skills at T2 when compared to students from the control group. However, the results of Step 4 for each of the three demographic variables showed that no evident interaction effect was present when looking at the whole sample data.

At-risk sample. The same analyses were rerun to examine the influence of demographic variables on those students identified at risk at baseline (n = 362). Table 8 shows the proportions of the responders and non-responders of the available T3 follow-up sample at T2 from the at-risk group at T1. The proportion of the at-risk sample represented in each of the ABRA conditions was as follows: 42% of the students were part of the control group classrooms, 12% of the students were part of the classrooms where ABRA was poorly implemented, and 46% of the students of whole sample were part of the classrooms where ABRA was implemented well.

One notable finding at T2 was that there were proportionally 7% less college educated non-responders in comparison to treatment responders. Furthermore, there were higher proportions of responder students with college-educated mothers for both the control condition (8% increase) and the good implementer condition (7% increase) in comparison to the non-responding students. This result indicates a similar risk for both the control and good implementation conditions.

Table 8

At-Risk Proportion Descriptives for T3 Sample Compared by Demographic Variable, Intervention Condition and if Participants were Responders or Non-Responders at T2

	T1 Entire At-risk Sample			**Responders vs Non-Responders at T2 from T1 At-risk sample**					
Predictor Variables	At risk at Pretest			Non-Responders			Responders		
Sex (*n* = 362)				**Sex (*n* = 362)**					
	Intervention Implementation Level			**Intervention Implementation Level**					
	0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
Female (%)	19	6	22	21	6	12	18	6	25
Male (%)	23	6	24	29	9	23	21	6	24
Subtotal (*n*)	362			84			278		
Ethnicity (*n* = 225)				**Ethnicity (*n* = 225)**					
	Intervention Implementation Level			**Intervention Implementation Level**					
	0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
White (%)	25	4	31	24	2	20	25	5	34
Asian (%)	7	1	9	17	0	11	4	1	8
Other (%)	11	1	11	15	2	9	10	1	12
Subtotal (*n*)	225			46			179		
SES (Mother's Ed.) (*n* = 271)				**SES (Mother's Ed) (*n* = 271)**					
	Intervention Implementation Level			**Intervention Implementation Level**					
	0	1	2	0	1	2	0	1	2
No postsecondary (%)	14	2	17	31	2	17	10	2	17
College/technical (%)	19	3	23	12	2	17	20	3	24
University degree (%)	10	3	9	10	5	4	10	3	11
Subtotal (*n*)	271			58			213		

Next, Table 9 provides a summary of the logistic regressions that were run with the at-risk group to examine the influence of demographic variables as predictors of reading success at T2. Similar to the results that were found when looking at the whole sample in Table 7, no clear effects of sex-differences were found on T2 blending ability. The independent main effects for ethnicity, SES, and ABRA implementation remained significant for the at-risk subsample at T2. As shown in Step 2 (demographic variable) under (B) Ethnicity, an increased change in the variance of the dependent variable can be explained by the ethnic differences in the at-risk sample (Nagelkerke R^2 change = 3.8% increase). This finding shows that the odds ratio of being in the responders group at T2 decreased by 0.36 for the Ethnic Asian variable, with a 95% confidence interval of likelihood between 0.15 and 0.84.

In Table 9, Step 2 under (C) mother's education as a proxy of SES suggests that a mother's education continues to be a significant predictor of the blending scores of the students at risk of reading difficulties at T2 (Nagelkerke R^2 change = 4.7%). For example, the results for the students in the at-risk group indicated that the college variable increased the likelihood of students being in the responders group and not in the non-responders group by 1.32 and 5.01 times. However, although both the College and the University coefficients were statistically significant for the whole sample, only the College coefficient for SES for the at-risk sample was significantly related to T2 blending word outcomes, X^2 (1, N = 271) = 7.67, p < .01. Nonetheless, the trend found in the at-risk data continues to support the finding that the odds of possessing stronger reading skills at T2 are greater for students with mothers who have some post-secondary education.

Table 9

Logistic Regression Results for Participants who were At Risk at T1

Dependent Variable (Reading 2) = Poor vs Good readers on T2 reading measure (below or above 30th percentile on blending words measure). Poor (non-responder) = 0, Good (responder) = 1

Predictor Variables		B	Wald	*p*	Odds Ratio Exp (B)	Nagelkerke R^2	H-L fit
A. Sex (*n* = 362)							
Step 1	LC1	.025	.374	.541	1.025	0.2%	.271
	Constant	.867	2.467	.116	2.379		
Step 2	Sex (F)	.412	2.619	.106	1.509	1.3%	.397
	Constant	.752	1.805	.179	2.121		
Step 3	IFM (Con)	--	5.718	.057	--	3.7%	.846
	IFM (Poor)	-.074	.038	.846	.929		
	IFM (Good)	.605	4.828	.028	1.831		
	Constant	.405	.477	.490	1.499		
Step 2	IFM (Con)	--	5.813	.055	--	2.6%	.167
	IFM (Poor)	-.061	.026	.871	.941		
	IFM (Good)	.611	4.966	.026	1.842		
	Constant	.523	.822	.365	1.688		
Step 3	Constant	.407	2.519	.112	1.502	3.7%	.846
	Sex (F)	.405	.477	.490	1.499		
Step 4	Interaction						
Sex * IFM (Con)		--	.905	.636	--	4.1%	.255
Sex * IFM (Poor)		.440	.330	.566	1.552		
Sex * IFM (Good)		.506	.808	.369	1.659		
Constant		.536	.786	.375	1.709		
B. Ethnicity (*n* = 225)							
Step 1	LC1	.056	1.219	.269	1.057	0.8%	.174
	Constant	.630	.876	.349	1.877		
Step 2	Ethnic (W)	--	5.692	.058	--	4.6%	.805
	Ethnic (A)	-1.027	5.595	.018	.358		
	Ethnic (O)	-.448	1.214	.271	.639		
	Constant	1.375	3.336	.068	3.956		
Step 3	IFM (Con)	--	5.235	.073	--	8.2%	.346
	IFM (Poor)	.902	1.244	.265	2.465		
	IFM (Good)	.776	4.757	.029	2.172		
	Constant	.753	.907	.341	2.124		
Step 2	IFM (Con)	--	5.246	.073	--	4.5%	.697
	IFM (Poor)	.932	1.364	.243	2.541		
	IFM (Good)	.753	4.676	.031	2.123		
	Constant	.073	.010	.919	1.076		

Dependent Variable (Reading 2) = Poor vs Good readers on T2 reading measure (below or above 30th percentile on blending words measure). Poor (non-responder) = 0, Good (responder) = 1

Predictor Variables		B	Wald	*p*	Odds Ratio Exp (B)	Nagelkerke R^2	H-L fit
Step 3	Ethnic (W)	--	5.684	.058	--	8.2%	.346
	Ethnic (A)	-1.038	5.643	.018	.354		
	Ethnic (O)	-.414	1.004	.316	.661		
	Constant	.753	.907	.341	2.124		
Step 4	Interaction						
Ethnic (W) * IFM (Con)		--	1.662	.798	--	10.4%	.424
Ethnic (A) * IFM (Poor)		20.885	.000	.999	1.176		
Ethnic (A) * IFM (Good)		.994	1.179	.278	2.703		
Ethnic (O) * IFM (Poor)		-1.045	.371	.542	.352		
Ethnic (O) * IFM (Good)		.291	.114	.735	1.338		
Constant		.571	.473	.492	1.770		
C. SES (*n* = 271)							
Step 1	LC1	.024	.261	.609	1.024	0.1%	.313
	Constant	.983	2.379	.123	2.672		
Step 2	No Post Sec.	--	8.269	.016	--	4.8%	.536
	College	.944	7.669	.006	2.569		
	University	.706	3.047	.081	2.025		
	Constant	.588	.795	.373	1.800		
Step 3	IFM (Con)		4.207	.122		7.2%	.415
	IFM (Poor)	.246	.193	.661	1.278		
	IFM (Good)	.657	4.200	.040	1.929		
	Constant	.200	.084	.772	1.222		
Step 2	IFM (Con)	--	4.006	.135	--	2.5%	.500
	IFM (Poor)	.286	.273	.601	1.331		
	IFM (Good)	.630	4.006	.045	1.877		
	Constant	.621	.875	.350	1.860		
Step 3	No Post Sec.	--	8.489	.014	--	7.2%	.415
	College	.958	7.758	.005	2.606		
	University	.752	3.349	.067	2.122		
	Constant	.200	.084	.772	1.222		
Step 4	Interaction					10.8%	.858
No Post Sec * IFM (Con)		--	6.429	.169	--		
College * IFM (Poor)		-1.332	.678	.410	.264		
University * IFM (Poor)		-2.059	2.069	.150	.128		
College * IFM (Good)		-1.357	3.555	.059	.258		
University * IFM (Good)		.107	.011	.915	1.113		
Constant		-.290	.156	.693	.749		

Note. Nagelkerke R^2 refers to the proportion of variance in the dependent variable (DV) that is explained by the covariates; H-L fit refers to the Hosmer-Lemeshow goodness of fit measure.

As mentioned previously, the independent main effect of the ABRA intervention for the students at risk of reading difficulties was significant (in Table 9, see Steps 3 [IFM]). Step 4 for each of the three demographic variables indicated that no clear interaction effects were evident for the at-risk subgroup, as also was the case for the whole sample. So even though no clear interaction effects were associated with the ABRA intervention, ABRA does appear to work equally well for both students at risk of reading difficulties and typically developing readers, despite sex-differences, ethnicity, and SES.

As noted earlier, to be thorough, the researcher reran the analyses by rearranging the stepwise order of variables to confirm whether demographic variable effects would still be present. So in Table 9, these additional analyses are provided—when the ABRA condition variable was entered in Step 2, the demographic variables of interest were entered in Step 3. The findings of these additional analyses illustrate that even after controlling for treatment quality, an overall effect of ethnicity and of mother's education still persist. The analyses of the full sample and the at-risk subsample data results suggests that an independent effect of treatment, mother's education, and ethnicity exists with respect to literacy outcomes at T2, although no interaction effect exists between the ABRA treatment and the demographic variables. These findings imply that even though the ABRA intervention has an independent main effect, the intervention does not appear to have an added impact on the odds of pupil-level demographic variables predicting which students, at risk of reading difficulties at T1, will respond to the intervention at T2.

Results for pupil-level question 3. *Does the ABRA program provide protection from long-term reading development problems at T3 for students at risk of reading difficulties at T1?*

Preliminary descriptive analyses were run to determine how many participants of the original T1 sample were available at T3 follow-up. To address this question, the number of students who were responders at T2 (n = 278) was identified. The T2 responders were then subdivided by the number of students who were or were not at risk at T3 (see Table 10).

A student was identified as an at-risk reader at T3 if their overall percentile score on the GRADE Total Comprehension score was below the 30th percentile. With respect to the complete T3 sample (N = 467), under a quarter of these students (n = 109) were at risk of reading difficulties at baseline, then responded to classroom intervention at T2 and were available at T3.

Table 10

Regression of Sample Size from the Total Sample at T1 to the Remaining Original At-Risk Sample at T3 by the Intervention Condition Group

		1. Total Sample at T1	2. At-risk sample at T1[a]	3. At-risk at T1 and responders at T2[b]	4. At-risk at T1, responders at T2 and part of T3 follow-up	4a) Typical Reader at T3[c]	4b) At-Risk Reader at T3[d]
						56 (51%) (*Inoculation*)	53 (49%) (Insulin)
		N = 1181	n = 362	n = 278	n = 109	n = 109	
Intervention Condition	Control	535 (45%)	151 (42%)	109 (39%)	40 (37%)	25 (45%)	15 (28%)
	Poor	96 (8%)	45 (12%)	32 (12%)	9 (8%)	4 (7%)	5 (9%)
	Good	550 (47%)	166 (46%)	137 (49%)	60 (55%)	27 (48%)	33 (62%)

Note. Numbers in parentheses indicate within column percentages.
[a] Baseline (T1) students at risk of reading difficulties had a reading composite score in the bottom 30% on three measures (Letter Sound Knowledge, FRY Words, and Blending Words subtests). [b]Responders at T2 were the students identified at risk at T1, but who had scores above the 30th percentile on the blending words measure at immediate post-test. [c]Typical reader at T3 follow-up had a score above the 30th percentile on the GRADE Total Comprehension measure. [d]At-risk reader at T3 follow-up had a score below the 30th percentile on the GRADE Total Comprehension measure.

From the T3 sample, 56 students (51%) remained responders, since they scored as typical readers on the T2 comprehension measure. Fifty-three students, (49%) of those who had scored as typical readers at T2, fell back to being at-risk readers at T3. Due to the reduced T3 sample, the following findings are reported but need to be interpreted cautiously, as the smaller cell sizes for some of the analyses could influence the goodness of fit measure of the logistic regression models ran (Leech et al., 2015) (see Table 11).

The possibility of small cell sizes impacting the T3 follow-up findings may be even more evident when examining Table 11, which illustrates the proportion of students at T3 subdivided by the demographic variables of interest and by the classroom intervention condition at T1. However, the researcher decided to run the analysis and share the findings to be consistent, while at the same time acknowledging that this analysis is likely to be very conservative.

For the T3 analysis, the dependent variable (Reading 3) measured whether a student was a typical reader based on their overall percentile scores on a standardized comprehension measure. A student was considered a typical reader if they scored above the 30th percentile on the T3 comprehension reading measure. *Yes* is equal to 1, and *No* is equal to 0.

To address the question of ABRA's long-term influence, binary logistic regressions were again conducted to evaluate how well the variables of interest predicted long-term reading outcomes. However, to ensure that models with sound goodness of fit measures were reported on, some modifications had to be made on the logistic regression models that were run. These modifications will be described later in this chapter when discussing the adjustments made to specific variables of interest in order to obtain better fitting models.

Table 11

Proportion of Students who were Either Typical or Atypical (At-Risk) Readers at T3--Proportions by Demographic Variables (Sex, Ethnicity, SES) and by Level of Intervention

	Good vs Poor Readers at T3 from whole population who were part of T3 follow-up						**Good vs Poor Readers at T3 from the T2 Responders subgroup**					
Predictor Variables	Typical Readers			Atypical (At-risk) Readers			Typical Readers (Inoculation)			Atypical (At-risk) (Insulin)		
Sex (*n* = 467)							**Sex (*n* = 137)**					
	Intervention Implementation Level						**Intervention Implementation Level**					
	0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
Female	94	7	90	19	2	35	18	3	20	10	2	19
	29%	2%	27%	14%	1%	25%	27%	4.5%	30%	14%	3%	27%
Male	71	11	56	34	4	44	13	1	11	15	4	21
	22%	3%	17%	25%	3%	32%	20%	1.5%	17%	21%	6%	30%
Subtotal (*n*)	329			138			66			71		
Ethnicity (*n* = 466)							**Ethnicity (*n* = 136)**					
	Intervention Implementation Level						**Intervention Implementation Level**					
	0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
White	124	11	116	40	3	49	19	2	21	15	3	22
	38%	3%	35%	29%	2%	35%	29%	3%	32%	21%	4%	31%
Asian	13	1	12	5	2	11	6	0	3	3	2	6
	4%	.3%	3.7%	4%	1%	8%	9%	0%	5%	4%	3%	9%
Other	27	6	18	8	1	19	5	2	7	7	1	12
	8%	2%	6%	6%	1%	14%	8%	3%	11%	10%	1%	17%
Subtotal (*n*)	328			138			65			71		
SES (Mother's Ed.) (*n* = 466)							**SES (Mother's Ed.) (*n* = 136)**					
	Intervention Implementation Level						**Intervention Implementation Level**					
	0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
No post-secondary	26	4	30	18	2	31	8	1	7	11	2	19
	8%	1%	9%	13%	1%	22%	12%	1.5%	11%	15%	3%	27%
College/technical	85	5	70	28	1	40	16	1	17	10	1	17
	26%	2%	21%	20%	1%	29%	25%	1.5%	26%	14%	1%	24%
University degree	53	9	46	7	3	8	6	2	7	4	3	4
	16%	3%	14%	5%	2%	6%	9%	3%	11%	6%	4%	6%
Subtotal (*n*)	328			138			65			71		

To examine this last pupil-level question, a similar stepwise pattern was used in the set-up to address the first two pupil-level questions, however with minor modifications. The dependent variable was the reading comprehension percentile cut-off score at T3 (as discussed previously). The meta-cognitive baseline measures of early reading skills and listening comprehension were placed in Step 1. The demographic variables (sex, ethnicity, and SES) were added in Step 2. The ABRA level of intervention condition (IFM) variable was added in Step 3, and the interactions between the demographic variables and the ABRA implementation condition were added in Step 4 (see Table 12).

Similar to the earlier logistic regression result tables, Steps 2 and 3 were repeated with the intervention condition variable in Step 2 and the demographic variables of interest in Step 3. Furthermore, all analyses were run and reported on twice. The first group of analyses looked at the entire follow-up sample (*N* = 467) (Table 12), and the second group of analyses reported on the regression results of the T1 sub-sample at risk of reading difficulties (*n* = 137) (Table 13).

Whole follow-up sample. Starting with the entire follow-up sample, the preliminary analyses produced poor fitting models. This analysis had a binary covariate for reading similar to that in Question 1 and 2, however, that the Hosmer-Lemeshow goodness of fit measure was significant ($p < .05$). To best address the long-term intervention effects with sound regression models, additional analyses were run with some minor adjustments to the variables of interest used. For example, for the Reading 1 variable, rather than using the binary variable for the latent scores for the three key T1 reading measures (0 = bottom 30% of scores; 1 = top 70% of scores), the continuous latent regression scores were used as a variable in Step 1. Reading 1 (Not Binary - NB) is the variable name representing the non-binary, composite pretest reading scores (see Table 12, Step 1).

Table 12

Analyses Results of the Whole Sample of Students who were part of the T3 Follow-up

Dependent Variable (Reading 3) *= If a student is a typical reader at T3 based on their overall percentile scores on a standardized comprehension measure (above or below the 30th percentile). 0 = No (at risk), 1 = Yes (typically developing reader)*

Predictor Variables		B	Wald	*p*	Odds Ratio Exp (B)	Nagelkerke R^2	H-L fit
A. Sex (*n* = 467)							
Step 1	Reading1 (NB)*	.887	39.960	.000	2.427	23.2%	.599
	LC 1	.198	21.621	.000	1.219		
	Constant	-1.827	8.821	.003	.161		
Step 2	Sex (F)	.681	8.995	.003	1.976	25.5%	.150
	Constant	-2.049	10.814	.001	.129		
Step 3	IFM (Con)	--	6.262	.044	--	27.1%	.644
	IFM (Poor)	.583	1.029	.310	1.789		
	IFM (Good)	-.472	3.969	.046	.624		
	Constant	-1.780	7.456	.006	.169		
Step 2	IFM (Con)	--	4.938	.085	--	24.5%	.209
	IFM (Poor)	.527	.851	.356	1.694		
	IFM (Good)	-.407	3.066	.080	.666		
	Constant	-1.582	5.999	.014	.206		
Step 3	Sex (F)	.741	10.323	.001	2.098	27.1%	.644
	Constant	-1.780	7.456	.006	.169		
Step 4	Interaction					27.2%	.526
Sex * IFM (Con)		--	.112	.945	--		
Sex * IFM (Poor)		-.348	.088	.767	.706		
Sex * IFM (Good)		.040	.007	.932	1.041		
Constant		-1.788	7.383	.007	.167		
B. Ethnicity (*n* = 466)							
Step 1	Reading1 (NB)*	.894	40.382	.000	2.445	23.4%	.561
	LC 1	.197	21.440	.000	1.218		
	Constant	-1.820	8.744	.003	.162		
	Ethnic (W)	--	.611	.737	--	23.5%	.485
Step 2	Ethnic (A)	.089	.055	.815	1.093		
	Ethnic (O)	-.207	.487	.485	.813		
	Constant	-1.855	8.616	.003	.156		
Step 3	IFM (Con)	--	4.855	.088	--	24.8%	.280
	IFM (Poor)	.574	1.003	.317	1.775		
	IFM (Good)	-.398	2.792	.095	.678		
	Constant	-1.619	5.976	.015	.198		

***Dependent Variable (Reading 3)** = If a student is a typical reader at T3 based on their overall percentile scores on a standardized comprehension measure (above or below the 30th percentile). 0 = No (at risk), 1 = Yes (typically developing reader)*

Predictor Variables		B	Wald	*p*	Odds Ratio Exp (B)	Nagelkerke R^2	H-L fit
Step 2	IFM (Con)	--	4.791	.091	--	24.6%	.283
	IFM (Poor)	.541	.895	.344	1.718		
	IFM (Good)	-.395	2.883	.090	.674		
	Constant	-1.585	6.021	.014	.205		
Step 3	Ethnic (W)	--	.691	.708	--	24.8%	.280
	Ethnic (A)	.071	.035	.852	1.074		
	Ethnic (O)	-.230	.591	.442	.794		
	Constant	-1.619	5.976	.015	.198		
Step 4	Interaction						
Ethnic (W) * IFM (Con)		--	4.810	.307	--	26.1%	.220
Ethnic (A) * IFM (Poor)		-1.713	1.075	.300	.180		
Ethnic (A) * IFM (Good)		-.609	.567	.451	.544		
Ethnic (O) * IFM (Poor)		.187	.016	.898	1.205		
Ethnic (O) * IFM (Good)		-1.124	3.027	.082	.325		
C. SES (*n* = 466)							
Step 1	Reading1 (NB)[a]	.894	40.382	.000	2.445	23.4%	.561
	LC 1	.197	21.440	.000	1.218		
	Constant	-1.820	8.744	.003	.162		
Step 2	Mom Ed (PS)[b]	.623	6.058	.014	1.865	24.9%	.169
	Constant	-2.138	11.120	.001			
Step 3	IFM (Con)	--	4.303	.116	--	26.0%	.655
	IFM (Poor)	.529	.839	.360	1.697		
	IFM (Good)	-.374	2.546	.111	.688		
	Constant	-1.900	8.016	.005	.150		
Step 2	IFM (Con)	--	4.791	.091	--	24.6%	.283
	IFM (Poor)	.541	.895	.344	1.718		
	IFM (Good)	-.395	2.883	.090	.674		
	Constant	-1.585	6.021	.014	.205		
Step 3	Mom Ed (PS)	.599	5.544	.019	1.820	26.0%	.655
	Constant	-1.900	8.016	.005	.150		
Step 4	Interaction					26.2%	.548
No Post Sec * IFM (Con)		--	.724	.696	--		
Post Sec * IFM (Poor)		-.746	.329	.566	.474		
Post Sec * IFM (Good)		.252	.228	.633	1.287		
Constant		-1.893	6.911	.009	.151		

Note. [a]Reading1 (NB) refers to the calculated (not binary) latent scores for the three key T1 reading measures (Letter-Sound Knowledge, FRY Words, Blending Words); [b]Mom Ed (PS) refers to Mother's Education binary variable, with the reference variable being no post-secondary (variable value = 0) and variable value of 1 for mothers with some post-secondary (PS) education.

Substituting the binary latent scores with the actual calculated latent scores of the T1 reading variable helped to improve the goodness of fit for the whole sample analysis at the time of follow-up for the Sex and Ethnicity regression models, but not the SES regression analysis model.

To obtain a sound SES model for the whole follow-up sample, the SES variable was also collapsed from the original three subcategories down to two. Two of the variable coefficients for the SES variable, *college* and *university* were combined to represent a single category of *Post-Secondary* (See Table 12-C).

The final analyses that were run, incorporating the adjustments made to obtain better fitting models, are reported in Table 12. By using a continuous covariate for reading, and collapsing the SES variable down to two subcategories, the model fits improved for the whole follow-up sample analysis, with the results not being markedly different from the preliminary analyses that were run.

The results for the whole follow-up sample in Table 12 (Step 1) indicated that initial T1 reading and listening comprehension skills are both significant predictors ($p < .001$) for later reading comprehension success. Early reading and listening skills explained 23% of the variance in reading comprehension 4 years later at T3. Specifically, the odds ratio of being in the typically developing reader group and not in the atypically developing reader group was increased by 2.45 times because of a student's earlier reading ability. The T1 listening comprehension variable increased the likelihood of being in the typically developing reader group by 1.22 times at T3.

While a Sex effect was not present at T2, a sex difference that favoured female participants was evident for the whole follow-up sample at T3 (see Table 12-A, Step 2). These findings suggest that being female increased the likelihood that a student would be in the typically developing reader group in later grades by 1.27 and 3.08 times.

An Ethnicity effect was not evident at time of follow-up, but an SES effect was found to significantly predict the group membership for the whole sample at T3, X^2 (1, $N = 466$) = 6.06, $p < .05$ (see Table 12-C, Step 2). These results indicate that students with mothers who had post-secondary education had an increased chance of 1.14 and 3.06 times of being in the typically developing reading group, rather than in the group struggling with reading comprehension.

The T1 at-risk sample that were part of the T3 follow-up study. With respect to the sample of students identified at risk of reading difficulties at T1 who were part of the T3 follow-up sample, preliminary analyses indicated that adjustments to the variables of interest needed to be made to produce sound fitting models. For the at-risk sub-sample analysis, a greater number of the regression models showed that the goodness of fit had been breached. Substituting the binary latent scores with the calculated continuous latent scores of the T1 reading variable created a better fitting model for the whole sample analysis, although this substitution was not effective for the at-risk subgroup analysis. Rather, maintaining the binary covariate for baseline reading similar to that in Question 1 and 2 produced sounder models for the at-risk group. The Hosmer-Lemeshow goodness of fit measure was not significant for the sex variable analysis run, so no additional adjustments were needed with the sex regression models (see Table 13-A).

For both the Ethnicity and SES factors, small cell sizes for the analysis appeared to be the contributing factor for the goodness of fit issue (refer back to Table 11). To resolve this issue, analyses were re-run with adjustments made to the *Intervention* (IFM) variable coefficients—the *Poor* implementers group was removed and *Control* versus *Good* intervention classrooms were retained for the Intervention variable conditions. Although the sample size for both the Ethnicity and SES models dropped by 10 participants (from $n = 137$ to $n = 126$), these models' goodness of fit conditions were no longer compromised.

Table 13

Analyses Results of the T1 Sample At risk of reading difficulties who were Part of the T3 Follow-up Study

Dependent Variable *(Reading 3) = If a student is a typical reader based on their overall percentile scores on the T3 standardized comprehension measure (above or below the 30th percentile). 0 = No (at risk), 1 = Yes (typically developing reader)*

Predictor Variables		B	Wald	*p*	Odds Ratio Exp (B)	Nagelkerke R^2	H-L fit
A. Sex (*n* = 137)							
Step 1	LC 1	.229	10.169	.001	1.258	11.1%	.058
	Constant	-3.233	10.095	.001	.039		
Step 2	Sex (F)	.651	3.224	.073	1.918	13.9%	.107
	Constant	-3.407	11.079	.0011	.033		
Step 3	IFM (Con)	--	1.973	.373	--	15.7%	.600
	IFM (Poor)	-.409	.288	.591	.664		
	IFM (Good)	-.533	1.951	.163	.587		
	Constant	-3.094	8.773	.003	.045		
Step 2	IFM (Con)	--	1.718	.424	--	12.6%	.534
	IFM (Poor)	-.361	.234	.629	.697		
	IFM (Good)	-.490	1.702	.192	.613		
	Constant	-2.947	8.028	.005	.053		
Step 3	Sex (F)	.684	3.475	.062	1.981	15.7%	.600
	Constant	-3.094	8.773	.003	.045		
Step 4	Interaction						
Sex * IFM (Con)		--	.316	.854	--	16.0%	.697
Sex * IFM (Poor)		.900	.309	.578	2.458		
Sex * IFM (Good)		.047	.004	.951	1.048		
Constant		-3.037	8.141	.004	.048		
B. Ethnicity (*n* = 126)							
Step 1	LC 1	.247	10.006	.002	1.280	11.7%	.062
	Constant	-3.486	9.895	.002	.031		
Step 2	Ethnic (W)	--	2.073	.355	--	13.8%	.698
	Ethnic (A)	.216	.141	.708	1.241		
	Ethnic (O)	-.587	1.659	.198	.556		
	Constant	-3.582	9.355	.002	.028		
Step 3	IFM (Good)	-.421	1.203	.273	.656	14.9%	.800
	Constant	-3.308	7.834	.005	.037		
Step 2	IFM (Good)	-.462	1.485	.223	.630	13.2%	.404
	Constant	-3.231	8.333	.004	.040		
Step 3	Ethnic (W)	--	1.797	.407	--	14.9%	.800
	Ethnic (A)	.185	.098	.754	1.203		

Dependent Variable *(Reading 3) = If a student is a typical reader based on their overall percentile scores on the T3 standardized comprehension measure (above or below the 30th percentile). 0 = No (at risk), 1 = Yes (typically developing reader)*

Predictor Variables		B	Wald	*p*	Odds Ratio Exp (B)	Nagelkerke R^2	H-L fit
	Ethnic (O)	-.554	1.481	.224	.575		
	Constant	-3.308	7.834	.005	.037		
Step 4	Interaction					15%	.473
Ethnic (W) * IFM (Con)		--	.117	.943	--		
Ethnic (A) * IFM (Good)		.023	.000	.985	1.024		
Ethnic (O) * IFM (Good)		-.314	.112	.738	.730		
Constant		-3.424	7.503	.006	.033		
C. SES (*n* = 126)							
Step 1	LC 1	.247	10.006	.002	1.280	11.7	.062
	Constant	-3.486	9.895	.002	.031		
Step 2	No Post Sec.	--	5.081	.079	--		
	College	.853	4.035	.045	2.347	16.7%	.071
	University	1.038	3.240	.072	2.824		
	Constant	-3.942	11.407	.001	.019		
Step 3	IFM (Good)	-.456	1.386	.239	.634	18.0%	.251
	Constant	-3.678	9.790	.002	.025		
Step 2	IFM (Good)	-.462	1.485	.233	.630	13.2%	.404
	Constant	-3.231	8.333	.004	.040		
Step 3	No Post Sec.	--	4.985	.083	--	18.0%	.251
	College	.858	4.040	.044	2.359		
	University	1.025	3.061	.080	2.786		
	Constant	-3.678	9.790	.002	.025		
Step 4	Interaction					19.9%	.069
No Post Sec * IFM (Con)		--	2.028	.363	--		
College * IFM (Good)		-.003	.000	.997	.997		
University * IFM (Good)		1.534	1.668	.197	4.636		
Constant		-3.913	9.614	.002	.020		

Note. Nagelkerke R^2 refers to the proportion of variance in the dependent variable (DV) that is explained by the covariates; H-L fit refers to the Hosmer-Lemeshow goodness of fit measure.

The results of Table 13 report the final analyses for the at-risk follow-up group, which incorporate the adjustments made to obtain sound regression models for the demographic variables of interest. Overall, the results of Table 12 and 13 suggest that ABRA intervention is not significant at T3 follow-up with the current sample as there are no main effects or interaction effects with IFM.

These results suggest that the sample's responsiveness to intervention does not seem to predict reading comprehension success in later years, neither for the whole follow-up sample nor the T1 sample at risk of reading difficulties.

Further, the T1 sample at risk of reading difficulties showed neither a Sex nor an Ethnicity effect at T3, although a slight trend towards a Sex effect (p = .062) was detected in Step 3. In contrast, early listening comprehension skills and SES factors were shown to be significant predictors of later reading ability for the follow-up subsample of T1 students at risk of reading difficulties. The odds ratio of being a typical reader at the time of follow-up increased by 1.09 and 1.49 times because of a student's aptitude in listening comprehension in the earlier grades. Students whose mothers were college-educated had an increased chance of 1.02 and 5.45 times of being in the typically developing reading group, rather than in the group struggling with reading comprehension.

Further exploration of the inoculation hypothesis. A closer examination of the possible long-term inoculation effects was also undertaken. Specific focus was on the notion of Matthew effects in reading—the widening achievement gap between strong and weak readers—which continues to attract a great deal of attention in education research (Pfost et al., 2014). Additional analyses on the pupil-level data were conducted to examine if the study could provide long-term support for the inoculation hypothesis and Stanovich's *Matthew effect* assertion that *the rich get richer* (Stanovich, 1986).

To further investigate the inoculation hypothesis, the proportion of the T1 sample at risk of reading difficulties from the intervention group (n = 211) that became genuinely *typical* readers at T2 immediate post-intervention was identified. For the purposes of these analyses, genuinely typical word readers at T2 were those students who scored in the 50th percentile or higher on the T2 comprehensive test of phonological processing (CTOPP) blending words measure. The top 68% of the intervention at-risk sample (n = 144) met this criterion. Next, a matching proportion reflecting the top 68% of the control at-risk sample at T2 (n = 103) was then identified.

The reading ability of this subgroup of students identified as genuinely typical readers at T2, regardless of demographic and cognitive factors, was the variable of interest. Independent-samples t-test comparisons were conducted to compare the reading outcomes of the control and intervention conditions. At immediate post-intervention, a significant difference occurred in the blending word scores for the control (M = 58.66, SD = 14.04) and intervention (M = 64.39, SD = 14.51) conditions; $t(245) = 3.10$, $p = .002$, with the intervention subgroup receiving higher scores. However at time of follow-up, T3, a significant difference did not exist in the GRADE Total Test percentile scores for the control (M = 34.24, SD = 22.37) and intervention (M = 32.44, SD = 23.08) conditions; $t(94) = .380$, $p = .705$. The results suggest that treatment effects are evident immediately after intervention, although these significant gains in reading outcomes for the intervention group are not maintained over the long-term.

Summary

This chapter examined the follow-up pupil-level data. The added contribution of three key demographic variables (i.e., socioeconomic status [SES], ethnicity, and sex) in predicting students' risk of reading difficulties after their involvement in the ABRA intervention study were examined. Significant independent main effects for the ABRA intervention—SES and Ethnicity—were found at immediate post-testing. Neither a main effect of Sex nor an Interaction between the levels of implementation of ABRA (IFM) and Sex, SES and Ethnicity were evident at T2.

At T3, an SES effect was again evident for both the whole sample and the T1 sample at risk of reading difficulties, which suggests that children whose mothers had some post-secondary education had a greater likelihood of being typical readers at the T3. For the whole sample, a Sex effect was evident, which suggests that female students were more likely to be typical developing readers in later grades; however, the Sex effect was not statistically significant, although a positive trend was present (p = .062) when examining the at-risk sample T3 results. In addition, an Ethnicity effect was not evident at the time of follow-up for either the whole sample or the T1 sample at risk of reading difficulties. Also, while immediate group differences existed between the control and intervention student groups at T2, a long-term treatment effect of ABRA on the reading skills of the intervention student sample group at T3 was not evident.

Chapter 8: Teacher Data Analysis and Results

Overview of Chapter

The second focus of this dissertation examined teachers' long-term responses to being involved in an intervention study. Follow-up data were collected from a sample of 22 teachers, 19 of whom provided multiple sources of data. This dissertation is a quantitative follow-up study with a nested qualitative component at the teacher level of data analysis. In this design, a dominant phase using measureable quantitative data from questerview responses was followed by a less dominant phase that examined open-ended qualitative questerview responses. Quantitatively, this second focus examined the relationship between teachers' original level of implementation and comfort with technology and their use of the ABRA program at the time of follow-up. The qualitative phase explored additional teacher-level and school-level factors that may have influenced teachers' long-term implementation of a new innovation. The objective of the addition of a qualitative descriptive piece was to obtain a richer picture of the complexities of the factors influencing teachers' responses to being part of an intervention study.

This chapter discusses the steps taken in the coding and analysis of the teacher data gathered through the questerview method process. The chapter concludes with a teacher results section that explores the ABRA teachers' responses to being part of the intervention study, and factors influencing their long-term use of the ABRA program.

Data Analysis

The aim of the teacher data analysis was to evaluate the teachers' responses to being part of an ABRA intervention study. The teacher-level data was subject to both quantitative and qualitative analysis.

Quantitative analysis. Quantitatively, data were examined using SPSS version 20 software that uses standard parametric tests. The chi-square test of independence, also called the chi-square test of association, was the primary test of analysis used to discover if any relationships were evident between the categorical variables of interest at the teacher-level of analysis (Field, 2009).

Qualitative analysis. Thematic analysis (TA) was the data reduction technique used with the qualitative data. TA involves the process of searching across a data set to find repeated patterns of meaning and identifying themes of significance in describing a phenomenon (Braun & Clarke, 2006; Miles & Huberman, 1994).

Deductive thematic analysis. Rather than an inductive TA approach in which the data coding and analysis is driven by the data (a bottom-up approach), this study employed a predominately **deductive TA** approach in which the coding and analysis was more of a top-down analytic process that extracted data to fit ***a priori*** theory-based concepts and themes (Braun & Clarke, 2012). When examining the teacher-level data, the predefined research questions of the present study provided the structure for coding and analysis (Braun & Clarke, 2012). The research questions guided the development of the coding strategies and the four theoretical propositions of the *Teacher Change Domains (external domain, domain of practice, domain of beliefs and domain of consequences)* were used to extract data that were mapped out and interpreted by the researcher. In comparison to the inductive TA approach, which is used to provide rich descriptions of data, the deductive TA approach provides more of a detailed analysis of the coded data (Yin, 2016).

Data Coding

The aim of the coding was to draw out data extracts from the teacher questerview data that could be relevant for answering the research questions posed by this dissertation. When going through the data, codes were used to give meaning to chunks of data and assist the researcher in summarizing, describing, and interpreting the transcribed data. The codes were organized using an *Excel* spreadsheet program. They were a mix of both descriptive and interpretative data (Braun & Clarke, 2012). Primarily, the analyses of the data used a deductive TA approach, although these analyses also tended to find subthemes that were emergent rather than *a priori*, and that had not been anticipated. Thus, the coding was predominately descriptive in nature for the initial stages of the coding, and then some interpretative coding was done during the final re-reads and the sorting of the coded data extracts into patterns of meanings (Braun & Clarke, 2006).

The researcher and her supervisor decided that a minimum of three repetitions of a particular topic could constitute a theme (Creswell, 2014). Thus, a potential influencing factor was coded and categorized into a theme or subtheme if more than three teachers made a reference to it. Any *neutral* comments made by the teachers were not categorized or analyzed. A *neutral comment* lacked an elaboration to support the response. For example, if a teacher—when asked about the training and support she received—merely stated that "Yes she had received the ABRA training" but did not elaborate on how she found the training to be, or what she gained from it, or what she found it lacked, then this teacher's response was not categorized.

This analysis of the teachers' open-ended qualitative text provided an opportunity to obtain a richer picture of the complex nature of influences on teachers' RtI from the educators themselves. Responses were coded quantitatively (percentages) and qualitatively.

First phase. The process of coding and analysis of the transcribed data began when the researcher initially became immersed in the data. To begin, the transcripts were all thoroughly read a few times without a conscious attempt at coding. However, during these initial read throughs, general questions and thoughts about what was being read were noted in the margins of the transcription hard copies.

Second phase. During the next few readings of the transcripts, a conscious effort was made to identify and begin the coding of data that could potentially be relevant to the research questions of interest. Emerging codes were noted in the margins of the transcribed text, and the associated portions of the text were highlighted with a running record of emerging codes, documented as they occurred (Braun & Clarke, 2012).

During this step an Excel spreadsheet document was created to begin categorizing the coded teacher data. The qualitative data was organized in Excel spreadsheets, since this software was most cost affordable and enabled the researcher to work with the data in a flexible and meaningful way that suited the analysis.

Using a *template approach*, six *a priori* conceptual categories representing the research questions of interest were initially used to organize the extracted data in the first Excel document created. A *template approach* is when "key codes are determined on an *a priori* basis…these codes serve as a template or 'bins', remaining flexible as the data analysis process proceeds" (Crabtree & Miller, 1992, as cited in Bloomberg & Volpe, 2012, p. 138).

These initial six *a priori* areas of focus included: (1) teacher comments identifying positive influencing factors; (2) teacher comments identifying negative influencing factors; (3) teachers' responses to being part of the intervention study, specifically, comments reflecting a change in beliefs, attitudes, and practice regarding technology use; (4) teacher comments towards the training and support they received during the study; (5) teacher comments reflecting their current use of ABRA; and (6) teacher comments reflecting their change in practice process during their teacher change model selection.

Third phase. Next, a second Excel workbook including the four domains of teacher change was created to further organize all the pulled data quotes and comments that had been collected and categorized initially in the first Excel document. This stage of coding involved taking the extracted data quotes from the six *a priori* areas of interest and further comparing and sorting them down to one of the four conceptual domains of professional growth (e.g., external, personal, practice, and consequence) delineated by Clarke and Hollingsworth (2002). A total of 276 separate data segments were identified for coding and analysis purposes.

Fourth phase. After each of the 276 data segments was categorized under one of the four teacher change domains, the process of grouping and categorizing the coded segments into themes and subthemes began. Based on the central idea within each pulled data extract, the teacher data extracts were coded and clustered into seven key conceptual themes and then further down to 16 subthemes under one of the four domains of teacher change (detailed descriptions and analysis of the themes and subthemes are presented later in this chapter in the *Results* section and are summarized in Table 16).

Prior to the final analysis, the themes were revisited several times to delineate and refine them (Yin, 2016). The end goal of this extensive process of data transformation was not only to identify patterns of influencing factors that impacted teachers' responses to being part of the ABRA intervention study, but also in turn to produce an *empirically-grounded* typology of teacher change influencing factors that could be compared and contrasted against Clarke and Hollingsworth's (2002) four domains of professional growth.

Reliability of Coding. A process was undertaken in order to validate the dependability of the conceptual categories developed during the coding and data analysis process. A reader, with a background in qualitative coding experience and familiar with the present study, independently reviewed all of the coded data segments to determine if the codes were logically and appropriately assigned. Prior to coding, this reader was provided with an introduction and background to the research and a brief description and example of each of the coded themes and sub-themes. The reader's feedback was used in a final review of the assigned coding and of the conceptual theme and sub-theme categories. The reader examined the definitions of all the themes, and the coded data points identified to support them. All 7 major conceptual categories and 16 sub-theme categories were agreed on.

After reviewing the 276 categorized comments, the reader suggested that 26 data coded segments could either be broken down further or perhaps support another sub-category. The researcher revisited these suggestions and determined that some coded segments did warrant movement. The reliability of coding measure resulted in a 91% agreement between the coding of the researcher and that of the reader.

Member checking. To assist further with the validation of the coded data, a process of member checking of the agreed upon data coded segments was carried out. The value of the member checking at this stage was to find out whether the coded data was congruent with teachers' interpretations of their experiences (Carlson, 2010). After the coding was complete, a random selection of over half of the participants (n = 16) were contacted via email. The teachers were provided with an update of the data analysis that had been carried out and how the data was coded, categorized, and sorted into themes and subthemes. They also were emailed a copy of the Excel workbook containing the final categorized data. Each of these Excel documents was individualized, so the specific comments a teacher had made were highlighted for her to review. Specifically, they were asked to review the coding of their data segments under the four domains of teacher change. Twelve of the teachers replied after reviewing their coded comments (n = 2 from Ontario, n = 4 from Alberta, and n = 6 from Quebec). The feedback received was positive with respect to the interpretation, coding, and categorization of the reviewed coded comments, and no additional changes or further suggestions were requested.

Results for the Teacher-Level Research Questions

Demographics of the follow-up teacher sample. Twenty-two of the 33 accessible teacher participants (67%) agreed to participate in some capacity with this follow-up study (n = 3 questionnaire only, n = 19 questionnaire and interview). The participating teachers were all female, and the majority identified their ethnicity as white (n = 21 White, n = 1 White and South Asian). At the time of follow-up, the teachers' ages ranged from 28 to 61 years (M = 44.50 years; SD = 9.14), and they all had over 5 years of teaching experience (M = 16.68 years; SD = 7.69), ranging from 6 to 32 years of experience.

Results for teacher-level question 4. *At the time of follow-up, had teachers integrated ABRA (technology) into their practice?* Up to 4 years after their involvement in the ABRA 2007–2009 study, over 70% of the teacher respondents (n = 16 out of 22 teachers) reported that the ABRA program continued to be part of their literacy practice in some capacity. Fifty-five percent (n = 12) of the respondents reported that they had used the ABRA tool in the academic year in which the follow-up study was done. Another 18% (n = 4 teachers) reported that although they were not using ABRA with their current group of students, the ABRA tool was still part of their literacy repertoire. Twenty-seven percent (n = 6 teachers) reported that after the completion of the study, they discontinued using the program as part of their literacy program.

The four teachers that were not using ABRA at the time of the follow-up, although ABRA was still part of their literacy repertoire, all reported systematic reasons for their current disuse. ***Teacher A*** was teaching a group of pre-K students for only half a day and found them too young/challenging to take up to the computer room during the time she had with them. She reported that next year, she was scheduled to teach full-day kindergarten classes, and she planned to use ABRA again when she took her students to the computers. ***Teacher T*** was teaching grade four students and found that the tool was not suitable for the literacy needs of her older students. She stated that if she ever taught the younger grades again, she would revisit the program. ***Teacher C*** had just returned from maternity leave during the month that her interview occurred, and she was continuing with the literacy program set in place by the previous teacher, which did not integrate the students' computer time with language arts practice. She stated that she would use ABRA at the beginning of the next academic year when she would be implementing her own literacy program. ***Teacher E*** reported that her kindergarten class did not have access to the computer room this year due to a "population explosion" in their school catchment area, which had resulted in only intermediate grades having scheduled computer time. ***Teacher E*** asserted that the ABRA program was actively used when her students had regularly scheduled computer-time.

Since these four teachers reported structural reasons for why they were not using ABRA that particular school year (but would otherwise be using it), it was decided to group them with the teachers still using ABRA (n = 12). In the subsequent analyses to follow, the group of 16 teachers who still had ABRA as part of their early literacy repertoire would be the comparison group for the group of 6 teachers who were no longer using ABRA in any capacity. Table 14 summarizes the comparisons of the key characteristics of the teachers who still had ABRA as part of their teaching repertoire versus those who were no longer using ABRA.

Table 14

Teacher Participant Characteristics -- Comparing 'ABRA is Part of their Repertoire' Group versus 'No Longer Using ABRA' Group

Characteristics	Overall	Part of Repertoire	No longer using ABRA
Teachers (*n*)	22	16	6
Age			
M	44.5	45.1	43.0
SD	9.1	10.1	6.5
Range	28-61	28-61	31-50
Years of teaching experience			
M	16.7	17.1	15.5
SD	7.7	8.4	5.7
Range	6-32	6-32	8-22
Female	100%	100%	100%
Ethnicity			
White	95.5%	94%	100%
Non-White	4.5%	6%	0%
ABRA Condition[a]			
Control Only	9%	6%	17%
Experimental Only	27%	31%	17%
Both	64%	63%	67%
ABRA IFM[b]			
M (0-4)	2.05	2.31	1.33
SD	1.05	.95	1.03

*Note***.** [a]ABRA Condition refers to the intervention group(s) the teachers were part of during the original study. [b]ABRA IFM refers to the Implementation Fidelity Measure score teachers received during the original study (scores ranged from 0 to 4, with 0 = Control Group and 4 = Differentiated Adaptation).

How often and in what capacity are teachers currently using ABRA? The 16 teachers who were still using ABRA as a part of their teaching repertoire varied in their responses to the question of how often their students worked with the ABRA program. The responses ranged from 12.5% reporting that they currently do not use ABRA at all, 19% reporting they use ABRA infrequently in an odd lesson once in a while, and another 12.5% stating that their students use ABRA occasionally for about 30 minutes per week. At the time of the follow-up, the highest percentage of the respondents (44%) reported that they were continuing to use ABRA regularly (30–60 minutes/week) in their LA lessons, and another 12.5% of the teachers claimed that ABRA was an integral part of their early literacy program, and they used it frequently with their students, at least 1 or 2 hours per week.

Fit between the ABRA program and teachers' existing LA curriculum. Teachers used a 5-point Likert-type scale (0–*not a good fit*, 1–*a poor fit*, 2–*a moderately good fit*, 3–*a good fit*, and 4–*a very good fit*) to rate the overall fit between the ABRA program and their existing language arts curriculum. Overall, the teachers (n = 22) reported that the ABRA resource itself was a good fit (M = 3.00) for their literacy program. Most (36%) reported that there was a *good fit* between ABRA and their literacy program, while 32% reported a *moderately good fit*, and another 32% reported that there was a *very good fit* between the ABRA program and the literacy skills their students were learning.

Results for teacher-level question 4(a). *Does a teacher's level of implementation of ABRA ELA lessons during the Tier 1 intervention influence her decision to continue using the ABRA tool after the conclusion of the study?* According to the four-component theory models of teacher change, the prediction is that those who implement ABRA well will carry on using it in the future. To address this question, the following two nominal variables were compared: teachers' treatment integrity level of implementation (IFM) at time of the original study (poor implementers/good implementers) and teachers who continued using ABRA long-term (Yes/No).

A *poor implementer* was identified as a teacher who received an overall IFM score of one, *Entry Level*. That is, during the study, a *poor* (low) implementer showed little to no evidence of teacher planning of ABRA ELA lessons nor teacher instructional guidance. ABRA exposure was primarily through unstructured lessons where students chose their own activities (i.e., during free-time). The majority of the low implementers in the original study were kindergarten teachers (n = 6 kindergarten teachers and 1 = first grade teacher).

Whereas a *good implementer* was a teacher who scored an overall IFM score of 2 (*Adoption* Level), 3 (*Adaptation* Level), or 4 (*Differentiated Adaptation* Level). A teacher identified as a *good* implementer would have shown at least some basic evidence of delivering structured ABRA lessons, and appeared comfortable navigating and guiding her students through the ABRA activities (for additional IFM criteria, see Appendix K, ABRA Implementation Rubric).

A teacher was considered to be still using ABRA if they reported that ABRA continued to be part of their teaching repertoire at the time of follow-up (Yes/No). Teachers who only taught a control (non-ABRA) classroom during the intervention study were excluded from the analysis (n = 2), leaving the data from 20 follow-up teachers to be compared (see Table 15).

The chi-square test of independence was performed to examine the relationship between teachers-level of implementation (IFM) of ABRA ELA lessons during the original study, and teachers' current use of the ABRA resource. Since this was a 2 x 2 variable analysis, the chi-square distribution was adjusted with Fisher's Exact test (χ^2 (1, N = 20) = 10.59, p = .009). The results suggested that the relation between these variables was significant. For chi-square analyses, the effect sizes are Cramer's V or phi (ϕ), with the following guidelines suggested by Cohen (1988) used to determine the magnitude of the effect: 0.1 is a small effect; 0.3 is a medium effect; and 0.5 is a large effect. Using effect size in the standard manner for categorical analyses, the results also showed that the effect size of the above contrast (ϕ = .73) was also large, potentially suggesting practical significance of the findings.

The teachers who were identified as poorer implementers of the ABRA program during the intervention phase of the original study were less likely to continue using ABRA long-term. Of the three follow-up teachers identified as poor implementers, two taught kindergarten while the other taught grade one during the implementation phase of the original study. This Grade 1 teacher pointed out that most of her students had attended a full-time French kindergarten program the prior year, and thus she felt more like a kindergarten teacher than a Grade 1 teacher when teaching the English LA program to her group.

Table 15

Comparing Teachers' Level of Implementation During the Intervention Study to their ABRA Use Four Years Later

Level of implementation during intervention study	*ABRA still part of repertoire* Yes (n = 15)	No (n = 5)
Poor (IFM score = 1)	0 (0%)	3 (60%)
Good (IFM score >1)	15 (100%)	2 (40%)

Note. Poor implementers were teachers who had an overall IFM score equal to 1, whereas a good implementer was a teacher who scored an overall IFM score of 2 (Adoption Level), 3 (Adaptation Level), or 4 (Differentiated Adaptation Level).

Results for teacher-level question 4(b). *In relation to treatment integrity levels, with a particular focus on the adaption of ABRA into their lessons during the intervention phase, is there a pattern in terms of teachers still using ABRA up to 4 years later?* The earlier examination found that the teachers who were poor implementers of the ABRA tool during the intervention phase of the study had a greater likelihood of not using the resource at the time of follow-up. A second chi-square adjusted with Fisher's Exact test was performed to determine if a pattern of long-term use would still be evident when focusing on teachers who implemented ABRA above just adopting it into their practice (scoring an IFM rating of 3–*Adaptation* or 4–*Differentiated Adaptation*) in comparison to their peers who incorporated ABRA into their practice at a basic level (scoring an IFM rating of 1–*Poor* or 2–*Adoption*). This statistic was not significant (χ^2 (1, N = 20) = .317, p = .517). Effect size was again calculated. The results showed that the findings (ϕ = .13) was small suggesting little practical significance of the contrast.

This analysis was rerun with Adoption only (n = 11) teachers versus Adaptation implementers (n = 6). This analysis did not include the Control teachers (IFM = 0) or the Poor implementers (IFM = 1). Note, with the removal of Poor IFM, only 2 teachers remained in the No Longer Using ABRA Group (one for Adoption one for Adaptation). The chi-square adjusted with Fisher's Exact test (χ^2 (1, N = 17) = .215, p = .60) also was not significant. Effect size was again calculated. The results showed that the effect size (ϕ = .11) was small, suggesting little practical significance of this contrast.

Results for question 5: Emerging patterns of influencing factors. *Are there common identifiable factors that have influenced teachers' decisions to integrate ABRA into their literacy practice?* The purpose of this question was to assess the educators' perspectives on what factors influenced their current use (or disuse) of the ABRA program. By employing a deductive thematic analysis approach, the aim was to draw out commonalities and patterns of meanings across the teachers' *questerview* question responses. Questerview question responses refer to the open-ended response data from both the teacher questionnaires and the teacher interviews. When sorting data extracts into conceptual categories, the extracts that were included for analysis were those that could be identified as either a positive (supporting) or a negative (hindering) influencing factor.

Emerging Patterns of Influencing Factors

From the teacher data, and through the coding process, patterned responses of positive and negative influencing factors were grouped into seven key conceptual categories, which fell under one of the four domains of interest (see Table 16).

Table 16

Summary of the Teacher Participants who provided the Data Comments that Shaped the Key Themes and Sub-Themes Analyzed in the Teacher Results Section

	EXTERNAL DOMAIN				**PERSONAL DOMAIN**			**DOMAIN of CONSEQUE.**		**DOMAIN of PRACTICE**						
	Nature of Training *n*=17		Nature of Support *n*=18		Pedagogical Beliefs *n*=22		Self-Eff[a] *n*=13	Flagging St. Interest[b] *n*=20		Environment Elements *n*=22				Environment Contributors *n*=16		
	Familiarize	Timing	Resources	In-Class	Alignment with resource	Value & Use of Technology	Improved with involvement	Student Friendliness	Student Engagement	Structural	Adaptability	Tech. Access	Lack of Time	Administration	Colleague	Technical
#	15	6	11	12	15	22	13	11	19	6	15	17	10	8	7	6
%	68	27	50	55	68	100	59	50	86	27	68	77	45	36	32	27
Teachers Currently Using ABRA																
B	**+**		**+**	**+**	**+**	**+**	**+**		**+**		**+**			**+**		
F	**+**	**+**			**+**	**+**	**+**	**+**	**+**		**+**	**B**				
H	**+**		**+**	**+**	**+**	**B**	**+**	**+**	**+**		**+**			**+**		
I	**-**		**+**		**+**	**+**	**+**		**+**	**-**		**B**	**-**	**-**		
L	**+**			**-**	**+**	**+**			**B**		**+**	**-**	**-**			
M	**+**		**+**	**+**	**+**	**+**	**+**		**+**		**+**	**-**				**+**
N	**B**		**+**	**+**	**+**	**B**	**+**		**+**			**-**			**+**	**B**
Q	**-**		**+**			**+**		**-**	**+**			**+**		**+**		
R			**+**		**+**	**+**		**+**	**+**		**+**				**+**	
S	**+**		**+**		**+**	**+**		**+**	**+**		**+**	**-**				**+**
U*	**+**		**+**		**+**	**+**			**+**			**-**	**-**		**+**	
V*			**+**			**B**			**+**		**+**					
Teachers Not Currently Using ABRA, But ABRA is Part of Teaching Repertoire																
A		**+**		**B**		**+**	**+**	**-**	**+**	**-**		**-**			**+**	**-**
C		**+**		**+**	**+**	**+**	**+**		**+**	**-**	**+**	**B**				
E	**+**	**+**			**+**	**+**			**+**	**-**	**+**	**B**	**-**	**-**	**-**	
T*	**+**		**+**	**+**		**+**		**+**	**+**	**-**	**+**					
Teachers No Longer Using ABRA																
D	**+**			**+**		**+**						**-**	**-**	**+**		**+**
G	**-**	**-**		**-**	**+**	**B**	**-**	**-**	**-**		**-**	**B**	**-**		**-**	
J		**+**			**B**	**B**	**-**	**-**			**-**	**B**	**-**	**-**		**-**
K	**B**			**+**		**B**	**+**			**-**		**-**	**-**		**-**	
O					**-**	**+**	**-**	**-**	**-**		**-**	**-**	**-**	**+**		
P				**+**		**B**	**-**	**-**	**-**		**-**	**-**	**-**			

Note. [a]Self-Eff. refers to Technology Self-Efficacy. [b]Flagging St. Interest refers to Flagging Student Interest. A **(+)** indicates positive comments made towards factors within the subtheme. A **(-)** indicates negative comments made. The letter **(B)** indicates when both negative and positive comments were made. An asterisk (*) by a teacher's letter denotes that the responses examined came from teacher questionnaires. The number and percentage of participating teachers making reference to a specific theme/subtheme is noted below each factor of influence category.

Under the **External Domain**, the two major conceptual themes were: (1) *Nature of the Training* and (2) *Nature of Support*. Under the **Personal Domain**, the two central themes were: (3) *Pedagogical Beliefs* and (4) *Self-Efficacy.* (5) *Flagging Student Interest* categorized the pattern of data extracts that emerged under the **Domain of Consequences,** and under the **Domain of Practice,** the key themes that emerged were: (6) *Elements of the Environment of Practice* and (7) *Contributors to the Environment of Practice*.

The common identifying patterns of the key concepts and factors that arose within these seven conceptual themes are examined next. Table 16 provides a summary of the 7 themes and 16 sub-themes, and specifies the proportion of teacher participants who provided the descriptive comments that shaped these themes. The table highlights the teacher comments made toward the various themes that were favourable (with a positive sign) and unfavourable (with a negative sign). If both negative and positive comments towards a particular theme were made, a letter ***B*** would be noted. The following section examines the collection of teacher comments and provides a detailed analysis of the contents of Table 16. In this section, the four domains of teacher change provide the framework for the review of the themes and subthemes of the teacher qualitative results.

External Domain

The data patterns of the influencing factors in the external domain were drawn primarily from the teacher responses to the following two questerview questions: (1) *During your involvement in the ABRA study, can you describe what key factors over the course of the training support phases influenced your use of ABRA in your classroom?* and (2) *Can you describe what key* *factors over the course of the training and support phases influenced your current professional use of the ABRA software?* Two themes captured the discussion of issues related to the external domain: *Nature of the Training* and *Nature of the Support*.

Nature of the training. A majority of the participants (77%) reported on the elements of the training they received during their involvement in the ABRA intervention study (n = 17 teachers made references towards factors within this theme; n = 14 made positive influence references, n = 5 made negative influence references). Most of this subgroup of respondents (82%) discussed features that they found favourable.

For example, the training overview of the literacy-content of the ABRA resource was noted as being relevant to meeting the district curriculum demands and fitting with their existing literacy practice objectives. Other key positive aspects of the training included the interactive nature of the sessions that allowed for hands on time with the ABRA resource. For some teachers, this time included opportunities to collaborate with their peers and develop ABRA lesson plans. Two central subthemes emerged within this theme: *Familiarization with the ABRA Tool* and *Timing and Duration of the Training*.

Familiarization with the ABRA tool. Fifteen participants (68%) made references to the ways in which the training enabled them to become more familiar with the ABRA program (n = 12 of these references were favourable, while n = 5 of these references were unfavourable). Favourably, participants reported that the interactive training enabled them to become familiar and comfortable with ABRA prior to using it in their practice with their students. For example, ***Teacher H*** (a 48 year old, Grade 1 teacher with 24 years of teaching experience) stated, “Initial training was excellent, [it] built my confidence on computers. [And they] taught me how to bookmark [the] ABRA site and how to use it in my program.” ***Teacher U*** (a 45 year old, Grade 1 teacher with 9 years of teaching experience) echoed how the training enabled her to become familiar with the program: “The training really helped me feel more comfortable with the characters, games, and stories. It also showed me how I could incorporate ABRA into my language program.” Another teacher noted, “The training provided time for me to explore and learn, so I could best match up students to activities to try to meet their needs*”* (***Teacher S,*** a 51 year old, Grade 1 teacher with 28 years of teaching experience). ***Teacher F*** (a 57 year old, Grade 1 teacher with 20 years of teaching experience) added:

> It [the training] was interactive and after the training we, at least I, felt ready and comfortable enough to use it with my class. I remember it was a good opportunity to really get to it [the ABRA program] and to develop some beginning lessons.

As mentioned previously, not everyone found the nature of the training to be sufficient. Some of the reporting subsample (29%) highlighted aspects of the training that they found lacking, and noted that they would have appreciated even more time to familiarize themselves with the program before the study began. For example, ***Teacher Q*** (a 43 year old, Kindergarten teacher with 20 years of teaching experience) shared the following:

> Remembering back to that time, I wish that we had a bit more time to kind of work through the actual program and see what [the] kids would be seeing. That was my hope at the time, because after the training, it then took a while of me just playing to get to know the program better.

Teacher I (a 33 year old, Grade 1 teacher with 9 years of teaching experience) highlighted how she would have benefited from more guided instruction and hands-on-time exploring ABRA prior to using it with her students:

> I don’t remember much guided instruction. Having a bit more of that would have been good. It was more [pause]… I felt like the responsibility was more on our part to kind of play with it, and then to ask for questions. So sometimes it was difficult because I was trying to learn it, but sometimes it was just difficult finding time.

Teacher G (a 50 year old, Kindergarten teacher with 10 years of teaching experience) echoed similar sentiments, "I remember it as a very brief training, and it was more our own hands-on training as far as going in and exploring more than anything else… I could have used more time to practice and play with it." The range of satisfaction with regards to the training the teachers' received is evident in these comments.

Timing and duration. The timing and duration of the training was noted favourably by some of the teachers (23%). Particularly, some teachers acknowledged that they appreciated the training being carried out during regular workday hours and that the training was more than an hour workshop. When looking at the topic of duration, the length of the teachers' overall involvement in the study was noted by some (14%) as having an impact on their use and comfort level with the program. Not specifically referring to the duration of the ABRA training session, some teachers who were part of the study for more than 1 year made positive comments regarding the ease of implementation during their second year of involvement.

For example, ***Teacher F*** (a 57 year old, Grade 1 teacher with 20 years of teaching experience) commented on how she believed her increased comfort level with the tool improved her delivery of it, "Well, I think the group after the first year would have benefited more than the group I had first year (laughter) because I was more comfortable with ABRA and more prepared the second year." While ***Teacher E*** (a 54 year old, Kindergarten teacher with 32 years of teaching experience) commented:

> I think my success with the program the year before encouraged me to try it out and to use it with a different groups of kids the next year… and continued onward from there to now, because even though I'm not using it this year, I did use it last.

Teacher C (a 34 year old, Grades 1 & 2 teacher with 8 years of teaching experience), who was part of a pilot study and the first year of the ABRA 2007–2009 intervention study, added that:

I felt much more prepared, I was like "Oh look at that…" So it was helpful in finding out more about how I could use it because the first year it was more like "Hmm, what do I do" kind of thing, and I was also still learning how to use it. And the second year seeing that I knew how to use the tool, and I knew quite a bit about the different areas that were in ABRACADABRA, I was more comfortable.

Nature of support. Eighty-two percent of the participants made references to influencing factors related to the nature of the support they received during the intervention implementation phase of the study (n = 18 teachers made references towards factors within this theme; n = 16 made positive influence references, n = 3 made negative influence references). Two distinct categories of support were identified: *Supporting Resources* and *In-Class Support*.

Supporting resources. Fifty percent of the teachers (n = 11) discussed how specific ABRA resources, which they were provided with during their training sessions, assisted them with the implementation of ABRA. From this sub-group of respondents, the following supporting resources were identified as supporting them during the implementation phase: the online support videos (9%), the teaching manual (36%), and the ABRA icons (55%).

The online support videos were noted as a useful resource for illustrating different ways that ABRA could be implemented, from whole class instruction with a SMART Board to using ABRA as part of a literacy center. The ABRA teacher's manual, both the hard copy and the online versions, was noted as a useful resource for providing lesson ideas and practical suggestions about how to address specific literacy skills using ABRA activities. As illustrated by ***Teacher V***'s (a 28 year old, Grade 1 teacher with 6 years of teaching experience) statement, "The teacher's manual really helped me plan lessons. I could easily pick a different literacy activity from each section."

Fifty-five percent of those commenting on the ABRA resources said they found the ABRA icon cards to be of value. The icon cards were an image resource of various ABRA characters and activities individually being represented on cardboard cards (3 x 5 inches in size). Teachers remarked that they used these cards as navigating visual aids for their students. For example, ***Teacher Q*** (a 43 year old, Kindergarten teacher with 20 years of teaching experience) noted how the icon cards were valuable for directing her students through the program to the specific activities she wanted them to explore:

> We had a computer lab at the time and so we would need to walk them through the various steps [Teacher Q visually demonstrates by clicking invisible buttons in the air]… I would show them that you have to go to the desk and put this here, but also to remember to choose this particular story if you want this activity… you know all of those little cues… so getting those little cards certainly helped to show the children what they needed to do, instead of me having to physically go around to each and every one. So that was a good resource.

Teacher S (a 51 year old, Grade 1 teacher with 28 years of teaching experience) discussed how she used the icons to create a system to track the students' progress on ABRA:

> The icons were provided after my training and I still use these. Because especially with pre-literate children, K-1, you need those icons so that they can navigate quickly to where you want them to go. I also developed a tool using the online icons and I made a tracking sheet, so I could actually track which activities the children had completed. And so I could see if we needed to do that again or could go on to another one.

In-class support. Twelve teachers (55%) made references to the in-class RA technical support provided during the implementation phase of the study. A majority of this subgroup (83%) made favourable references, with many of the educators noting how the in-class support assisted in building their comfort level with the ABRA program. However, 25% of the references highlighted times and situations when teachers stated that they could have benefited from additional support. The following teacher comments illustrate the reported range of satisfaction with the in-class support provided during the intervention phase of the 2007–2009 ABRA study. For example, ***Teacher C*** (a 34 year old, Grades 1 & 2 teacher with 8 years of teaching experience) commented on the value of having in-class support during the initial stages of implementation:

> I was happy to have people come in and watch me do it [use ABRA] because I wanted to know if it made any sense, you know [pause] was I doing it correctly, were the kids getting it, so I liked that… and because of all of the different technical issues also, it was nice to have somebody to come in and help. Help with the computers or plug something in, or look for something. It was nice just to know somebody was there from the program to kind of release the stress or the tension of "Oh my God, I'm doing this all by myself! I can't do this"… you know it was just good to have somebody there if I needed something I could just say, "Do you mind, could you help out"… so that, I thought that was great.

Teacher M (a 53 year old, Grade 1 teacher with 7 years of teaching experience) added that the in-class support allowed for timely trouble shooting of issues as they arose:

> The support was very helpful when the students were stuck. It provided good training for the students so they could work independently. The training and support also allowed me to feel comfortable with using ABRA as I was able to get immediate responses to questions.

Teacher H (a 48 year old, Grade 1 teacher with 24 years of teaching experience) also discussed how the technical support provided positively assisted her with the implementation of ABRA during the study:

> It [using ABRA] was made easy for me to implement by the trainers on the program. Most important was the support. They set up the links and showed me how to put them on a menu bar. Continued support for the follow-up weeks was good too… so I could ask questions and really get it going. I can't remember exactly how long or for how many lessons, but I was very impressed with how much they were helping. I remember getting a ton of support and really that is why I did it [participate in the study].

Teacher B (a 44 year old, Grade 1 teacher with 24 years of teaching experience) noted how she continued to receive support when needed, even 4 years later at the time of the follow-up:

> I can honestly say that anytime I needed any kind of support, the team was there. Whether I needed someone to pop ideas off of or was having trouble with a game, even now. Anytime I called they answered me right away, and like I said, they still answer me right away anytime I call. Whether it is a tech problem or an understanding problem. Basically anything to do with the program, everyone was always there to help me.

Some teachers also acknowledged that prior to their involvement in the ABRA study they were not comfortable with technology, so having the in-class support available was something they valued. For example, ***Teacher P*** (a 45 year old, Kindergarten teacher with 19 years of teaching experience) noted, "As I am not a 'techie' I would have been less than enthusiastic about using a new program without support." ***Teacher N*** (a 55 year old, Grade 1 teacher, teaching 20 years) recalled:

> My comfort with using computers is much better than it used to be because at first I wasn't comfortable using it [ABRA and computers in class] and that is why I was so grateful for the support of the people coming in.

However, not all of the comments regarding the in-class support were positive. As mentioned previously, some of the teachers who discussed the influence of this factor stated ways in which the in-class support did not meet their particular needs during their time in the study. For example, ***Teacher L*** (a 34 year old, Kindergarten teacher with 10 years of teaching experience) recalled the following, "There was very little in-class support, I think they came once? I don't remember much though, but I do remember wanting a bit more support at times." ***Teacher G*** (a 50 year old, Kindergarten teacher with 10 years of teaching experience) discussed how she would have perhaps benefited from additional in-class support:

> Without support, I felt that I was not using ABRA effectively. At times, I had difficulty navigating through the site, which impacted my students' use of the site and their engagement of the material presented. Yeah [pause] I think if I had some of the support you mentioned that would have influenced my use definitely… I just sort of felt that at times I was sort of floundering on my own, and so that definitely influenced my use of it.

Another teacher commented on how although she felt supported by the team, she was often overwhelmed by their presence during classroom visits. Specifically, ***Teacher A*** (a 61 year old, Kindergarten teacher with 17 years of teaching experience) pointed out her frustrations of being observed in the computer room. She stated that the observations added undue pressure on her in an environment with which she was not comfortable (i.e., in a poorly functioning computer room rather than in her own classroom). ***Teacher A*** added that during the initial stages of implementation, she could have benefited from more support, since she faced a number of technical issues due to outdated computers and poor Internet service at her school:

> Now, I did feel supported but I also felt very frustrated, and I also felt overwhelmed by the people coming in to observe. They were all very nice people, but they were coming to observe in a lab that had a lot of technical problems. In a classroom I am super confidant and totally in control, but in a lab when it is the first time I'm presenting a lesson, to have people sitting and noting things… now I thought they were basically noting how the children were using the software, but I basically always walked away with the impression that they were noting what I was saying and doing. I found that put a lot of pressure on me, and I didn't particularly enjoy that part. I have to commend the people the years that I worked with ABRA. The team of people was amazing and really supported our efforts. And although I've said just previously about needing more support, I felt they were stretched because they had all of these schools and these programs and they had so much bouncing around to do… and I don't know how they did it and I do commend their efforts, but sometimes I felt that I personally needed a little more support.

Another pattern that emerged from the responses on the nature of the training and in-class support was that many reported an increased level of confidence with technology. They attributed the support they received to an increase in their comfort level with not only the ABRA program but with the use of technology in general. This pattern of increased comfort with technology relates directly to the theme of ***Self-Efficacy*** under the Personal Domain of Influence, which is discussed in more detail in the next section.

Personal Domain

The following questions from the teacher questerviews shaped the data that was examined for patterns of influence in the teachers' personal domain: *(1) What were your attitudes about using information and computer technology (ICT) in the classroom before your involvement in the ABRACADABRA literacy study? (2) Has your involvement in the ABRA study changed your attitudes and beliefs toward using computer technology in your language arts (LA) lessons with your students? Explain. and (3) Could you explain in what ways the ABRA program fit or did not fit with your teaching philosophy and existing LA program?* The issues relevant to the teachers' Personal Domain of Influence were categorized into the following two major themes: the *Pedagogical Beliefs* and *Self-Efficacy*.

Pedagogical beliefs. *Pedagogical beliefs* refer to a teacher's attitude and viewpoint of what she values as essential components of her teaching practice (Ertmer, 2005). For this study, the theme of Pedagogical Beliefs primarily focused on the teachers' comments regarding their beliefs around the value of using computer technology in their teaching practice. Data coded segments discussing the alignment of the ABRA tool with the teachers' existing pedagogical beliefs also were captured under this theme. Ultimately, this theme encompassed the subthemes of *Alignment Between Teacher Beliefs and the New Resource* and the *Value and Use of Technology.*

Alignment between teacher beliefs and the new resource. This subtheme captured a pattern of teacher comments made regarding ways in which the ABRA tool fit or did not fit with their teaching beliefs. Fifteen participants (68%) made references to how the ABRA resource aligned with different aspects of their teaching philosophies (n = 14 of these references were favourable, while n = 2 of these references were unfavourable). From the sub-population of teachers (n = 15) commenting on this theme, some of the key teaching philosophies mentioned as being supported by ABRA were the following: it was a multifaceted tool that facilitated their delivery of a well-balanced literacy program (33%, *n* = 5); it was a teaching resource that enabled differentiated instruction (20%, *n* = 3); it made learning interactive for the students (20%, *n* = 3); and it was a resource that helped in establishing a stronger home and school connection (13%, *n* = 2). For example, the successful alignment of the ABRA resource with aspects of the teachers' pedagogical beliefs was evident in comments made by ***Teacher N*** (a 55 year old, Grade 1 teacher with 20 years of teaching experience) regarding how she valued being able to meet the diverse needs of her students through differentiated instruction:

> I like to provide variety in my LA program. I find ABRA a wonderful tool. It has a terrific variety of activities and some are more challenging than others, and so it addresses multi-levels of learning, and I can differentiate my lessons with it.

Teacher B (a 44 year old, Grade 1 teacher with 24 years of teaching experience) provided an example of how ABRA fit with her goal to deliver a balanced literacy program:

> …but it [ABRA] wasn't like my sole tool [pause] and that is important to state too, because in order to have a balanced literacy program you have to provide many opportunities and different situations where children can learn. So ABRA was, and still is, a great literacy resource for my class, as it hones in on all of the early literacy skills I'm working on with them.

Teacher S (a 51 year old, Grade 1 teacher with 28 years of teaching experience) emphasized her belief in using resources that provide interactive learning opportunities for her students by stating, "I believe learning is doing and doing is learning. When students are interacting with one another through play or with ABRA they are engaged and learning." ***Teacher M*** (a 53 year old, Grade 1 teacher with 7 years of teaching experience) provided the following comments on how ABRA was used to foster a stronger home-school connection with her students and their parents:

> Also, the children would use it at home. I've given the website out to the parents to use it home at the beginning of the year, and I have even set it up on the school website. I also at one point had a little blog where I was putting in the homework… so I had different websites there, and so the kids could just click on it with their parents and get extra practice whenever they wanted.

In contrast, some of the teachers who no longer used ABRA noted how they did not find it to be an effective resource for them. For example, ***Teacher O*** (a 31 year old, Kindergarten teacher with 8 years of teaching experience) discussed how ABRA did not align well with one of her key beliefs about being a facilitator to learning:

> All of the building blocks of reading were useful, but any student that required remedial use found the program confusing as it had too many options. Students couldn't remember the icons meanings, so lessons ended up being very adult directed and this doesn't fit with my current pedagogical belief that my role is that of a facilitator. I want students to learn from exploration and collaborative work.

While ***Teacher J*** (a 45 year old, Grade 1 teacher with 20 years of teaching experience) discussed how she prefers a quieter learning environment for her students, and that she found that her experience using the computer-based tool failed to facilitate an environment for optimal learning:

> Well first of all, I'm not that comfortable with going to the computer lab, number one… and using ABRA, number two. My experience with ABRA in the computer room wasn't a very positive one. I don't feel that the kids are ready yet to go up there. And when we are all in the lab at the same time with two or three kids sharing a computer, and not being able to hear properly, no [pause]… no, not in that kind of situation. It was just too much noise and they can't learn that way.

Value and use of technology. This subtheme was comprised of teachers' comments on the value and use of technology in their practice. During the questerviews, the researcher asked all of the respondents (100%) two direct questions concerning computer technology use (see the first two questions above), which provided some data for this subtheme (n = 22 teachers made references towards factors within this theme; $n = 22$ teachers made positive references, $n = 6$ teachers made some negative references).

When asked if their involvement in the ABRA study led to a change in their attitudes around the value of technology in the classroom, 50% of the teachers ($n = 11$) said yes it did; 23% stated that their ABRA involvement reinforced their existing beliefs about the value of technology in the classroom; and 27% stated that their involvement in the ABRA study did not influence their existing attitudes regarding the value of technology in their classroom practice.

When responding to questions concerning their beliefs about the value of technology, teachers also provided comments regarding their *pattern of use* of technology, before and after ABRA (see Table 17 for a summary of teachers' self-reporting of their technology use before and after their involvement in the ABRA study).

The teachers' responses were split evenly when reporting on their use of technology prior to their involvement with ABRA, with 50% ($n = 11$) reporting that they did use technology, and the other 50% reporting that they did not ($n = 11$). Of those teachers who answered *yes* that their involvement in the ABRA study had led to positive changes in their technology use, most (82%, $n = 9$) had avoided using technology prior to ABRA, while others (18%, $n = 2$) stated that they had used it minimally. Furthermore, at the time of follow-up, the number of participants who claimed to have integrated technology into their classroom practice had moved from 23% before ABRA to 77% after their involvement in the ABRA study.

As mentioned previously, some of the teachers reported that their beliefs regarding the value and usefulness of technology in the classroom did change after their involvement in the ABRA intervention study. For example, ***Teacher N*** (a 55 year old, Grade 1 teacher with 20 years of teaching experience) commented that prior to being part of the study, she did not consider technology to be a suitable teaching tool, but this changed after she began using the ABRA resource:

> Before using ABRA I did not think using technology was appropriate... before the ABRA study, well I guess I really hadn't investigated and I didn't know what was out there. I have since changed my mind. I think it is a great tool to use to reinforce phonics and decoding. It also helps to develop comprehension.

Table 17

Summary of Teachers' Self-Reporting of their Technology Use Before and After their Involvement in the ABRA Intervention Study

ABRA Use at Follow-up	Currently Using ABRA												Part of Repertoire				No Longer Using ABRA					
	B	F	H	I	L	M	N	Q	R	S	U	V	A	C	E	T	D	G	J	K	O	P
Before ABRA Study Involvement – Use of Technology																						
Avoidance	*	*	*				*					*	*	*			*	*	*	*		
Minimal					*	*		*	*		*											*
Integration				*						*					*	*					*	
After ABRA Study Involvement - Use of Technology																						
Avoidance																			*			
Minimal											*						*	*				*
Integration	*	*	*	*	*	*	*	*	*	*		*	*	*	*	*				*	*	
Study involvement influenced this Change																						
Yes, No or Reinforced	Y	Y	Y	R	Y	R	Y	N	R	R	Y	Y	Y	Y	R	N	Y	N	N	Y	N	N

Note. Summary of teachers' self-reporting of their feelings about their own use of technology before and after their involvement in the ABRA intervention study. Categories of technology use were adapted from: Mumtaz, S. (2006). Factors affecting teachers' use of information and communications technology: A review of the literature. *Journal of Information Technology for Teacher Education, 9*(1), 320.

Teacher V (a 28 year old, Grade 1 teacher with 6 years of teaching experience) also questioned the value of technology in her first grade classroom:

> Before, I questioned whether technology was really necessary in a Grade 1 classroom where the focus was learning to read. I was very happy to get a program that does contribute to learning to read. It made me much more willing to use technology in the classroom.

Teacher F (a 57 year old, Grade 1 teacher with 20 years of teaching experience) described how before her exposure to the study, she doubted if a computer-based program could adequately supplement her literacy program:

> Before being involved in the study, I was not using technology in my literacy program… I didn't believe there was anything out there that would supplement as well with my existing program. During the study, I could see first-hand the value of kids learning now through the use of technology. I could see how engaged they were and how much they enjoyed it and how it was never a struggle to get them to work on ABRA.

Some teachers reported that prior to their involvement in the ABRA study, they were using other computer-based programs in their classroom; however, their use of computers was limited, and it was primarily for technology exposure for their students, rather than to enhance their students' literacy program. For example, ***Teacher K*** (a 41 year old, Grade 1 teacher with 14 years of teaching experience) noted, "When we did go to computers it was more to get the students used to being around computers and learning the basics… like how to log on and use a mouse… that sort of thing. We used KIDPIX mostly..."

Teacher H (a 48 year old, Grade 1 teacher with 24 years of teaching experience) shared how her use of technology in her class has changed:

> Before, I generally only used the computer to play KIDPIX or STARFALL. I used it as a reward, so it was a reward to go on the computer, for example, for early finishers and for free time. [It was] not part of my academic program. But now, it is like another subject [pause] math, science, computers, and so on. ABRA is a vital part of my reading program.

Other teachers noted that one reason they decided to participate in the study was because they already had an existing strong belief in the value of technology prior to their involvement in the intervention study. For example, ***Teacher O*** (a 31 year old, Kindergarten teacher with 8 years of teaching experience) commented:

> My objective is to achieve our outcomes. Not to change necessarily the make-up of the child… they come with that kind of background [technology background], so I have to play with that game. So I am a very big proponent of technology in the classroom. We, myself, and the school got involved in the ABRA study because it coincided with our existing beliefs about technology; that is, it is a reality of today's youth and it needs to be integrated into our work with them.

Teacher D (a 46 year old, Grade 1 teacher with 22 years of teaching experience) added that despite not being technically confident, she decided to participate in the study, since she saw the value of introducing technology into her classroom practices:

> Well obviously I thought it [technology] was very important because I embarked in something that I had no I had no idea I was embarking in to [pause]. Right? I had no idea what it would take for me to learn because I'm not computer savvy at all, so that is why I wanted to do it. I wanted something different for the kids, and if the world is moving that way then I really wanted the kids to be part of that movement as well.

Teacher A (a 61 year old, Kindergarten teacher with 17 years of teaching experience) echoed similar sentiments:

> The reality is that these children have to learn about computers; this is the reality of the future. So just because I'm old school doesn't mean they have to be old school, and if I want to stay in teaching then I have to constantly challenge them and expose them to the newest things…And being involved in the study did impact me [as] it really opened my eyes to the possibilities for this style of learning [learning with computers].

Other teachers, like ***Teacher M*** (a 53 year old, Grade 1 teacher with 7 years of teaching experience), noted how being in the study reinforced her existing beliefs about technology use in the classroom:

> I always believed that technology was great for students as they were able to be more engaged… and I used technology to show the students different websites where they could learn while playing games. So that is what I was using it for before, and ABRACADABRA just was another positive example.

This sentiment of reinforcing existing beliefs about the value of technology use was further supported by ***Teacher I*** (a 33 year old, Grade 1 teacher with 9 years of teaching experience), who commented, "For me it didn't change what I believed in about technology but it just gave me another example [pause] like support for what I already felt about technology and using it in class." ***Teacher R*** (a 48 year old, Grade 1 teacher with 27 years of teaching experience) also stated, "I did think technology was appropriate but I was not always able to find materials or programs for students to use independently. And ABRA just facilitated my beliefs as it has been a very successful and easy to use program."

Self-Efficacy. In the context of this study, the theme of *Self-Efficac*y refers to *technology self-efficacy* and the comments made by the teachers that reflected their belief in their capacity to integrate technology into their language art lessons and successfully execute literacy lessons using a computer-based program such as ABRA (Bandura, 1997). Over half of the teachers (59%) made direct descriptive references to how their involvement in the study subsequently impacted their comfort level with computers (n = 13 teachers made direct references to the self-efficacy theme; n = 9 teachers made positive influence references; and n = 4 teachers remarked that their involvement in the ABRA study had a negative or no impact).

Comfort level improved with involvement. This theme of Self-Efficacy is closely related to the earlier examination regarding how the teachers' *amount of technology use had* changed from before and after their involvement in the ABRA study. However, this theme is distinct in that the pattern of data comments that made up the Self-Efficacy theme focused on teachers' descriptive remarks about how their *comfort level* with the ABRA tool and technology was impacted due to their involvement in the study. For example, ***Teacher A*** (a 61 year old, Kindergarten teacher with 17 years of teaching experience) discussed how her comfort with technology improved over the 2 years she used ABRA:

> Generally, my personal experience has influenced my use of ABRA, as I could see it in the 2 years as I worked with it… I started to become more familiar with the different games and activities, knowing more the parts of the program that were best geared to my children… and my comfort level with technology started to improve. So prior I was very nervous about the whole idea [of using computers] because I really was a novice at the whole computer thing, and when I'm nervous about something of course then the confidence is not there… but now I'm able to do my reports and get the kids online no problem… I'm no whiz on the computer, but I feel that I have and can now keep up to the technology.

Teacher I (a 33 year old, Grade 1 teacher with 9 years of teaching experience) added that her existing aptitude and interest in technology was strengthened by her involvement in the study:

> Well, I have always liked technology and I think it is something that is never going to leave us. So for me, my personal interest in technology got me involved in the study and then working with ABRA and the ease and use of the program just made me more comfortable. So for me, I think liking technology and feeling more comfortable with it over time encouraged me to continue using it.

Teacher M (a 53 year old, Grade 1 teacher with 7 years of teaching experience) echoed similar sentiments about how her comfort with the ABRA program increased over time:

> Yes, I'm still using it. I have become more comfortable with using ABRA and I find it easier to incorporate it into my lessons. After using it the last few years I also know it more by heart and what is available. I don't worry about my computer time anymore. I'm confident that I can find activities on ABRA that will work with my group. So it is easier to say "Yeah, okay I can use it in this way or do this activity when we are doing let's say rhyming or sound-matching."

While the following response from ***Teacher B*** (a 44 year old, Grade 1 teacher with 24 years of teaching experience) aptly summarizes how her involvement in the ABRA study facilitated her comfort level with not only the ABRA tool but with technology use in general:

> Before ABRA [pause] I mean, I really could not figure out how to plug the projector into the laptop. I just didn't use the computer in my class that way and now I'm so comfortable with technology. In fact, I'm the computer research person at the school (Teacher B laughs) and I'm the lead teacher for the school board for my sector! And now I use technology every day in my classroom. My laptops are always on. I have five laptops from Concordia that I borrowed and four computers in my classroom and they are always on and always accessible to the kids. It's a very technology-based practice now.

However, 67% of the teachers at follow-up who reported that they no longer use ABRA stated that their involvement in the ABRA intervention study did not improve their technology self-efficacy. For example, ***Teacher P*** (a 45 year old, Kindergarten teacher with 19 years of teaching experience) stated how ABRA did not change her opinion about technology use nor did it influence her existing comfort level with using computers in her practice:

> Use of ABRA did not influence my attitude or opinion about ICT. I have always been cautious about over using computers with primary children. And being in the ABRA study did not change my opinion, and my comfort with computers is still the same as it was before. I see a place for technology, but just in moderation.

Teacher O (a 31 year old, Kindergarten teacher with 8 years of teaching experience) discussed how she already had a strong technology-based background and that her involvement in the study did not influence her comfort level with computers:

> In terms of using technology, it was never an issue for me. I was hired on at this school because of my background in technology. So no. That was never a barrier for me. The content on ABRA was there, so I saw merit in the program, but it just did not work with my students. So I really never used it again after the study ended. And so, no, using the program didn't really change how I felt about technology.

Teacher J (a 45 year old, Grade 1 teacher with 20 years of teaching experience) discussed how her experience in the computer lab during the study was not favourable, and how she would rather still use non technology-based resources over computers:

> As I already talked about earlier, I was never comfortable going to the computer lab… then during the study, my experience with ABRA in the computer room was never enjoyable. It took too long to get the kids up there and then when we were up there we had so many problems with the computers… and because they had to share computers, sometimes three kids to one computer it was just too noisy for them to even hear the activities. So it was frustrating and I never did enjoy it. But I also think I am a little bit of a dinosaur. So for example, I like and I go to what I'm familiar and more comfortable with. So using the easel with flipchart paper or the dry-erase board and the Big Books… you know that kind of thing… I rather use hands-on things more than the computers.

The general pattern to emerge from the comments on self-efficacy was that the teachers who noted that their involvement in the ABRA study did not positively impact their comfort level with technology and computers were also part of the group of teachers who reported that they were no longer using ABRA.

Domain of Consequences

The Domain of Consequences encompasses the influence of salient outcomes on teachers' decisions to continue using a new resource or not. Although acknowledging that teachers can value a diverse array of outcomes, this follow-up study focused specifically on teachers' comments about how their involvement in the study impacted student-related outcomes. The data patterns of influencing factors in the domain of consequences came primarily from teacher responses to the following three questerview questions: (1) *What role did your observations about student learning play in your attitudes towards implementing ABRA into your LA program? Explain.*; (2 Figure 2. Core Conceptual Framework for Studying Effects of Professional Development Figure 2. Core Conceptual Framework for Studying Effects of Professional Development) *Has your involvement in the ABRA study changed your attitude and beliefs towards using computer technology in your language arts (LA) lessons with your students? Explain*.; and (3) *Are you currently using ABRA? Explain what factors have influenced your decision to continue to use/not use the ABRA program*.

Flagging student interest. The theme of *Flagging Student Interest,* coined by Barnes (2005), best captured the discussion of issues and influencing factors of this domain. Flagging Student Interest reflects teachers' comments regarding how the use of the ABRA resource was able to, or not able to, address the needs and the interest of their students. Twenty of the participants (91%) made references to this theme. Their comments were categorized into two subsequent subthemes. The subthemes of *Student Friendliness* towards a new resource and *Student Engagement* further encapsulated the key variables of influence teachers reported as valuable.

Student friendliness. Half of the respondents (n = 11) made references to student friendliness towards the ABRA resource (n = 5 of these references were favourable, while n = 6 of these references were unfavourable). Of those who did not find the resource to be very student friendly, a majority of the respondents (83%) were kindergarten teachers. For example, ***Teacher G*** (a 50 year old, Kindergarten teacher, teaching 10 years) found the ABRA resource not to be as student-friendly compared to other computer-based resources:

> However, sometimes I got bogged down in trying to get 18 little bodies on the same page… you know... finding their way through it [the program]… and some kids were good and we buddied-up but it was difficult at times with the age… And so now, at this time, I do not use ABRA. There are many literacy programs and some are very child friendly. They are easier to use and less complicated for young children.

Teacher P (a 45 year old, Kindergarten teacher, teaching 19 years), during her narrative echoed similar sentiments of why she didn't find the resource as suitable for her kindergarten students:

> It wasn't as easy as I thought for the kids to navigate through, in comparison to other programs… well, not for the early part of the year, I thought it was more suitable at the end of the year, or more for the kids that were functioning at a higher level at the beginning. For early K not as much, I found the kids needed to be more technologically savvy… So [pause] it did not seem to be as effective as others for early K. And later when it is more effective, like January plus, [the] kids are already immersed in and engaged with [other] programs that have a broad range of challenging activity levels. So it has just been easier to continue with what we're already using. Also, I found the graphics a bit child "unfriendly" in comparison.

Other teachers commented on how the student-friendliness of the ABRA resource was one of the positive aspects of the program for them. This positive aspect was evident in ***Teacher F's*** (a 57 year old, Grade 1 teacher with 20 years of teaching experience) observations regarding the ability of students to work and manoeuver independently through the ABRA literacy activities:

> It is user friendly and the kids don't need someone always telling them what to do next and how to do it. Because once they have been shown how to get on it, they can usually manage very well on their own and explore new activities and figure them out. And once they are on, then all of the learning blocks of literacy are accessible to them at their fingertips, as ABRA contains basically everything you need to know to learn how to read.

This sentiment, was supported by ***Teacher H*** (a 48 year old, Grade 1 teacher with 24 years of teaching experience) as well:

> The key reasons for why I continue to use ABRA… well first, it meets all the needs of my students. Second, ABRA is easy to use and they can do it independently and finally, I feel it is an excellent program. I love the way the program makes corrections when a child gets the wrong answer.

Student engagement. Nineteen participants (86%) made comments regarding how the resource facilitated their students' learning outcomes. Many of these teachers specifically discussed whether their students found the ABRA resource to be engaging (*n* = 16 of these references were favourable, while *n* = 4 of these references were unfavourable).

All of the teachers (*n* = 16) who reported that the ABRA resource was still part of their teaching repertoire made positive comments regarding the ABRA tool's effects on their students' learning outcomes. For example, ***Teacher U*** (a 45 year old, Grade 1 teacher with 9 years of teaching experience) noted that she continues to use ABRA because of the following reasons: "I'm using ABRA because the kids love it and are actively engaged in level appropriate activities to boost their reading. I was able to see how much the students enjoyed the games and how the games impacted their reading success." Student engagement was echoed by ***Teacher T*** (a 33 year old, Grade 1 teacher with 13 years of teaching experience) who added, "The program was developmentally appropriate, easy to use, and the students were very engaged" and by ***Teacher B*** (a 44 year old, Grade 1 teacher, teaching 24 years) who provided the following comment:

> I'm using it because it works number one. The children love [ABRA] and are engaged in reading... they are attracted to the characters and they are attracted to the technology. They think they are not working and they actually think they are playing and in reality they are learning while they are playing… so that is probably the main reason I use ABRACADABRA.

The students' reaction to the ABRA program, and their enthusiasm while working on the program, were also noted favourably by ***Teacher C*** (a 34 year old, Grades 1 & 2 teacher, teaching 8 years):

> And just the reaction of the kids when you did take it out or had them go to the computer room… it was something a little bit different from regular textbooks and paper and pencils... it was a little bit more interactive, a little bit more fun… and they are still learning… even though they don't know that they are learning (Teacher C laughs)… so they are engaged. Also the kids' enthusiasm for it, made me want to use it.

Teacher M (a 53 year old, Grade 1 teacher with 7 years of teaching experience) added how the students' engagement in the ABRA activities facilitated improvements in their reading skills:

> The students' enthusiasm for ABRA and the subsequent improvements in their reading and writing skills always motivated me to plan my lessons with ABRA. I love to see the students feel success when playing the activities, especially the weaker students.

While ***Teacher Q*** (a 43 year old, Kindergarten teacher with 20 years of teaching experience) discussed observing similar positive consequences when her students were on-task and engaged in the ABRA activities: "When the kids were playing on ABRA they didn't realize that they were learning, they were just having fun… but I could notice that the more they became engaged in the program that their reading skills were improving."

In contrast, ***Teacher G*** (a 50 year old, Kindergarten teacher with 10 years of teaching experience), who no longer is using ABRA, noted how she found her students rather disengaged due to difficulties manoeuvring through the program:

> I did not feel that the program benefited my students more than it frustrated them. The students were quite young and did not find it easy to transition from one area to another. They would get stuck in the program and then lose interest.

Students finding success. Another pattern of responses worth noting was the teachers' responses to the following question: *Did you notice that some students more than others appeared* *to benefit more from its use? If so, could you describe the students whom you found were successfully using ABRA?* Every teacher responded to this question. So, although the responses did not uncover an additional factor influencing the teachers' decisions to use or not use ABRA, they highlighted individual differences at the pupil-level to the ABRA intervention.

Teacher responses varied across the board regarding which group of students they believed benefitted most from ABRA use. The responses ranged from struggling students—the students who still needed to develop their literacy foundation skills (n = 7 teacher references made) or the students with attention issues (n = 2 teacher references made)—to stronger students, the students who were academically stronger (n = 3 teacher references made) or the students identified as engaged learners who demonstrated a strong motivation to learn (n = 2 teacher references made) or the students identified as risk-takers who appeared more comfortable with computers (n = 4 teacher references made). Many of the teachers also added that they believed that overall success could be found at both ends of the spectrum, and that every student using the program benefited in some way (n = 7 teacher references made).

Domain of Practice

The following questions from the teacher questerview survey provided the data that was examined for patterns of influence in the teachers' domain of practice: (1) *If you are currently not using ABRA, could you explain the factors that have influenced your decision not to use the ABRA program? Explain* (2) *If you are currently using ABRA, what factors have influenced your decision to continue using ABRA to support your current literacy program?* When extracting teacher comments for factors positively or negatively influencing their domain of practice, the initial overarching theme to emerge was *Supportive School Environment*. This initial theme was subsequently broken down into the following two encompassing themes that were used to categorize the factors related to the domain of practice: *Elements of the Environment of Practice* and *Contributors to the Environment of Practice*.

Elements of the environment of practice. The theme of *Elements of the Environment of Practice* captured teacher comments regarding aspects of their school environment and practice that either hindered or supported their integration of the ABRA resource. This theme was further divided into four distinct subthemes: (1*) Structural Factors*, (2) *Adaptability of the Resource*, (3) *Technology and Computer Access*, and (4) *Just Not Enough Time*.

Structural factors. This subtheme captured comments made by teachers discussing structural and organizational changes within their schools that had impacted their current use of the ABRA tool. Six participants (27%) made references to this subtheme. Each of these participants made comments suggesting ways in which structural changes to their practice had hindered their current use of the resource.

This subtheme was touched on previously when addressing Question 1 in the "Teacher Results" section. Structural factors were discussed when describing the reasoning behind grouping the four teachers (A, C, E, and T) who were not currently using ABRA, but who said that it was still part of their teaching repertoire, with those teachers who were still using ABRA at the time of the follow-up (see page 182). As previously noted, two of the teachers mentioned that they were teaching grades for which they felt ABRA was not suitable, while a kindergarten teacher reported that her school was experiencing a population explosion that resulted in the primary classes receiving limited computer time that year. ***Teacher I*** and ***Teacher K*** also commented on how their current assignments, to teach older students, impacted their use of the ABRA program. For example, ***Teacher K*** (a 41 year old, Grade 1 teacher, teaching 14 years) stated that:

> It [ABRA] was very useful but as I changed levels, it became more difficult to keep using it… I was teaching older grades and once I started teaching grade three and grade four, I didn't see my older kids needing to work with ABRA.

Teacher I (a 33 year old, Grade 1 teacher, teaching 9 years) described how her teaching assignments, along with other factors, influenced her use of the resource from year to year:

> I can say the greater pressure to hit curriculum expectations in a different manner and the limited amount of availability on the computer, along with teaching higher grades has influenced me using it less over the years since I was involved in the study… but this year I'm back to teaching the younger ones and we have a new sign-up procedure for the computer room, so I'm using it a bit more this year than last year.

Adaptability of the resource. The theme of *Adaptability of the Resource* captured teachers' comments regarding the ease of integration of the ABRA resource into their existing literacy curriculum. Fifteen participants (68%) made references to this theme (*n* = 11 made positive references, while *n* = 4 made negative references). The four references regarding the lack of ease of integration of the ABRA resource were made by teachers no longer using ABRA.

As mentioned previously, teachers that still had ABRA as part of their literacy repertoire were those who commented positively on the adaptability of this resource. For example, ***Teacher R*** (a 48 year old, Grade 1 teacher with 27 years of teaching experience) commented on the ease of use of the program and how it fit into her existing literacy practice:

> It has been a very successful and easy to use program. The students enjoy working on it and it fits in well with my literacy program. ABRA is easy to use in the lab or my classroom. It was just good to find something they could learn on and do fairly independently.

Teacher S (a 51 year old, Grade 1 teacher with 28 years of teaching experience) echoed similar sentiments while noting the flexibility of the resource:

> It is relevant, flexible and meets the needs of a wide range of abilities from beginner to at grade level. It is excellent for intervention work, skill building, and review. It is web based and accessible. It fits with what I believe about literacy. Kids love it and it fits the curriculum.

The ability of the resource to complement existing literacy practices already in place was captured by ***Teacher F***'s (a 57 year old, Gr. 1 teacher, teaching 20 years) comment:

> Well, we would teach a lesson and then they would go into ABRA and work on an activity, which would reinforce the literacy concept we were working on in class. So it fits perfectly… it really complements the two main literacy programs I mentioned earlier… it breaks apart all of the skills that you need for literacy and learning to read and write and the kids have fun learning as they play the games on the computer.

Teacher C (a 34 year old, Grades 1 & 2 teacher with 8 years of teaching experience) noted how the adaptability of new resource is something she looks for:

> It works better if it [a new resource] can mesh with what I'm already doing… I like add-ons like ABRA that makes what I already have a little more fun, a little bit more exciting or extends the knowledge of what I'm already doing in a different way.

Teacher L (a 34 year old, Kindergarten teacher with 10 years of teaching experience) added how ABRA fit well with her existing literacy curriculum:

> It worked because the lessons were related to the concepts I wanted to teach. There is quite a lot of phonemic awareness as well as phonological awareness activities and I work a lot with phonemic awareness. I use it when I introduce concepts, like syllables, when I introduce rhyming… I'll introduce it as a mini-lesson.

Teacher B (a 44 year old, Grade 1 teacher with 24 years of teaching experience) also discussed the ways in which the resource easily integrated into her practice:

> Overall I find it a very useful tool and it very easily integrated into my classroom. ABRA covers the basic skills that we need to teach the children in Cycle One in early literacy and I

> really think that the program fits this perfectly. It has everything that we are trying to teach with regards to early literacy.

For those teachers who stated that they were no longer using the ABRA resource, the lack of ease in the integration of the resource into practice was mentioned as a significant deterrent. ***Teacher J*** (a 45 year old, Grade 1 teacher with 20 years of teaching experience) suggested that, "…it [any new resource] has to be simple and not time consuming… because if it is, I probably will not do it… and it [ABRA] wasn't for me, it wasn't easy for me to use." ***Teacher P*** (a 45 year old, Kindergarten teacher with 19 years of teaching experience) added that she opted to use other resources that she found easier to integrate with her kindergarten class:

> After the study we became familiar with other computer programs, such as STARFALL and SUPERSTAR online, which our principal actually brought to us from a conference that she went to in the States. These programs seemed to be a lot easier to manipulate for the children. As they didn't need that one-on-one direction as much to work on these programs as it seemed they needed with ABRA… or a least initially with ABRA. And that made using these other resources easier and more convenient.

Technology and computer access. This theme captured comments highlighting one of the key constraints the teachers identified as hindering their successful integration of the ABRA. Seventeen participants (77%) contributed comments towards this theme (n = 1 teacher provided a positive reference, n = 10 made negative references, while n = 6 provided both positive and negative references). Teachers shared how access to technology impacted their use of the ABRA resource either at the time of the original intervention study or at the time of follow-up. References were made to the lack of computers in their schools (n = 2); to problems with the technology in the schools (n = 6), from the Internet crashing to the computers being outdated; and to computer-time access both prior to and/or after their involvement in the study (n = 10).

For example, ***Teacher A*** (a 61 year old, Kindergarten teacher with 17 years of teaching experience) discussed how the outdated technology in her school made it difficult to implement ABRA during her time in the study:

> We had a lot of problems due to our computers in the school. A lot of the problems were technologically based which is not my forte… There was a lot of this that was frustrating, especially the computer glitches. I used to have problems with the computers freezing… the projector was unreliable… and when you have troubles with the computers you quickly lose the interest of the students.

Teacher D (a 46 year old, Grade 1 teacher with 22 years of teaching experience) also noted encountering technical issues with the computers at her school:

> I'm also not really using ABRA now because of the problems I had with the usage of our school computers… we have older computers here so being able to use them without a glitch is an issue… so technology problems is a factor for sure.

Some teachers noted that they had greater access to their school's computer lab time during their involvement in the intervention study than they do now. Many also shared that in their schools the computer time was primarily for the intermediate classes. For example, ***Teacher F*** (a 57 year old, Grade 1 teacher with 20 years of teaching experience) explained how her access to the school computers had been limited in comparison to her time in the study:

> Well at our school, the first year that we did it, we were able to get into the computer room every day and do it the way you [the ABRA research team] wanted us to do it. And years after that, because we got so big… we got so big because the school population has exploded over the last few years, we [the primary classes] are never able to have any computer time.

Teacher F added that this limited availability to computer time was a key variable to her current decreased use of the ABRA tool:

> So this lack of access to the computer lab really stopped the whole process of using ABRA as it was intended. I still use ABRA but I can't say I use ABRA the way we used it the first year… Basically, I would use ABRA more if I had more computer time.

Teacher M (a 53 year old, Grade 1 teacher with 7 years of teaching experience) also made reference to the barrier of having limited access to the school computers:

> The only thing really that effects our use is not having as much access to the computers as I would like to. Access to computers is limited here… and we have some negative issues with reliable Internet access too. We can't get the program sometimes when the Internet is down, so in that way too technology limits how much I use it [ABRA].

However, other teachers did mention that in general the number of computers and the quality of technology had improved in their schools in the years since their involvement in the original ABRA study. For example, ***Teacher C*** (a 34 year old, Grades 1 & 2 teacher, teaching 8 years) shared how improvements in the accessibility and the quality of the technology in her school had influenced her current feelings towards the use of technology:

> I'm not as scared now of technology. In the past it was more difficult as I found the schools weren't as technologically advanced at that time. I think now with SMART Boards and everything… if everything is set up it is such an easy tool to use now… because you don't have to worry about screens instead your whiteboard is there… and you just click, click, click and it's in everyone's face right away… so having access to all of this updated technology makes it easier and it's not as scary anymore because now I'm used to it more as I see it more and I've been using it more… so like everything else, practice makes perfect and it makes it easier.

Teacher G (a 50 year old, Kindergarten teacher, teaching10 years), even though she is no longer using ABRA, discussed how her use of technology in her classroom had increased since her time in the study:

> At that time, working with the Kindergarten students was difficult… well, we had to move half way across the school as there were no computers in the classrooms… and it would take such a while to get the little ones organized to go over there… that so it was just easier not going once we were done with the study. And before the study… no, I didn't use technology much then. I knew it was the way to go… I just wasn't sure how to practically get it into my classroom because [pause] well the availability was the issue… as things have progressed and we now no longer have a computer lab, we have computers in the classrooms all of the time, we have the SMART Board… so like, we have more now and we have more regular access to it. So now, I use it [technology, the SMART Board and computers] as a support to what is going on in the classroom all of the time.

Just not enough time. The subtheme of *Not Enough Time* was another key constraint brought up by the participants during the interviews. This theme captured a pattern of the teacher comments made regarding how time impacted their use of ABRA. Ten participants (45%) made references to how time was an influencing factor on their use of the ABRA resource. The references made were unfavourable in nature and demonstrated how time was an inhibiting factor. For example, ***Teacher E*** (a 54 year old, Kindergarten teacher with 32 years of teaching experience) highlighted how other curricular demands limited the amount of time available for exploration on computer-based resources like ABRA:

> We don't have the time. Time is such a valuable thing. We have a new curriculum focus mandated down from the superintendent and that focus is strong on assessment… so strong that it is taking up a lot of our teaching time… and taking up the student learning through exploration time… time they could be on the computers playing and learning on ABRA.

Some teachers added that during their time in the study, and afterwards, due to existing curriculum demands, they found it difficult to find the time to familiarize themselves with the ABRA program as much as they would have liked to. For example ***Teacher K*** (a 41 year old, Grade 1 teacher with 14 years of teaching experience) shared the following:

> For me, I found that I enjoyed it [ABRA]. I thought it was a great idea, but it was also challenging at the same time and I didn't have the time to get to really explore ABRA too much myself before using it with my students… so because of everything else I had to do I didn't have the time to really do it well and to see what it could offer. So for timing purposes… the planning of time was probably why it was really hard to try to make it fit, especially with everything else I have and want to do.

Teacher D (a 46 year old, Grade 1 teacher with 22 years of teaching experience) suggested how due to the limited LA time she had, she could not afford to work on a computer-based program if they encountered technical problems:

> The problem with the year after the study ended was finding the time to use ABRACADABRA. This is probably the biggest problem here… it is the time factor… we only have an hour of LA a day and when those laptops don't work or when the system is down or for whatever the reason is, it becomes very frustrating.

A logistic issue that a number of teachers (n = 5) alluded to was that even though they may be exposed to a new innovation that appears worthwhile to try out, they often chose not to integrate it, since they already have a full curriculum. So, even though they might wish to integrate a new innovation into their practice, they only have so much time in the day to actually do so. For example, ***Teacher J*** (a 45 year old, Grade 1 teacher with 20 years of teaching experience) shared the following:

> I'm not using it because one main reason is the issue with time… because the time in our school is limited to how much we spend on LA… and I'm already using other programs that I believe work well, but they also take up a lot of my time, and if I use ABRA then I have to decide what needs to go.

Teacher K (a 41 year old, Grade 1 teacher with 14 years of teaching experience) revealed similar reasons for using other literacy resources rather than ABRA:

> So a lot of the things I was doing and [was] exposed to were very similar in that the skills they were targeting… and I did like aspects of it [ABRA]… but it was almost like I didn't have enough time to do it all… that is I didn't have the teacher planning time to actually do something with the ABRA program… and because what I was already doing tied in well with the ABRA program… and so I already have a very strong early literacy program, so instead I use these other programs that target the same skills that ABRA does.

Contributors to the environment of practice. The theme of *Contributors to the Environment of Practice* was used to categorize teacher comments regarding individuals within their school environment and practice that supported or influenced their experience with the ABRA resource during and after the intervention study. The teacher comments concerning the key contributors fell into one of the following three subthemes: (1) *Administration Support*, (2) *Ability to Collaborate with Colleagues*, and (3) *Technical Support Providers*.

Administration support. Thirty-six percent of the participants made references to this subtheme (n = 5 made positive references, while n = 3 made negative references). For example, ***Teacher B*** (a 44 year old, Grade 1 teacher with 24 years of teaching experience) shared the following example of how valuable she found her principal's support during her time in the study:

> For me the role that administration played was very important. He was very supportive and I think you have to have a principal who believes in research-based projects, and my principal did. He also was very helpful when it came to setting up the technology in order to make ABRA work. I didn't know anything. I didn't know how to set-up the SMART Board or even how to set-up the projector. Well he [the school principal] came in and set it all up. And he actually sat in with two of the parents at the beginning of my lesson and showed the kids ABRA too and helped with the technology… he really helped with everything. And if it wasn't for him, I don't think I would have moved so quickly through the ABRA program.

It was mentioned earlier that access to the computer lab in some schools was limited for the primary grades in comparison to the intermediate grades. In schools where there was just a sign up for lab time, some of the primary teachers voiced the opinion that their students' computer "exploring" time often took a back seat to the higher grades computer research time. However, some teachers shared how their school administrators were able to influence this variable by ensuring a fairer access to the computers through the implementation of a computer lab schedule. For example, ***Teacher H*** (a 48 year old, Grade 1 teacher with 24 years of teaching experience) noted, "Yeah, our principal's support was very important when we first started because that was what gave us blocked out computer time… so if we didn't have that school administration support we would not have been able to run it."

However, not all of the comments concerning administration involvement were positive. References were made to actions taken by the administration during the ABRA study that negatively impacted the professional autonomy of a few of the participating teachers.

For example, ***Teacher E*** (a 54 year old, Kindergarten teacher, 32 years teaching) shared her story of not having a say about the placement of new technology in her classroom. ***Teacher E*** shared an example of how she wished her administration had consulted her before they installed a SMART Board in her classroom. She stated that the location of the SMART Board installation completely disrupted the design of the classroom in which she had taught for the last 10 years:

> If someone from the board or administration is pushing something onto me without consulting me first, this is something I'm not happy about. It bothers me that people sometimes make decisions for me without consulting me. I like to be consulted before someone hands me something over. My classroom is set around interactive centers and carpet time. Where the SMART Board was placed is completely away from this area… it makes no sense to me… and that is why I probably don't use it as much as I should, or would like to.

Teacher J's story evidently demonstrated how school administrators' actions have the potential to negatively impact a teacher's professional autonomy as well as adversely influence a teacher's experience of being involved in a research intervention study**.** ***Teacher J*** (a 45 year old, Grade 1 teacher with 20 years of teaching experience) recounted her experience of getting involved in the intervention study as not truly voluntary, but rather, as she put it, she was "volun-told" that she would be participating:

> Yeah, I felt I was told to do this and it wasn't my choice… that year I had come from teaching kindergarteners for a long-time, and then all of a sudden I was teaching a very challenging one-two split group. And I felt like that I was told to do it… and I was told it was a voluntary thing and at that point in time I would have rather volunteered to not do it because I felt I had enough on my plate that I didn't need any more. So I didn't really want to be part of the study, but I felt that I couldn't say no… so basically I was volun-told to do this when I had a very difficult group that year. So that was a very negative experience for me. So I went into the study sort of dragging my feet a little bit.

Ability to collaborate with colleagues. Seven participants (32%) made direct references to this theme (n = 4 made positive references, while n = 3 made negative references). Teachers mentioned that when it came to integrating a new innovation, like ABRA, into their practice, having an opportunity to collaborate with one's colleagues was, or could have been, beneficial. Having another teacher at the school or at least at a neighbouring school to share with and bounce off implementation ideas and lessons is something they felt was lacking in their experience during their involvement with the ABRA study. For example, ***Teacher A*** (a 61 year old, Kindergarten teacher with 17 years of teaching experience) shared the positive support she received from one of her colleagues during her time in the study:

> I think I got the most support from the grade one teacher who I was working with… we were both working on the study together… it was a positive thing. It is excellent to have colleagues who you can talk things through with, when engaged in something new like that [the ABRA study].

Teacher R (a 48 year old, Grade 1 teacher with 27 years of teaching experience) added how she found that having a number of her colleagues part of both years of the study was beneficial:

> I don't think we needed that [in-class support] as much at our school, especially the second year of the study when I was using ABRA… just because we already had people who had been trained on it… so I think we just kind of relied on each other and helped each other out when needed. So like, if you didn't understand something, or if you wanted clarification or whatever, we would just ask someone from the year before or another teacher who was also using it.

Teacher U (a 45 year old, Grade 1 teacher with 9 years of teaching experience) shared an example of how she and her colleagues worked together during the implementation phase to solve issues as they arose:

> Yes, because I was able to problem solve with other teachers in my school who were also participating in the study. For instance, we came up with a Popsicle stick system to help the children remember how to login.

Others who lacked the collegial support made references to how this was something they wish they had had during their initial stages of using ABRA. For example, ***Teacher K*** (a 41 year old, Grade 1 teacher with 14 years of teaching experience) shared the following:

> Another thing I found hard was that I was the only teacher doing it [using ABRA] at my school. So maybe if there was somebody else there may have been time for us to work together… like… I could make a game and you could make a game or I make a lesson and you make a lesson and we could share it so now we each have two lessons.

Teacher G (a 50 year old, K teacher with 10 years of teaching experience) also pointed out her lack of opportunity to collaborate with her colleague with this comment:

Because I was the only one using it at my school, and I don't even think anyone in my neighbouring schools were doing it… it becomes a bit more difficult as you have no one to bounce ideas off of or share ideas and frustrations with.

Teacher E (a 54 year old, Kindergarten teacher, teaching 32 years) noted that although other teachers at her school were using ABRA, she still lacked any collaboration time with them:

> There were other teachers at our school using it but I was the only kindergarten teacher… I know some of the others worked together and planned lessons and such… but my break time, and time with my students didn't make it possible to work with the grade one teachers…so yes, I remember thinking sometimes… especially at the beginning… I wish that there had been another "K"-teacher to work with.

Technical support providers. *Technical Support Providers* was the third subtheme under *Contributors to the Environment of Practice.* Six participants (27%) provided comments towards this theme (n = 4 positive references were made, while n = 3 negative references were made). The degree of tech support varied from school to school. Some had in-school support available, so technical issues could be addressed in a timely manner, whereas others had to rely on a school district technician who would make a school visit every other week. For example***, Teacher S*** (a 51 year old, Grade 1 teacher with 28 years of teaching experience) shared her positive experience with the level of the tech support she received:

> We have staff people here, and through the division office. So if I was having trouble, technical troubles, I would phone the Division Office, and the Division Liaison to the project would address it right away. But we didn't have many issues.

Teacher J (a 45 year old, Gr. 1 teacher, teaching 20 years) was in a school where the technology was a bit outdated, and computers freezing up was an issue, so she shared how she would have appreciated more technical support:

> More access to technical support would have been helpful because the computers we had in the school were upstairs and they just kept crashing and also we didn't have reliable Internet then… so that was a big problem… And even though I'm more a bit more computer literate now than I was then… I so was not comfortable with computers back then, so I needed the technical support… so it would have been better if there was someone there that knew what they were doing (**Teacher J** laughs).

Teacher N (a 55 year old, Grade 1 teacher with 20 years of teaching experience) reported on her mixed experiences with the technical support she received during her involvement in the study, noting that it took a while for them to come in, but once the service she requested was completed, it made accessing the program on the computers a lot more efficient for her students:

> The main difficulty I had was that the grade ones don't always know how to log onto the computers. You know they are there with their codes and all, and then we always had to type in the ABRA address and that wasted so much time. And so then I said to my principal, "Can we please get it put onto favourites"' He put in the request right away, but it took a long time for them to actually send someone. It took a couple of weeks before the school board techie finally came and put it on our favourites… that was frustrating. But then when it was done, it was much easier for my students to get on quickly after that.

Integration of Themes and Subthemes

Previous sections have shared the response-to-intervention (R*t*I) narratives of the follow-up teacher participants. Each of the subthemes presented above represented factors of influence that ultimately impacted the ABRA teachers' decisions to either continue or discontinue using the ABRA tool after their participation in the ABRA intervention study ceased. A visual representation and summary of the major conceptual themes and subthemes from the teachers' responses regarding factors that influenced their teacher change process is presented in Figure 7. The four Domains of Teacher Change are represented at the top of the top of the diagram, followed by the seven key conceptual themes drawn from the teacher interviews (e.g., self-efficacy, pedagogical beliefs, nature of training, etc.). Also illustrated are the 15 subthemes and 7 key concepts encompassed within the seven larger conceptual themes.

This diagram of themes highlights the key factors that teachers reported influenced their integration of a new technological resource (ABRA) into their long-term practices. In addition, this visual representation illustrates the interconnectedness of influencing factors on teachers practice, and shows how researchers can play a role in influencing teacher change. For example, it shows how the nature of in-class support teachers receive (External Domain) during the implementation of a new resource could lead to an increased comfort with the intervention tool, which in turn could positively impact teachers self-efficacy beliefs (Personal Domain). Taken together, this diagram effectively highlights the key factors influencing the teachers' R*t*I experience and provides a visual summary of the positive and negative factors that impacted the ABRA teachers' teacher change process and their responsiveness to the ABRA tool.

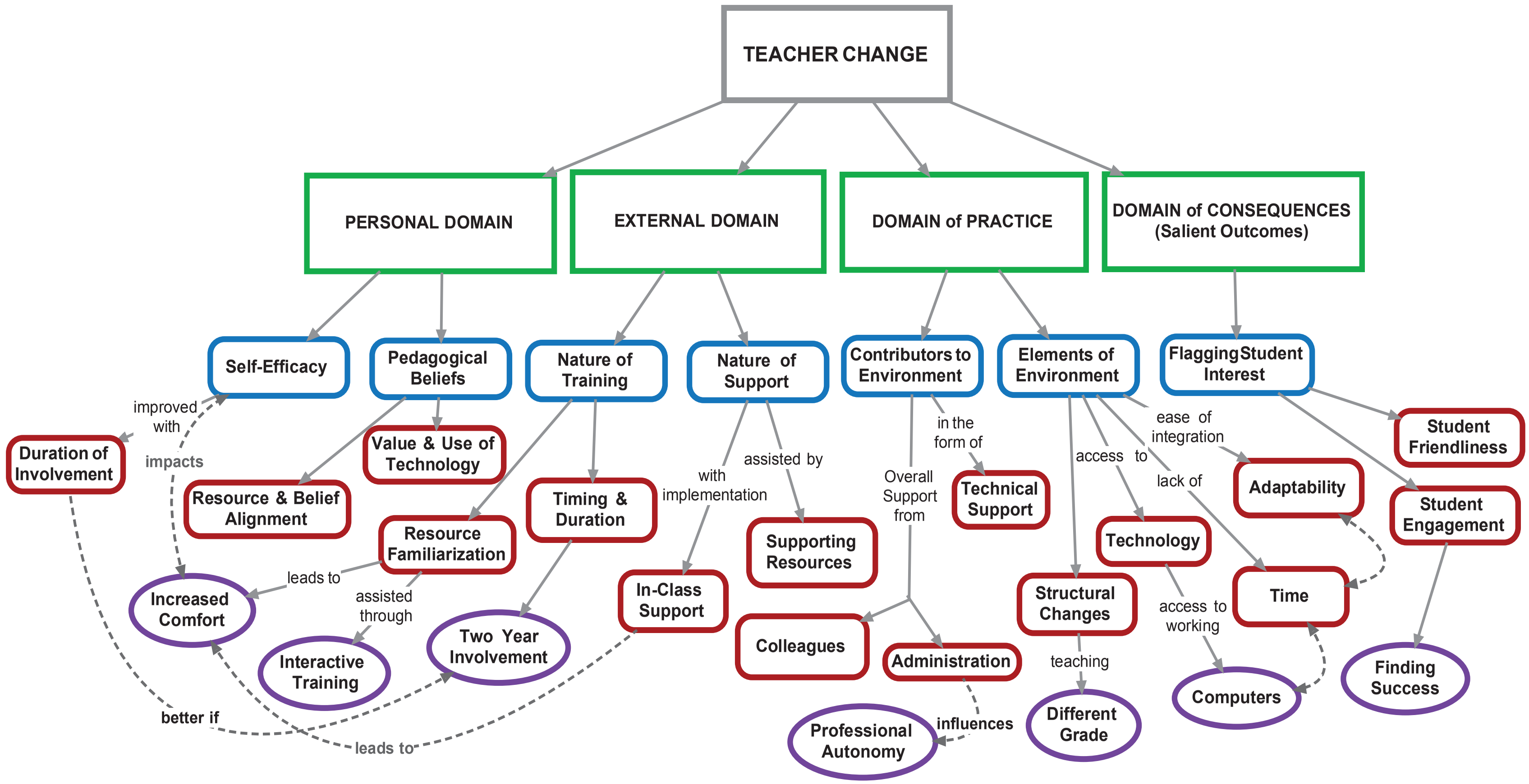


Figure 7. Illustration of categories of themes and subthemes from the teacher data. This figure is a visual representation of the major conceptual themes and subthemes from the teachers' responses regarding factors that influenced their teacher change process. The top green rectangles represent the four Domains of Teacher Change. Next, the blue level of rectangles represents the seven key conceptual themes within the domains of change. The red rectangular nodes represent the subthemes encompassed within the seven larger themes. The purple circular nodes represent some of the key coded concepts within the subthemes.

Results for teacher-level question 6. Is there an identifiable trajectory order of the four teacher change domains (i.e., external domain, domain of practice, domain of experience, and domain of consequence) that predicts whether teachers will integrate a technological innovation into their long-term practice? At the end of the teacher interviews, participating teachers (*n* = 19) were asked to select a linear trajectory teacher change model that they felt best represented their own process of incorporating a new innovation into their literacy practice long-term. As noted previously, Model C, the cyclical interconnected model of teacher professional growth, was never verbally described or visually shared with any of the participating teachers. Model C is a more full and more intricate model, so the concern was that the participants would automatically favour the cyclical model over the two linear models.

The teachers were asked to choose between **Model A** (Change in External Domain ➔ Change in Personal Domain➔ Change in Domain of Practice ➔ Change in Domain of Consequences) that proposes that teachers believed in the value of a new resource *before* implementing it into practice; or **Model B** (Change in External Domain ➔ Change in Domain of Practice ➔ Change in Domain of Consequences ➔ Change in Personal Domain) that suggests that teachers believed in the value of the new resource *after* trying it out and seeing the positive outcomes of the new resource on desired salient outcomes.

Model B was the trajectory change model selected by 58% of the participants, while 42% selected Model A. Teachers who reported that ABRA was still a part of their teaching repertoire were more inclined to select Model B (62%) over Model A (38%). The teachers who reported that they were no longer using ABRA were evenly split (50% selected Model A, and 50% selected Model B) when selecting the trajectory model that best reflected their teacher change process. The chi-square test of independence also was performed to examine whether a statistically significant relationship existed between teachers' current use of the ABRA resource and the teacher change trajectory model they selected as best reflecting their decision to continue using ABRA as part of their teaching repertoire or not. The chi-square distribution was adjusted with a Fisher's Exact test (χ^2 (1, N = 19) = .224, p = .506). The findings indicated a trend towards a greater proportion of teachers reporting that they were more likely to believe in the value of a new resource after seeing the benefits of the new resource on student outcomes; however, this finding was not statistically significant in the present sample.

When the 19 teachers had an opportunity to create their own drawings reflecting their personal teacher change processes, many of them modified the two linear trajectory options (Model A or B) to make them a more interconnected process between the four domains of teacher change. Specifically, 26% of the participants (n = 5) created a circular model, reflective of Model C, in which all four domains of the teacher change process were interconnected. Another 47% of the participants (n = 9 teachers) illustrated various degrees of interconnectedness and reciprocal relationships between the four domains of teacher change. For example, some of the participants illustrated interconnected and reciprocal trajectories between their Domains of Practice and of Consequences (26%); their Domain of Practice and Personal Domain (21%); their Personal Domain and Domain of Consequences (12%); and their Personal Domain and External Domain (11%).

Summary of Teacher-Level Findings

This chapter has chronicled the data analysis process and identified key findings regarding factors influencing teachers' long-term responses to being involved in a computer-based intervention study. Over 70% of the respondents maintained that ABRA continued to be part of their teaching repertoire, and that the tool was a good fit with their existing literacy program. The findings revealed that the teachers who were identified as poorer implementers of the ABRA resource during the intervention phase were less likely to continue using the program after the intervention study ended.

This chapter also dealt with organizing and managing large amounts of data, in part by applying a template approach and using a deductive thematic analysis. This process yielded a rich and detailed account of the experiences the participating teachers shared regarding factors that impacted their decisions to continue or discontinue using the ABRA resource. The seven conceptual categories discussed in this chapter highlighted the key factors influencing the teachers' R*t*I experience, whereas the 16 subthemes provided a richer description of the nuances and the positive and negative experiences that impacted these teachers' teacher change process. The following chapter considers the potential meanings underlying the data examined in Chapters 7 and 8, and offers recommendations and suggestions for future research directions.

Chapter 9: Discussion

Overview

The purpose of this dissertation was to apply a response-to-intervention (R*t*I) lens and examine both student-level and teacher-level long-term responses to being part of an intervention study. The R*t*I analysis was at a Tier 1 level of investigation, since the ABRA program was introduced as a supplement to the already exisiting Tier 1 classoom-level literacy programs. At the pupil-level, this follow-up study examined the enduring effectiveness of the ABRA intervention, and investigated whether the short-term reading gains achieved by students were maintained up to 4 years later. In addition, students' R*t*I and the impact of potential demographic factors (ethnicity, socioeconomic status, and sex-differences) on their short-term and long-term literacy gains were assessed. At the teacher-level, this study explored teachers' responses to being part of the original ABRA intervention study and the factors influencing their decision to continue or discontinue using the ABRA program. Previous chapters introduced the investigation undertaken, presented and analyzed the literature related to this thesis, defined and described the methodology, and presented the results of this follow-up study. The goal of this section is to highlight and discuss the study's findings in relation to theory and previous research, and to draw appropriate implications and conclusions. The limitations of this dissertation study, its unique contribution, as well as suggestions for future research directions also are provided.

Pupil-Level Discussion of Findings

One of the primary focuses of the pupil-level analysis was to examine the added contribution of demographic variables in predicting students' short- and long-term likelihood of being at risk of reading difficulties after exposure to a classroom-level literacy intervention.

Binary stepwise logistic regressions were run to examine the added predictability of the demographic variables of interest (ethnicity, SES, and sex) in identifying students' R*t*I outcomes.

At baseline, from the original sample (*N* = 1181) the lowest performing students (*n* = 362) were identified as at risk of reading difficulties if their baseline reading-composite scores on three literacy measures (Letter Sound Knowledge Assessment, FRY words, and CTOPP Blending Words [BW] subtest) were in the bottom 30%. At TIME 2 immediate posttest, students' whose percentile scores on the BW reading measure at TIME 2 immediate posttest were below the 30th percentile were identified as non-responders and those with scores above the 30th percentile were considered good decoders. While at time of T3 follow-up, students were identified as at-risk readers if their overall percentile scores on the GRADE Total Comprehension measure were below the 30th percentile.

Students' RtI and potential influencing variables. The first question aimed to see if there were pupil-level variables that could identify which group of students at risk of reading difficulties were more likely to respond to classroom-level intervention: *Are there identifiable pupil-level variables of students at risk of reading difficulties who had a greater response to Tier 1 (classroom-level) intervention?*

In addressing this question, empirical support was found for previous findings in the R*t*I literature. Specifically, this study found supporting evidence that pre-intervention reading-related variables were stronger predictors of responsiveness to intervention than demographic variables (Al Otaiba & Fuchs, 2002; Lam & McMaster, 2014; Nelson et al., 2003). For example, at TIME 2 immediate posttest, students' TIME 1 baseline pre-reading skills were a significant predictor of growth in blending word attainment, explaining approximately 15% of the various found; whereas, the demographic variables contributed approximately 2 to 5% of variance at immediate posttest. While for the whole sample at four year follow-up, baseline pre-reading and listening comprehension skills were both significant predictors for later reading success; explaining 23% of the variance in reading comprehension four years later. Students' earlier reading ability increased their likelihood by 2.45 times of being in the typically developing reader group, and the baseline listening comprehension variable increased the likelihood of being in the typically developing reader group four years later by 1.22 times. Whereas in comparison, the demographic variables of interest continued to only modestly predict anywhere from 2 to 5% of the variance in the reading outcomes at four year follow-up.

Lam and McMaster (2014) found that demographic variables generally did not impact children's responsiveness to early reading intervention and noted that demographic variables were not as relevant as reading-related variables such as phonemic awareness and the alphabetic principle on later reading outcomes. While in an earlier review, Nelson et al. (2003) found demographic variables to be the least predictive in comparison to the other factors examined. The results partially coincide with those of Lam and McMaster in that the demographic factors of interest did not influence students' responsiveness to intervention in as much magnitude as the baseline reading-related variables. However R*t*I is about growth and the findings of this study do show that demographic variables still have some effect on the student samples' growth in later reading attainment, a finding that is more inline with the earlier findings of Nelson et al. (2003). This is a valuable finding, as this is the 21st Century and with all of the advances made in reading education research, policy, and practice, still finding any demographic influences on Canadian students' reading outcomes is noteworthy and deserves a closer examination and further discussion.

Ethnicity effect. At TIME 2 immediate posttest, for both the whole sample and the baseline sample at risk of reading difficulties; there was a small yet significant Ethnicity main effect evident. Ethnicity was found to explain a proportion of the variance in the blending scores at immediate posttest between responders and non-responders of the whole group (1% of the variance) and between the responders and non-responders of the at-risk group (4% of the variance). Specifically, statistical support, be it small, was found for students of Asian background having a raised risk of not responding to the intervention and remaining in the at-risk group in comparison to their White peers.

However, at the four-year follow-up, an Ethnicity effect was no longer evident in either the whole group sample or the baseline sample at risk of reading difficulties. The lack of statistical significance for an Ethnicity effect at four year follow-up could partially be attributed to attrition. Even though selective experimental mortality due to students' ethnicity did not appear to be an issue for the four year follow-up sample, overall attrition of the follow-up sample did result in a smaller representation of minority group participants in this follow-up study. Another possible reason for an Ethnicity effect no longer being present is that a majority of this study's Asian sample came from inner city schools in Montreal with higher concentrations of new immigrants and thus the ethnicity effect may also have been a second-language issue. Other ESL language studies have found that potentially at-risk kindergarten students from ESL homes can eventually catch up to their non-ESL peers by grade 3 if they are provided early enough with a solid literacy foundation and they do not have any preexisting learning issues (Foster & Miller, 2007; Stuart, 2004). It would be reasonable to assume that additional compensatory work may have taken place with the students at risk of reading difficulties over the last four years. Suggesting that an Ethnicity effect has the potential to fade away as a result of positive educational experience. However, this suggestion is only speculation, as data on students at risk of reading difficulties receiving additional compensatory reading support over the last four years was not collected for this study.

Socioeconomic status effect (SES). Overall, the current study demonstrates that SES is predictive of stronger reading outcomes, as the influence of mother's education (this study's proxy for SES) was evident on the reading outcomes at both immediate post-testing and at time of follow-up. At TIME 2 immediate posttest and TIME 3 four year follow-up, approximately 2% of the variance in reading outcomes between responders and non-responders from the whole group, and 5% of the variance between the responders and non-responders of the group at risk of reading difficulties, was attributed to an SES effect. Specifically, the odds of having stronger reading skills increased for students with college and university educated mothers, as this group was twice as likely to respond to classroom intervention than students whose mothers had no post-secondary education. These reported results imply that mother's education was found to be a reliable predictor of growth in reading skills development.

While acknowledging that family income and education do not directly shape children's academic achievement per se, the SES findings, using mother's education as the proxy of SES, showed that a portion of the growth in attainment in literacy skills can be predicted by mother's education support previous studies that have found significant and positive associations between family SES and children's later language and reading abilities (Aikens & Barbarin, 2008; Foster & Miller, 2007; Hagans & Good III, 2013).

Sex-differences effect. At TIME 2 immediate post-testing there was no significant difference between males and females in their response to being part of the intervention study. As the immediate R*t*I literacy outcomes were found to be similar for the male and female students in both the whole sample and the sample at risk of reading difficulties analyses. Though the early literacy literature base largely cite the sex advantage in favour of females (Below, Skinner, Fearrington, & Sorrell, 2010; Robinson & Lubienski, 2011), other studies have demonstrated that the difference is less significant or nil in earlier grades (Matthews, Ponitz, & Morrison, 2009; Savage & Carless, 2004; Singleton & Thomas, 1999). The results support and extend the early intervention research findings of Savage and Carless (2004) and (Singleton & Thomas, 1999) who similarly did not find a sex-effect in the early years of grade school. One possible explanation for such contrary finding in the literature could be in identifying which type of reading instruction is used, which early reading skills are targeted, and how are they assessed in the early literacy intervention studies; as Logan and Johnston (2010) found that the assessment outcomes for boys were just as favourable as their female peers when phonological-based intervention approaches and assessments that focused on decoding skills were used.

At TIME 3 follow-up the findings were consistent with the greater body of previous reading research literature that has shown that achievement disparity in reading outcomes between males and females in later grades (Robinson & Lubienski, 2011; Savage et al., 2007). At TIME 3 four year follow-up the results showed that for the whole sample an effect in favour of female students was evident. Approximately 2.5% of the variance in TIME 3 follow-up reading outcomes between responders and non-responders from the whole group, and 3% of the variance between the responders and non-responders of the at-risk group, could be attributed to sex differences.

Specifically, the analysis found that an effect of sex-differences appears to be a predictive factor of future reading difficulties. As children progress through to the later grades there was a trend of boys having a greater likelihood of demonstrating reading comprehension difficulties.

This finding supports earlier study findings (Robinson & Lubienski, 2011; Savage & Carless, 2004; Savage et al., 2007) showing that as children progress through grade school there is a greater probability of females developing stronger reading skills than their male peers. Suggesting more work is needed with boys on their listening and reading as early as kindergarten and grade 1, in order to prevent boys from falling behind their female peers in later years.

Influence of teacher's level of implementation on pupil-level outcomes. The second pupil-level research question addressed was: *Does a teacher's level of implementation of ABRA ELA lessons during Tier 1 intervention stage impact the influence of pupil-level variables on student learning?*

In comparison to the original 2007-2009 ABRA study (Savage et al., 2013) this follow-up study used a slightly larger dataset, as it included all participating classes at baseline analysis including those classrooms that were not part of the final random control trial paper. Similar to the original paper results, the overall findings of this study found robust effects for level of implementation. Consistent with previous studies examining the effectiveness of technology-based interventions in the classroom (Archer et al., 2014; Musti-Rao et al., 2015), the outcomes with this larger dataset reiterate the fact that technology use needs to be well implemented. As significant intervention effects were only found for students from classrooms where technology was adopted and above, and not for those students who were part of classrooms where the technology was poorly implemented.

The proportion of treatment responders at TIME 2 immediate posttest in this present study (77%) is within the response rate range of between 60 and 80% of other published studies for at-risk poor readers (Coyne et al., 2004; Torgesen, 2000); thus this finding demonstrates the effectiveness of the ABRA intervention in the hands of the classroom teachers. For both the whole sample and the sample at risk of reading difficulties there was a main effect of the ABRA intervention. Specifically, the results found that students were twice as likely to have stronger blending skills if they were from classrooms where the level of implementation of the ABRA resource was stronger (IFM-Good). Thus, the intervention was successful in terms of where implemented well it was showing up as effective at immediate post-intervention. Despite the fact that Suggate (2014) found that experimenters tend to produce larger short-term effects out of their interventions than classroom teachers, the findings of this study clearly demonstrate that sound literacy interventions implemented well by classroom teachers can be effective in producing significant effects on students learning outcomes.

However, even though an independent effect of the ABRA intervention was found, ABRA does not appear to be narrowing the gap between responders and non-responders. As the outcomes of this study support previous research that found that reading interventions tend to benefit majority of readers despite students at risk status at pretest (Suggate, 2014). The findings indicate that the ABRA intervention did not selectively improve the reading outcomes of those children who were at a greater risk of reading problems. For example, the non-responders at TIME 2 immediate posttest had a greater likelihood of being children with mothers with lower educational backgrounds and from certain ethnic groups; however, while there was a main independent effect of treatment at immediate posttest, there was no interaction effect between the ABRA treatment and the demographic variables, suggesting that the ABRA intervention is not adding extra protection for students who are at risk of reading difficulties based upon social class and ethnicity. In this study, the short-term structural demographic effects continue to exist even after students were exposed to an effective intervention. While the ABRA intervention did have an overall impact on all student groups it did not have a selective impact on the responsiveness to intervention of those students who were identified most at risk of reading difficulties.

ABRA and long-term intervention effect. The third pupil-level research question addressed in this study was: *Does the ABRA program provide protection from long-term reading development problems for students at risk of reading difficulties*? In this follow-up study, immediate post-test responsiveness to intervention did not predict students reading comprehension success in later years. Neither for the whole sample or the baseline sample at risk of reading difficulties was an intervention effect apparent at the time of follow-up. While the short-term effects of the ABRA intervention on students reading outcomes were evident and produced similar patterns as those reported in earlier studies of ABRA (Comaskey et al., 2009; Savage, Abrami, Hipps, et al., 2009; Wolgemuth et al., 2011), the long-term effects of the intervention were not statistically significant four years later.

In a recent ABRA intervention effectiveness study, Di Stasio et al. (2012) in their 1-year follow-up of a within-classroom RCT efficacy study reported significant findings of growth in reading skills at a delayed posttest; however, with a substantially longer interval between TIME 2 immediate posttest and TIME 3 four year follow-up, the results were not able to replicate similar patterns of growth in this study's sample. Although the researcher had accurately predicted that the students who received a higher quality delivery of the ABRA literacy intervention would demonstrate greater short-term reading gains, this study failed to find significant support for the long-term effectiveness of the ABRA intervention. Specifically, the outcomes failed to show that students' responsiveness to the ABRA intervention at immediate follow-up could predict students' later reading growth and attainment. Therefore with the present sample, the ABRA intervention does not appear to be inoculating, as ABRA did not have a selective impact on those most at risk of having reading difficulties, nor does it appear to provide protection from later reading problems from developing. The results support the insulin hypothesis that while positive short-term gains, such as the elimination of at-risk status, may be evident after exposure to an effective reading intervention, these gains are not enduring and can only be maintained with continued evidence-based support (Blachman et al., 2014).

However, the lack of inoculating effect by the ABRA intervention is cautiously reported, as we do not know what happened between immediate posttest and four year follow-up. As noted earlier, another explanation for the lack of any long-term effects of the ABRA intervention is that some compensatory work may have taken place with the control group students at risk of reading difficulties over the last four years that this researcher was not aware of. As there was no longitudinal tracking of additional remedial assistance this study's student sample may have received since their involvement in the original study, and so there is a possibility that some diffusion of treatments for this study's sample could have occurred (Blachman et al., 2014).

Furthermore, while at immediate posttest ABRA did have a significant impact on students' phonological awareness and letter knowledge (Savage et al., 2013) and yet these positive effects did not result in a transfer to broader reading skills post-intervention. One possible explanation for the lack of significant long-term intervention findings on students' later reading comprehension skills could be that the ABRA intervention was conducted towards the beginning of the academic year when many of the ABRA teachers are focusing primarily on the word-level (phonic based) activities over the higher text-level (comprehension) activities with their beginner readers. ABRA is designed to be a balanced literacy resource, yet despite the ABRA research team's deliberate efforts to encourage the use of the higher order skilled activities during the implementation phase (Savage et al., 2013) few teachers actively utilized the

ABRA comprehension activities. A possible explanation for the lack of utilization of the comprehension activities is that many of our teachers were using ABRA with non-readers (i.e., with Kindergarteners or at the beginning of Grade one). Thus these teachers during the implementation phase may have been more focused on word reading and phonics activities while leaving work on comprehension skills for later in the year (Savage et al., 2013).

The data does support a general pattern that baseline decoding and listening comprehension are longitudinal predictors of reading comprehension (Pressley, 2002; Savage, Burgos, Wood, & Piquette, 2015). Early reading and listening seem to be important factors that need to be addressed as early as grade one as they appear to impact long-term reading comprehension development when looking at later reading comprehension success. One practical implication of this finding is that it suggests that early literacy teaching should not focus solely on phonics development; rather a balanced literacy approach would be of more benefit in the long-term. The findings support that a balanced literacy foundation of both early literacy and listening comprehension skills are needed for long term success as a reader (Pressley, 2002; Savage et al., 2015). Furthermore, the pattern observed of early word reading and listening comprehension ability predicting later reading comprehension success supports a longitudinal Simple View of Reading model. As this model suggests that a good foundation for later reading comprehension is predicted by a firm basis in decoding and listening comprehension early on (Aouad & Savage, 2009; Gough & Tunmer, 1986; Savage et al., 2015).

Finally, despite the documented effectiveness of the ABRA intervention on students' early literacy skill development (Piquette et al., 2014; Savage et al., 2013; Wolgemuth et al., 2011), the intervention does not appear to be narrowing the achievement gap between the lowest and highest performing children. Rather in this study, subtle growth in attainment was found to be consistent across readers, from those identified as poorer or at-risk to those identified as stronger readers at baseline. For example, as a group, the students identified more at risk of reading difficulties at baseline (*n* = 137), continued to be poorer readers at TIME 3 four year follow-up, in comparison to their peers (*n* = 330). Thus, no ‘catch-up’ effect was evident in this study’s data results, indicating that we are not providing any additional support for those who are at more of a disadvantage (Coyne et al., 2004). As those students who were at risk prior to the intervention remained at risk at time of follow-up.

This is not an original finding as a continued achievement gap between weaker and stronger students through the schooling years has been well documented (Hoff, 2013; Pfost et al., 2014; Robinson & Lubienski, 2011; Tran et al., 2011). Nonetheless, this finding for a large sample in a Canadian context in this millennium is new and one worth highlighting as it is very relevant to policy. Canada is a developed country with access to good early literacy starts and early reading intervention. Yet, despite this, the findings demonstrate that there continues to be a need for our Canadian school systems to address the issue that some students enter our schools at an academic disadvantage and continue to remain behind in later schooling years.

Further, that the lack of significant long-term effectiveness findings demonstrate that even with the use of well-implemented early literacy interventions there is no certainty of any lasting impacts on students’ later reading outcomes. This study’s results are a reminder that even in the most optimal environments, with a well-implemented and well-documented intervention that has demonstrated to be effective, we are not actually changing the impact of societal influences in the way that we had hoped. This suggests that structural factors continue to be a root cause of inequity in our school systems, and thus there continues to be a need for researchers and teachers to be consciously aware of such inequalities when deciding which evidence-based practices to implement in order to best meet the diverse needs of all their students.

Teacher-Level Discussion of Findings

When it comes to responsiveness to intervention (R*t*I) studies, researchers have generally focused on the "R" in R*t*I by measuring the extent to which students are responding to an intervention. This study set out to demonstrate that a similar emphasis on the extent to which teachers implement the intervention with fidelity and the factors that influence their response and subsequent long-term use of the intervention is also of value and requires examination.

At the teacher level, the primary objective was to explore what factors influence teachers' R*t*I. That is, identifying the factors that play a key role in influencing teachers' response to being part of an intervention study and their subsequent long-term integration of a new resource into their teaching practice. Follow-up data collected from 22 teachers provided the content for this analysis. In addition to the use of standard parametric tests to quantitatively examine the data, this study employed a deductive thematic analysis (TA) approach (Braun & Clarke, 2006) to identify patterns of commonalities across the teachers' *questerview* responses (Adamson et al., 2004). The analysis of teacher *questerview* responses permitted a more robust elaboration of issues teachers identified as impacting their use of the ABRA tool. The aim of the analysis was to identify influencing factors researchers should be mindful of addressing when developing a technology based literacy tool; especially if the end-goal of the RCT study is to produce an effective literacy intervention resource that teachers will implement into their long-term practice.

Teachers' integration of ABRA at follow-up. The first teacher-level research question addressed was: *At follow-up, have teachers integrated ABRA (technology) into their practice*?

The participants reported that overall there was a favourable fit between the ABRA program and the literacy skills their students were learning. Of the 22 participants, 68% reported that there was a good or very good fit, while the remaining 32% of the teachers reported a moderately good fit. Consequently, at the time of follow-up over 70% of the teacher respondents (*n* = 16 out of 22 teachers) reported that the ABRA program continued to be part of their literacy practice in some capacity. Having a majority of the follow-up teacher participants change their practice by incorporating ABRA into their literacy program, places into context that the long-term effects of being part of a field-based RCT intervention study are not exclusive to student outcomes. This finding provides a solid example of teacher change and thus offers support for Slavin et al. (2009) stance of how teacher change is a critical factor that needs addressing when examining the effects of technology interventions on student reading outcomes.

Teacher's level of implementation and long-term use of ABRA. This study also addressed the following question: *Does a teacher's level of implementation of ABRA ELA lessons during Tier 1 intervention influence a teacher's decision to continue using the ABRA tool upon conclusion of the study?*

The results found that the relationship between teachers' level of implementation (IFM) of ABRA English language arts (ELA) lessons during the original study and teachers' current use of the ABRA resource to be significant. In general, the teachers who were identified as poorer implementers of the ABRA program during the intervention phase of the original study were less likely to continue using the program once their time in the study ended.

This finding in part could be attributed to a higher value assigned to the tool by the teachers who implemented well and are still currently using ABRA. This finding supports Mumtaz (2006) hypothesis that the teachers more likely to transform their teaching practice are those who place a higher value for ICT and perceive it to be useful. In addition, this finding supports the four-component theory models of teacher change prediction that those that make the effort to learn about a new resource that aligns well with their pre-existing curricular goals and alter their practice to integrate it well, then they will have a greater likelihood of seeing positive consequences to their efforts and subsequently more likely to continue using the resource (Girvan, Conneely, & Tangney, 2016; Thomas et al., 2009).

From the original ABRA RCT study nearly all of the low implementers were Kindergarten teachers (*n* = 6 kindergarten teachers, *n* = 1 first grade teacher). At time of follow-up, 100% of the teachers identified as low implementers (*n* = 3), were no longer using ABRA. And of those no longer using ABRA at time of follow-up 50 % were kindergarten teachers (*n* = 3 kindergarten teachers, *n* = 3 Grade 1 teachers). Reviewing the follow-up comments of the low IFM teachers who were no longer using ABRA it became clearer to the researcher why many of the teachers rated poorly on the ABRA IFM rubric scale (see Appendix K, ABRA Implementation Rubric) during the implementation phase. For one, several teachers noted that they found many of the ABRA activities to be a bit too advanced or difficult for their students who were nonreaders. Subsequently these teachers primarily utilized the word-level activities and were less likely to explore the text-level and fluency activities offered on ABRA. And though the use of ABRA at both word-level and text-level was encouraged it was not presented to teachers as a requirement of faithful adoption of ABRA (Savage et al., 2013); however, not employing the higher level ABRA activities and not integrating a more balanced-literacy use of the program would have prevented these teachers from rating any higher than a 2 (adoption) on the ABRA IFM scale.

Furthermore, teachers' pedagogical beliefs and teaching styles influenced their integration and use of technology in their classroom. Some teachers noted that they saw their role as the facilitator of creating a learning environment where their students could explore and not be directly taught or be part of structured lessons; instead a number of the kindergarten teachers stated they believed that their students at this age learn best through exploration and play. Thus in some incidences such teaching beliefs evidently would have had the teachers scoring lower on the implementation scale; especially if the observers did not witness any directive teaching incidents and rather primarily observed 'unstructured playtime' on ABRA. For example, during the intervention phase, the sole low implementing Grade 1 was teaching at a bilingual school and her students were predominately coming from full-time French immersion Kindergarten classrooms. This teacher noted that her philosophy of teaching English reading skills on ABRA was for her students to learn to use the ABRA tool through exploration and play, as she contended that they did not yet have the literacy foundation for direct instruction, a similar philosophy held by some of the other kindergarten teachers.

The second part of the enquiry regarding the impact of teachers' level of treatment integrity addressed the following question: *In relation to treatment integrity levels, with a particular focus on adaption of ABRA into their lessons during the intervention phase, is there a pattern in terms of teachers still using ABRA three years later?* This question was intended to find out if a notable difference could be detected specifically between the teachers who ***adapted*** ABRA into their lessons versus those who merely ***adopted*** ABRA during the intervention phase. At follow-up, no statistical evidence of differences could be found between the teachers who ***adapted*** ABRA into their literacy curriculum versus those who merely ***adopted*** ABRA. The practical effect size in the chi-square test of independence analysis ran indicated that a small practical effect, $\phi = .11$ was present even though not statistically significant. Thus there is little evidence of difference in statistical analyses or in analyses of effect size.

Overall, the findings illustrate that the teachers that implement poorly are less likely to continue using ABRA. While if efforts and commitments were made during the intervention study to actively incorporate ABRA into practice, then there was a greater likelihood that teachers would continue to keep the new resource as a part of their teaching repertoire. This finding was evident regardless of if a teachers' implementation was at the *Adoption* or at the *Adaptation* level, as both levels of implementation facilitated a sustained change in teachers' literacy practice.

Factors influencing teachers' R*t*I and long-term use of ABRA. The third teacher-level question aimed to identify potential factors that influence teachers' decisions to integrate ABRA into literacy practice: *Are there common identifiable factors that have influenced teachers' decisions to integrate ABRA into their literacy practice?*

This present study retrospectively examined the process of integration of new technology-based innovations by exploring potential barriers and supports from the perspective of teachers involved in a cluster randomized control effectiveness study of ABRA computer-based intervention study (Savage et al., 2013). Thematic analysis yielded seven larger themes that were used to categorize the teacher data of *influential factors within the four domains of teacher change.* Through the coding and analysis process the researcher found that the teacher participants' perceptions were not easily classified on a simple dichotomous range. Rather their comments suggesting the benefits and concerns regarding potential influencing factors were found to be varied and situational. However, detectable patterns in the participants' responses were found and provided support in identifying potential variables that discriminated between teachers still using and no longer using the ABRA resource at time of follow-up. The next section will describe the contents of the four domains of teacher change, followed by a discussion of the interaction of the four domains in a general model afterwards.

External domain factors of influence. (Nature of Training & Nature of Support)

The majority of the responses regarding the ABRA training and support came largely from teachers still currently using ABRA and were mostly positive in nature. The lack of comments by teachers no longer using ABRA was primarily due to their lack of recall of their ABRA intervention experience, suggesting that participating in the study was something they tried but their involvement in the study did not leave a lasting impression on them.

The general pattern of responses from teachers in this study corroborate findings from Granger et al. (2002) that the goal of introducing new ICT resources needs to go beyond simply providing information about new tools to teachers. Specifically, efforts need to be made to also assist in the transferring of the skills to actual classroom implementation. This study found that teachers who reported that the training they received was valuable and the ongoing support was sufficient to meet their needs felt more prepared and were overall more successful at integrating the tool than their peers who suggested that they would have benefitted from additional training. Thus the quality of training and support during the implementation of a classroom based intervention study is essential to the successful transfer of skills to practice.

In addition, the teachers acknowledgement regarding the need of sufficient time to experiment and explore with a new innovation prior to introducing it to their students coincides with earlier findings by Frank et al. (2011) and Desimone (2009) who found that the quality and duration of exposure to a new resource enabled teachers to develop the necessary knowledge of ways in which they could effectively modify and improve their practice. Similar to previous findings by Wood et al. (2008) and Archer et al. (2014) the teachers reported that increased comfort with the intervention tool of interest and with technology in general also impacted their long-term level of integration of the ABRA program.

Personal domain factors of influence. (Pedagogical Beliefs & Self-Efficacy)

Overall, the findings within this domain provide supporting evidence for Clarke and Hollingsworth (2002) assertion that change in teacher beliefs and attitudes would be evident in the increasing value that teachers attached to the new teaching strategy. In the case of this study's participants, one clear response to being part of the ABRA technology-based intervention was the new pedagogical knowledge that it is possible to provide supplemental literacy instruction using a technology-based resource. All of the teacher respondents noted they saw the value of integrating technology into their literacy practice, however a majority of them noted that prior to their ABRA involvement they did not know how. Many of the teachers also added that having the opportunity to be trained on a resource that would enable them to integrate technology into their practice was their primary incentive to participate in the original study. Several teachers shared that prior to participating in the ABRA study they felt they lacked strong technology skills and had not had the opportunity to develop their technology-supported pedagogy skills. Many were drawn to the study as it grounded their experiences with computers with a content-connected technology based innovation. This finding supports previous study outcomes suggesting that grounding learning experiences in content-connected technology examples can assist in overcoming barriers in technology integration (Hew & Brush, 2007; Hughes, 2005).

As noted, despite several teachers noting their lack of previous exposure and comfort with technology, the majority of the teachers became involved in the ABRA study because of their interest in introducing technology into their literacy practice. This finding coincides with that of Wheatley (2005) who found that teachers' low self-efficacy and lack of confidence in their existing technology abilities could be a motivating factor for change in order to improve their teaching skills and practice. For many, the opportunity to be involved in a computer-based intervention study came just at the right time. As this 'just-in-time' opportunity (Granger et al., 2002) provided them access to a resource, and the needed support to address their immediate concerns of learning how to integrate a computer based program into their existing paper and pencil curriculum.

Furthermore, literature on the topic of teachers' pedagogical beliefs and the value teachers see in technology suggests the need for an alignment between the two (Ertmer et al., 2012; Mumtaz, 2006). Teachers decided to participate in the ABRA study as they saw the potential of introducing a technology-based literacy resource as being congruent with their beliefs about learning. Specifically, teachers noted that they viewed the use of technology as a valuable supplementary source of knowledge that has the potential to complement their existing balanced literacy program, as it can provide an interactive dimension to their pedagogic practice. In addition many noted that technology is the way of the future so they would be doing their students a disservice by not providing them the exposure when it was available.

The transition from teaching academic skills without technology, to teaching with a technology-based resource requires new pedagogical understanding through training, hands-on experience and on-going support (Archer et al., 2014; Mumtaz, 2006). This study found evidence of a transfer effect that could also be deemed as a *halo effect* of ABRA, as involvement in the ABRA study appeared to provide a catalyst of change for a general increase in the integration of technology use in the participants' classrooms. With only 23% of the teachers reporting that they actively used technology prior to their involvement in the ABRA study and increasing to 77% reporting that they were actively integrating technology into their classroom lessons at the time of follow-up. Further, 73% of the respondents reported that their involvement in the study directly influenced their existing beliefs about the value of technology in the classroom.

An identifiable pattern of teacher responses between teachers who continued on with ABRA versus those who discontinued their use of the program was also evident on their comments around self-efficacy. Self-efficacy according to Bandura (1997) is what a person believes about herself and her ability to persevere in a situation or accomplish a task. One key notable difference was that the group of teachers no longer using ABRA were more likely to report that their involvement in the intervention study did not positively impact their self-efficacy beliefs in their ability to use technology in their classroom. Specifically, the general pattern to emerge was that teachers who continued to report feeling uncomfortable in their ability to effectively integrate technology into their literacy practice were less likely to continue using ABRA in the long-term.

Furthermore, teachers existing pedagogical beliefs played a key role in their decision not to fully integrate ABRA into practice. Respondents still using ABRA highlighted ways in which the tool aligned with their existing beliefs; while those no longer using ABRA noted their disuse of the program was in part due to a lack of such an alignment. This finding coincides with those of Opfer et al. (2011) who found that teachers instructional decisions were greatly impacted by their past experiences and prevailing pedagogical beliefs and extend those of Clarke and Hollingsworth (2002) as well as Ertmer et al. (2012) who noted that teachers are less willing to experiment with a new resource, or in the case of the ABRA teachers implement it well and continue its use, if they find that the resource does not complement their existing teaching philosophy.

Domain of consequences factors of influence. (Student Outcomes/Flagging Student Outcomes – Student Friendliness & Student Engagement)

Clarke and Hollingsworth (2002) contend that when teachers implement a new resource into their existing practice their interpretation of the change they observe in this domain is a reflection of their existing value systems. The authors add that the change teachers observe in the domain of consequences is tied to the inferences they draw from their efforts to implement a new teaching practice and how these efforts of implementation influence the salient outcomes they value.

Not surprisingly, the results of this study found a distinct pattern of differences between teachers using and no longer using ABRA was clearly evident in their inferences of how their students responded to the ABRA tool. Teachers continuing with their use of the program, spoke positively about how their students received and interacted with the ABRA program and how they found the program to a good-fit match with the literacy skills they were targeting in their lessons. They mentioned that the tool was engaging and student friendly. Consistent with earlier research the teachers commenting positively about ABRA discussed how ABRA and technology in general added richness to their literacy instruction (Musti-Rao et al., 2015; Wood et al., 2008). Richer in that it was motivationally appealing to young readers, dynamic for attention-struggling readers and facilitated both collaborative and independent learning. While those who were no longer were using the program, predominately the kindergarten teachers, spoke less favourably about the student friendliness of the program and the engagement level of their students. These teachers noted that while the interactive word-level activities matched well with the early literacy skills being taught, they found the selection limited in comparison to the text-level activities, which they found too difficult for their age group. Others noted that due to previous lack of exposure to technology and less developed fine motor skills, many of the kindergarten students had difficulties maneuvering through the program resulting in their disengagement.

The teachers provided a wide-range of responses in their identification of the student groups that they perceived as benefiting the most from interacting with the ABRA program. Responses ranged from students at risk of reading problems and struggling readers, to engaged and highly motivated learners, to risk-takers and those with more developed computer skills, to all of the students equally benefiting. While no identifiable pattern emerged in these responses between teachers continuing to use ABRA or not, the varied teacher responses appeared to be a reflection of each teachers individual pedagogical beliefs of who to teach too and who benefits most from technology-based supplemental interventions.

This evident contrast in teachers' attitude and perspective towards the same tool provides support to others who have found that teachers are more likely to continue using a resource they observe as positively being received by their students (Barnes, 2005; Clarke & Hollingsworth, 2002). This finding also supports those of Thomas et al. (2009) who highlighted the central role of students in the successful implementation of a new resource. Further, the outcomes in this domain support Barnes (2005) findings that students' response to a new resource generates a formative assessment of the resource and this in turn influences the teachers' attitude about the value of the resource and their subsequent use or disuse of the resource.

Domain of practice factors of influence. (Environment Elements & Environment Contributors)

As noted, prior to their involvement in the ABRA study, the teachers generally supported the integration of computers into their classroom practice, however they also identified a number of concerns and limitations that deterred them from using technology in their practice. By examining the teachers' qualitative responses some of the key discriminatory factors between teachers who continued to use ABRA versus those who no longer had ABRA as part of their repertoire fell within the domain of practice environment: adaptability of the resource; technical access and time.

Specifically evidence was found supporting previous research that stressed the value teachers placed on the adaptability and the ease of integration of new resources into teachers' existing practice (Thomas et al., 2009; Wood et al., 2008). Just as Thomas et al. (2009) cited that the flexibility and customizability of a new innovation is an essential facilitating factor in fostering a good 'fit' between a new innovation and teachers existing practice, many of the teachers still using ABRA also commented favourably on how the ease of integration and adaptability of the program into their literacy practice was a key deciding factor for their continued use of ABRA. While the lack of ease in integration of ABRA into practice, was noted as a crucial deterrent for those teachers who stated that they were no longer using the resource.

Consistent with earlier literature, the teachers' comments provided an identifiable pattern suggesting that the factor of *time* can be a key constraint in impacting the integration of the technology-based resource (Ertmer et al., 2012; Wood et al., 2008). The comments made regarding time were unfavourable in nature and demonstrated how time was an inhibiting factor. Specifically, the lack of time to integrate ABRA into their existing practice due to other curriculum demands or because they already had a full curriculum was noted by teachers no longer using the resource.

Although technology and computer access in North American schools has improved immensely over the last decade (Ertmer et al., 2012), problems with technology access continued to be one of the key hindering constraints identified by 73% of the follow-up teachers. Though this factor did not discriminate between those still using versus no-longer using ABRA, it was a definite deterrent for those no longer using ABRA. Of particular interest was the finding that even though provincial curriculums continue to place a greater emphasis on earlier technology exposure for our students, in a number of schools the access to the computer labs or the availability of in-class technology for primary students was often limited in comparison to the older grades. Highlighting previous findings that at the institutional level, external barriers and school contextual factors such as access to technology, lack of technical support and timetabling factors continue to have a great influence on teachers practice and their successful implementation of technology-based resources (Ertmer & Ottenbreit-Leftwich, 2010; Ertmer et al., 2012; Hew & Brush, 2007; Mumtaz, 2006).

A third of the follow-up teachers noted that having a colleague who was also using the tool during the intervention phase or working with a colleague who had had positive previous experience with the tool was an influencing factor for them. This finding supports Thomas et al. (2009) and Rogers (2003) who found that teachers were more willing to implement an innovation already demonstrated as showing some success and that they valued the opportunity to be able to discuss lesson ideas and troubleshoot off one another. Thus, the value of school administrators and research teams facilitating an environment that fosters collaboration should not be overlooked, as teachers benefit from being able to plan, problem-solve and learn from each other.

Testing of theory. For the teacher level of analysis, the theoretical question examined was: *Is there an identifiable trajectory order of the four teacher change domains that predicts whether teachers will integrate a technology innovation into their long-term practice?*

The teachers were asked to choose between **Model A** (Change in External Domain ➔ Change in Personal Domain➔ Change in Domain of Practice ➔ Change in Domain of Consequences) that suggests that teachers believed in the value of a new resource *before* implementing it into practice; or **Model B** (Change in External Domain ➔ Change in Domain of Practice ➔ Change in Domain of Consequences ➔ Change in Personal Domain) that proposes that teachers believed in the value of the new resource *after* trying it out and seeing the positive outcomes of the new resource on desired salient outcomes (Clarke & Hollingsworth, 2002). This study found a trend in a greater proportion of the teachers selecting Model B over Model A, suggesting that the sample of teachers were more inclined to believe in the value of a new resource only after witnessing positive effects of the resource on student outcomes; therefore, favouring the trajectory of influence that proposes that it is the teacher's action and the observations of the consequences of his/her actions that leads to a change in the teacher's attitude (Clarke & Hollingsworth, 2002; Guskey, 1986). However, this finding favouring Model B was not statistically significant in the present sample; that is a teacher's change process did not statistically predict whether a teacher would or would not use ABRA long-term. The chi-square adjusted with Fisher's Exact test (χ^2 (1, N = 19) = .224, p = .506, ϕ = .10) was not significant.

Nevertheless, while this study did not find strong empirical support for whether teachers' attitude leads to action or their action leads to a change in their attitude, there was a distinct pattern of influences from the Domain of Consequences and Domain of Practice on the teacher samples' ultimate decision to continue using ABRA. Specifically, how the teachers perceived the engagement level of their students, the adaptability and ease of integration of ABRA, technology access and availability of time were identified as key inhibiting factors for this study's sample of teachers who were no longer using the ABRA resource. Furthermore, similar to Barnes (2005) findings, many of the teachers who kept ABRA as part of their ELA repertoire indicated that the opportunity to field-test ABRA and observe firsthand their students' positive response to the program provided them the incentive to continue with the change process and modify their literacy practice to include the ABRA resource as part of their reading curriculum. These findings suggest that some teachers need to value the outcomes of using a new resource with their students in order to be motivated to continue with it's long-term use (Clarke & Hollingsworth, 2002).

We acknowledged from the onset of this study that the teacher change process is complex, and even more so with the addition of integrating technology innovations into ones' practice. Therefore, not surprisingly, when the teachers were provide the opportunity to create their own representations of their personal teacher change process models, a majority of the teacher participants (73%) elaborated beyond the two suggested linear trajectory models and indicated various degrees of interconnectedness and reciprocal trajectories between the four domains of their teacher change models. This finding shows that teachers themselves are aware that their decision making process is more complex in nature as they did not choose simple causal models as the best fit. The teachers' comments and illustrated outcomes provide additional support to previous research findings that have argued that the flow of influence between the domains are reciprocal in nature and are better reflective of a more complex teacher change process, rather than an unidirectional linear model (Clarke & Hollingsworth, 2002; Desimone, 2009).

Accordingly, Clarke and Hollingsworth's (2002) teacher change model, **Model C**, may be the more fitting analytical tool to depict the majority of the participants' teacher change process, as model C fits best with the self-descriptions and the absence of a clear winner in the forced choice of linear models. According to this model, change in one of these four domains is mediated and reflected by changes in the other three domains; for instance, Model C predicts that changes in teacher beliefs concerning the efficacy of using a new innovation are mediated by teachers' inferences linking the use of the new tool to salient outcomes. According to Clarke and Hollingsworth, these salient outcomes will predictably reflect a teacher's existing thoughts regarding the objectives of early literacy instruction, and of acceptable literacy practices—that is, the teacher's knowledge and beliefs. Thus, Clarke and Hollingsworth's (2002) model recognizes the value/impact of different contexts on teachers' responses to change, and it explores the complex and interactive nature of teacher change by acknowledging multiple pathways and the individualistic nature of teacher growth.

Limitations of the Current Study

The results of this study should be interpreted with consideration of the following limitations.

Pupil-level limitations. The findings examining the demographic variable influences on students reading skills at four-year follow-up are reported with caution. Though the sample for this study was representative of the original study, there exists a need to report cautiously due to the limited numbers of participants in the follow-up analysis, as the sample size for analysis at TIME 3 four-year follow-up was based on those students who were at risk of reading difficulties at TIME 1 baseline (n = 362), then responded to the intervention at TIME 2 immediate posttest (n = 278), and then were available at four-year follow-up (n = 109).

This follow-up study's power calculation was depended on the original RCT research study's design (outlined in Savage et al., 2013). Due to attrition, the TIME 3 four-year follow-up analysis of the reduced sample at risk of reading difficulties is therefore likely to be underpowered and very conservative for some of the analyses as the small cell sizes have the potential to influence the goodness of fit measure of the logistic regression models run, and thus should be noted and interpreted cautiously (Leech et al., 2015). However, the analysis were conducted in order to be consistent, while acknowledging the methodological limitations for the pupil-level questions focusing on the subsample of students at risk of reading difficulties at baseline available at TIME 3 four-year follow-up.

Further, as only 26% of our four-year follow-up sample identified as being non-white (n = 123) it was necessary to group together a diverse array of ethnicities in order to explore possible ethnicity effects. As noted, the ethnicity *Asian* subgroup included participants who identified as one of the following -- Chinese, South Asian, Southeast Asian, Arab, Korean, Japanese, West Asian and Filipino (Statistics Canada, 2011); while the ethnicity *Other* subcategory included students who identified as either Aboriginal, Latin American, or Black. The limitation of such broad groupings is that it neglects to take into consideration the diversity within each subcategory. However, such groupings were needed in order be sizable enough for analysis, and consequently the subcategories were able to provide some insight on the effectiveness of reading interventions between the white and non-white participants of this study.

The preliminary descriptive analyses indicated that for mother's education as the proxy of SES, there appeared to be a slightly higher proportion of the follow-up sample to have mothers with some postsecondary education rather than those with no postsecondary education. However, despite this more common overall SES trend, this study at four-year follow-up lost a greater number of participants with university-educated parents, which is an unusual trend in research studies. One possible explanation for the loss of the participants with higher educated parents might be that these parents are more mobile as they have more opportunities and incentives to move.

Another possible limitation, is that the follow-up sample tended to be a slightly higher performing group than the original sample at baseline, as mean differences existed on the pretest literacy measures in favour of the four-year follow-up sample versus the sample of students not participating in the follow-up study. This trend could partly be explained by a methodological issue around not having access to the students from two schools in Ontario. Furthermore, another explanation could be that as students get older those students with more reading difficulties are often those less likely to return consent forms to be part of a literacy study. Nonetheless, both of these trends suggest that the results need to be cautiously interpreted.

Further, as there was no direct evidence about the variability or nature of instruction post-intervention, there exists the uncertainty of not knowing what happened in the 4 years between TIME 2 immediate posttest and TIME 3 follow-up. As is the case for the majority of follow-up papers, there was no longitudinal tracking of the possible additional interventions or remedial assistance the student sample may have been exposed to since their involvement in the original study. Therefore one recommendation would be for a stronger longitudinal design in order to incorporate ongoing observations. However, it should be noted that consistent and reliable longitudinal tracking is a difficult undertaking and it almost would have been impossible to find the funding, time and people needed to follow 1000 students across three provinces every year for 5 to 10 year. Nevertheless, without taking into account any additional interventions or services the children may have received, some diffusion of treatments for the sample was a possibility and could have impacted the long-term results (Blachman et al., 2014).

Finally, this study did not attempt to explain why effects were found or not, rather the aim of this study was to highlight if demographic effects, to some degree, play a factor in children's response to reading interventions. The hope is that future researchers can extend the analyses of this study and explain the mechanisms of these effects.

Teacher-level limitations. The power calculation of this follow-up study depended on the original ABRA RCT study's design and as noted above, given the nature of this study's longitudinal design, one limitation of the current research study was that attrition impacted the follow-up sample size. Specifically, for the quantitative analysis the follow-up teacher sample ($N = 22$) was relatively small and thus the results of the chi square analyses were conservative and should be interpreted with this in mind as the contrasts examined at the teacher-level were likely to be underpowered.

However, Howell (2013) contends that the recommended expected frequencies of 5 is a conservative position and that there should be few reservations about violating them; he also adds that the Fisher's Exact Test, which was employed in this study, applies well to cases with small expected frequencies. Furthermore, a qualitative sample size of 22, representing 35% of the original teacher sample, was more than ample for the thematic analysis portion of this thesis and in providing a richer picture of the teachers' experiences (Braun & Clarke, 2012; Creswell, 2014).

Another potential limitation was the retrospective nature of the teacher questionnaires and interviews, as teachers were asked to recall back up to five years. This researcher understood going into the study that complete memory recall for many of the participants might have not been possible. Issues with recall concerning the training and support received during the intervention phase did tend to be more prevalent for teachers who upon the completion of their involvement with the original ABRA study ceased to continue their use of the tool.

The methodological choice not to categorize or analyze any neutral comments made by the teachers during their interviews should be noted, as this researcher acknowledges that the potential to learn from qualitative data that is not mentioned can also be of value (Ryan & Bernard, 2003). A suggestion for future studies would be for researchers to consciously be aware that neutral comments may arise during their participants' responses, and to have a system in place during the interview stage to recognize such comments as they happen and ask additional probing questions.

Finally, when it came to using the teacher change models, it needs to be noted that limitations exist regarding using broad domains to group potential influencing factors together. Specifically, the respective four domains used in this study may not have adequately reflected the richness of the various contexts that could be influencing teachers' adoption of a new innovation into their long-term practice. For example, contextual factors—such as a side-ways transfer of knowledge and support from colleagues using a new innovation and sharing its value (Frank et al., 2011); school principals taking a pro-active role in promoting the use of technology in their schools during and after their involvement in an intervention study (Savage et al., 2010; Wood et al., 2008); and a school district's policy and mandate around technology use (Russell, Bebell, O'Dwyer, & O'Connor, 2003)—still need to be fully explored in order to provide a richer understanding of teacher change within a R*t*I. Continued research examining such factors in greater detail is suggested; such an examination of possible contexts could lead to more of an ecosystemic model of teacher change exploration, which may have been restricted at present by grouping multiple contexts together into one domain.

Implications and Directions for Future Research

At the pupil level, this study empirically addresses the limitations noted in previous R*t*I literature reviews of the need for additional studies examining the short-term and long-term influences of demographic factors on students' R*t*I (Al Otaiba & Fuchs, 2002; Lam & McMaster, 2014; Nelson et al., 2003). One of the key implications of this paper is that it highlights that even in the 21st Century the influence of demographic variables on children's educational success continues to persist, indicating that further work on how best to effectively address the academic needs of our diverse populations is still warranted. To the best of this researcher's knowledge, this study is the first time that demographic variables have really been looked at in the response to intervention context for reading in Canada and have shown to contribute to the prediction of educational outcomes. So while some of the outcomes replicate existing findings, the results extend what has been found into a context that it has not yet been explored. Further, the follow-up findings are a reminder to researchers and educators that we cannot assume that implementing empirically supported reading interventions will completely eradicate reading difficulties in our youth. Simply employing effective reading intervention strategies within our classrooms does not some how chip away at the inequalities that exist in society, as identifiable sociological inequalities continue to impact the educational outcomes of students identified as being at risk.

Another fundamental implication of this study's pupil-level findings is that it highlights the need for ongoing remedial intervention work with those most at risk of reading difficulties in the earlier grades. The findings suggest that a one-shot inoculation model of intervention is not likely to transpire for those most at risk. Furthermore, the insulin-inoculation hypothesis as set up by Coyne and others (Coyne et al., 2004) focuses on the phonics aspect of literacy and neglects the morphological aspects of the deep orthography of the English. Savage and Cloutier (2016) suggest that such selective focus on phonics instruction may hinder students later reading comprehension success, as many words in the English written language deviate from simple one-to-one letter-phoneme correspondence, making it more difficult to pronounce deep orthographies based on how they are written. Therefore, in order to better prevent later reading comprehension and spelling difficulties one recommendation is to support on-going researcher efforts for multi-year sustained intervention efforts, that focus on both phonics training and morphology instruction, in early literacy education (Fletcher & Wagner, 2014; Savage & Cloutier, 2016). Furthermore, in order to better assess the effectiveness of reading interventions, future research should not only move towards multi-year sustained intervention efforts but also continue tracking both the intervention effectiveness with more long-term follow-ups and include on-going monitoring of any additional remedial assistance participating students may receive.

At the teacher-level, this study has practical implication for the field of education, as it pertains to addressing teacher concern related to participating in technology-based classroom level intervention studies and introducing technology based resources into their teaching practice. This study's summary of findings provides an added insight of potential factors impacting teachers' change process. This information could be used to assist change facilitators in gaining a better understanding of intervention development and how better to support teachers in the long-term implementation of a new resource which has shown to positively affect student outcomes.

Education policy directives aim to improve the quality of instruction that students are receiving. Therefore, then researchers who engage teachers in intervention studies must consciously set out to improve the quality of teacher learning. Understanding factors that teachers identify as influencing their decision to implement changes to their practice is of value if we wish to improve the quality of teacher learning. As developers and researchers of new teaching tools, we need to be consciously aware of who will be using the tools we create. The focus must not just be on the creation and development of the tool but equally as important is how can we best support the teachers in implementing the tool into their existing practice. Of specific importance, is providing a tool that aligns well with teacher's pedagogy, is flexible and easily adaptable into different learning environments. This study has shown that a greater awareness of the influencing factors that impact a teacher's long-term use of new teaching resource is of value. As the ultimate goal for developing research-based literacy interventions is so that the tool can be effectively used and easily adapted into a teacher's long-term practice.

In addition, training and support needs to continue being of a focal point in the design and reporting of intervention studies (Archer et al., 2014). If there is one key take home message from this study then that is this, as if researchers are truly interested in the long-term success of technology-based reading interventions then they need to ensure that if planning on asking teachers to open up their classrooms that the supports are in place to make sure teachers gain the most out of their experience. Researchers and school-based professional development efforts should continue focusing on strategies that facilitate change in teachers' attitudes and beliefs towards the use of technology with primary children. These efforts should ensure that new resources can easily be assimilated into existing ELA curriculums. Specifically, teachers should be supported in the integration of the new resource long enough to assess students response to the resource as well as the impact of the resource on student learning outcomes. Ideally, at the teacher level, one recommendation when involving teachers in research studies is that rather than a one-shot participation period, a two-year involvement commitment appears valuable in improving the likelihood that a teacher will continue using the education intervention they were trained on. Thus, it is highly recommended that future researchers consider implementing a minimum two-year involvement loop of teachers, especially if they believe in the value of the educational resource under investigation, and are interested in teachers' continued use of it.

Conclusions and Contributions

Teaching children to read is one of the most significant accomplishments of primary education, as literacy skill development in early childhood provides the necessary foundation for children's long-term academic success (Pfost et al., 2014). Therefore, as long as there are children that need to be taught fundamental reading skills, the development of effective literacy interventions to assist teachers in accomplishing this task will continue to be of importance to literacy researchers, educators, parents and students.

Future classroom-based intervention researchers should continue incorporating a dual subject level of focus in their response to intervention analysis. Specifically, an effort to seek ongoing feedback from the teachers during and following integration of interventions in order to identify implementation-inhibiting factors is recommended. As understanding how both students and teachers respond to being part of researchers' intervention studies strengthens the value of classroom-level research participation, especially if it leads to the successful long-term implementation of empirically supported reading innovations.

An evidence-based early literacy program, such as ABRACADABRA, implemented well, has demonstrated that it can assist in developing the fundamental skills necessary for learning to read (Piquette et al., 2014; Savage et al., 2013; Wolgemuth et al., 2013). The implementation of sound research-based interventions can increase the quality of classroom instruction, especially if the fidelity of implementation of the new resource is high (Archer et al., 2014). Strong fidelity of implementation in addition to the monitoring of students response to the classroom instruction is the key in establishing an effective delivery of Tier 1 literacy instruction.

This study provides a richer picture of the value of factors influencing students and teachers long-term response to being part of a true experiment (RCT) reading intervention. The value of using RCTs to test the effectiveness of literacy interventions is undeniable, as well-executed RCTs are the gold-standard in determining if an innovation works well or does not (Haynes et al., 2012). While knowing if a new innovation can be effective in our classroom is of value, who responds to the intervention is just as valuable. There is also value of extending the research focus of Tier 1 interventions beyond just investigating the effects of the interventions on the student outcomes. An examination on how teachers implement Tier 1 instruction and how they respond to being part of an intervention study is also worth assessment. This study builds upon the existing effectiveness literature of ABRA by providing a longitudinal analysis of demographic factors impacting students' short-term long-term reading outcomes. The mixed methods approach used in this study alongside a gold RCT design was very powerful in getting the rich picture behind an apparently causal IV-DV link. In addition, this study contributes to the ABRA literature by providing a more comprehensive understanding of potential factors that may influence teachers' use of a technology-based resource. In conclusion, this long-term follow-up study extends our understanding of the potential factors that can impact both students' and teachers' response to being part of an intervention study.

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., Savage, R. S., Wade, C. A., Hipps., & Lopez, M. (2008). Using technology to assist children learning to read and write. In T. Willoughby & E. Wood (Eds.), *Children's learning in a digital world* (pp. 129-172). Oxford, UK: Blackwell.

Adamson, J., Gooberman-Hill, R., Woolhead, G., & Donovan, J. (2004). 'Questerviews': using questionnaires in qualitative interviews as a method of integrating qualitative and quantitative health services research. *Journal of Health Services Research & Policy, 9*(3), 139-145.

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*(2), 235-251. doi:10.1037/0022-0663.100.2.235

Al Otaiba, S., Calhoon, M. B., & Wanzek, J. (2010). Response to intervention: Treatment validity and implementation challenges in the primary and middle grades. In Thomas E. Scruggs & M. A. Mastropieri (Eds.), *Literacy and Learning (Advances in Learning and Behavioral Disabilities, Volume 23)* (pp. 1-28). Emerald Group Publishing Limited.

Al Otaiba, S., Connor, C. M., Folsom, J. S., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. K. (2014). To wait in tier 1 or intervene immediately a randomized experiment examining first-grade response to intervention in reading. *Exceptional Children, 81*, 11-27. doi:10.1177/0014402914532234

Al Otaiba, S., Folsom, J. S., Schatschneider, C., Wanzek, J., Greulich, L., Meadows, J., . . . Connor, C. M. (2011). Predicting first-grade reading performance from kindergarten response to tier 1 instruction. *Exceptional Children, 77*, 453-470.

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention. *Remedial and Special Education, 23*, 300-316. doi:10.1177/07419325020230050501

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? *Journal of Learning Disabilities, 39*, 414-431. doi:10.1177/00222194060390050401

Al Otaiba, S., Kim, Y., Wanzek, J., Petscher, Y., & Wagner, R. K. (2014). Long-term effects of first-grade multitier intervention. *Journal of Research on Educational Effectiveness, 7*(3), 250-267. doi:10.1080/19345747.2014.906692

Al Otaiba, S., Wagner, R. K., & Miller, B. (2014). "Waiting to fail" redux understanding inadequate response to intervention. *Learning Disability Quarterly, 37*(3), 129-133. doi:10.1177/0731948714525622

Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., . . . Lang, T. (2001). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Annals of Internal Medicine, 134*, 663-694.

Anderson, A., Wood, E., Piquette-Tomei, N., Savage, R. S., & Mueller, J. (2011). Evaluating the impacts of just-in-time instructional support for teachers introducing a web-based reading program for primary grade children. *Journal of Teaching and Teacher Education, 19*, 499-525.

Aouad, J., & Savage, R. (2009). The component structure of preliteracy skills: Further evidence for the simple view of reading. *Canadian Journal of School Psychology, 24*, 183-200. doi:10.1177/0829573509336280

Archer, K., Savage, R. S., 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. doi:10.1016/j.compedu.2014.06.001

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Barnes, R. (2005). Moving towards technology education: Factors that facilitated teachers' implementation of a technology curriculum. *Journal of Technology Education, 17*(1), 6-18.

Below, J. L., Skinner, C. H., Fearrington, J. Y., & Sorrell, C. A. (2010). Gender differences in early literacy: Analysis of kindergarten through fifth-grade dynamic indicators of basic early literacy skills probes. *School Psychology Review, 39*(2), 240-257.

Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum.

Bishop, M. J., & Santoro, L. E. (2006). Evaluating beginning reading software for at risk learners. *Psychology in the Schools, 43*, 57-70. doi:10.1002/pits.20129

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2014). Intensive reading remediation in grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology, 106*, 46-57. doi:10.1037/a0033663

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72*, 101-130. doi:10.3102/00346543072001101

Bloomberg, L. D., & Volpe, M. (2012). *Completing your qualitative dissertation: A roadmap from beginning to end* (2nd ed.). Los Angeles, CA: Sage Publications.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15. doi:10.3102/0013189X033008003

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper (Ed.), *APA handbook of research methods in psychology: Vol. 2. Research Designs* (pp. 57-71): Washington, D.C: American Psychological Association. doi:10.1037/13620-004

Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91*, 403-414. doi:10.1037/0022-0663.91.3.403

Byrne, B., Coventry, W. L., Olson, R. K., Samuelsson, S., Corley, R., Willcutt, E. G., . . . DeFries, J. C. (2009). Genetic and environmental influences on aspects of literacy and language in early childhood: Continuity and change from preschool to Grade 2. *Journal of Neurolinguistics, 22*, 219-236. doi:10.1016/j.jneuroling.2008.09.003

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *The Teachers College Record, 112*(3), 579-620.

Campbell, M. K., Elbourne, D. R., & Altman, D. G. (2004). CONSORT statement: Extension to cluster randomised trials. *British Medical Journal, 328*, 702-708. doi:10.1136/bmj.328.7441.702

Campuzano, L., Dynarski, M., Agodini, R., Rall, K., & Pendleton, A. (2009). Effectiveness of reading and mathematics software products: Findings from two student cohorts. *National Center for Education Evaluation and Regional Assistance*, 1-111.

Canadian Council on Learning (CCL). (2010). The Future of Literacy in Canada's Largest Cities, (pp. 1-4). http://www.ccl-cca.ca/ccl/Reports/ReadingFuture/index.html

Carlson, J. A. (2010). Avoiding traps in member checking. *The Qualitative Report, 15*, 1102-1113.

Catts, H. W., Nielsen, D. C., Bridges, M. S., Liu, Y. S., & Bontempo, D. E. (2015). Early Identification of Reading Disabilities Within an RTI Framework. *Journal of Learning Disabilities, 48*(3), 281-297. doi:10.1177/0022219413498115

Chambers, B., Abrami, P. C., Tucker, B., Slavin, R. E., Madden, N. A., Cheung, A., & Gifford, R. (2008). Computer-assisted tutoring in Success for All: Reading outcomes for first graders. *Journal of Research on Educational Effectiveness, 1*, 120-137. doi:10.1080/19345740801941357

Chen, C. H. (2008). Why do teachers not practice what they believe regarding technology integration? *The Journal of Educational Research, 102*, 65-75. doi:10.3200/JOER.102.1.65-75

Cheung, A. C. K., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review, 7*(3), 198-215. doi:10.1016/j.edurev.2012.05.002

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*, 947-967. doi:10.1016/S0742-051X(02)00053-7

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Vol. 2nd). Hillsdale, NJ: Lawrence Erlbaum.

Coleman, L. O., Gibson, P., Cotten, S. R., Howell-Moroney, M., & Stringer, K. (2015). Integrating computing across the curriculum: The impact of internal barriers and training intensity on computer integration in the elementary school classroom. *Journal of Educational Computing Research, 1*, 1-20. doi:10.1177/0735633115616645

Comaskey, E. M., Savage, R. S., & Abrami, P. C. (2009). A randomised efficacy study of Web-based synthetic and analytic programmes among disadvantaged urban Kindergarten children. *Journal of Research in Reading, 32*, 92-108. doi:10.1111/j.1467-9817.2008.01383.x

Coyne, M. D., Kame'enui, E. J., & Simmons, D. C. (2001). Prevention and intervention in beginning reading: Two complex systems. *Learning Disabilities Research & Practice, 16*(2), 62-73. doi:10.1111/0938-8982.00008

Coyne, M. D., Kame'enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin. *Journal of Learning Disabilities, 37*, 90-104. doi:10.1177/00222194040370020101

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Los Angeles: Sage.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage.

Cuban, L. (2001). *Oversold and underused: Computers in the classroom.* Cambridge, MA: Harvard University Press.

Cullen, T. A., & Greene, B. A. (2011). Preservice teachers' beliefs, attitudes, and motivation about technology integration. *Journal of Educational Computing Research, 45*(1), 29-47. doi:10.2190/EC.45.1.b

Cunningham, A. E. (2006). Accounting for children's orthographic learning while reading text: Do children self-teach? *Journal of Experimental Child Psychology, 95*, 56-77. doi:doi:10.1016/j.jecp.2006.03.008

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934-945. doi:10.1037/0012-1649.33.6.934

Davidson, M. R., Fields, M. K., & Yang, J. (2009). A randomized trial study of a preschool literacy curriculum: The importance of implementation. *Journal of Research on Educational Effectiveness, 2*(3), 177-208. doi:10.1080/19345740902770028

Deault, L., Savage, R. S., & Abrami, P. C. (2009). Inattention and response to the ABRACADABRA Web-based literacy intervention. *Journal of Research on Educational Effectiveness, 2*, 250-286. doi:10.1080/19345740902979371

Denton, C. A. (2012). Response to intervention for reading difficulties in the primary grades some answers and lingering questions. *Journal of Learning Disabilities, 45*(3), 232-243. doi:10.1177/0022219412442155

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*, 181-199. doi:10.3102/0013189X08331140

Di Stasio, M. R., Savage, R. S., & Abrami, P. C. (2012). A follow up study of the ABRACADABRA web based literacy intervention in Grade 1. *Journal of Research in Reading, 35*, 69-86. doi:10.1111/j.1467-9817.2010.01469.x

Diamond, K. E., & Powell, D. R. (2011). An iterative approach to the development of a professional development intervention for head start teachers. *Journal of Early Intervention, 33*, 75-93. doi:10.1177/1053815111400416

Ecalle, J., Magnan, A., & Calmus, C. (2009). Lasting effects on literacy skills with a computer-assisted learning using syllabic units in low-progress readers. *Computers & Education, 52*, 554-561. doi:10.1016/j.compedu.2008.10.010

Edwards, J. A., & Lampert, M. D. (Eds.). (1993). *Talking data: Transcription and coding in discourse research.* Hillsdale, NJ: Erlbaum.

Ehri, L., Nunes, S., Willows, D., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*, 250-287.

Ertmer, P. A. (1999). Addressing first-and second-order barriers to change: Strategies for technology integration. *Educational Technology Research and Development, 47*(4), 47-61.

Ertmer, P. A. (2005). Teacher pedagogical beliefs: The final frontier in our quest for technology integration? *Educational Technology Research and Development, 53*(4), 25-39. doi:10.1007/BF02504683

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education, 42*, 255–284.

Ertmer, P. A., Ottenbreit-Leftwich, A. T., Sadik, O., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers & Education, 59*(2), 423-435. doi:10.1016/j.compedu.2012.02.001

Farstrup, A. E. (2007). RTI: A vital concern for reading professionals. *Reading Today, 25*, 17.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage Publications Inc.

Fien, H., Smith, J. L. M., Smolkowski, K., Baker, S. K., Nelson, N. J., & Chaparro, E. (2015). An examination of the efficacy of a multitiered intervention on early reading outcomes for first grade students at risk for reading difficulties. *Journal of Learning Disabilities, 48*, 602-621. doi:10.1177/0022219414521664

Fletcher, J. M., & Wagner, R. K. (2014). Accumulating knowledge: When Are reading intervention results meaningful? *Journal of Research on Educational Effectiveness, 7*, 294-299. doi:10.1080/19345747.2014.925310

Foorman, B. R., & Wanzek, J. (2016). Classroom reading instruction for all students. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The sciene and practice of multi-tiered systems of support* (2nd ed., pp. 235-252). New York: Springer.

Foster, W. A., & Miller, M. (2007). Development of the literacy achievement gap: A longitudinal study of kindergarten through third grade. *Language, Speech, and Hearing Services in Schools, 38*, 173-181. doi:10.1044/0161-1461(2007/018)

Frank, K. A., Zhao, Y., Penuel, W. R., Ellefson, N., & Porter, S. (2011). Focus, fiddle, and friends: Experiences that transform knowledge for the implementation of innovations. *Sociology of Education, 84*, 137-156. doi:10.1177/0038040711401812

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*, 93-99. doi:10.1598/RRQ.41.1.4

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention a decade later. *Journal of Learning Disabilities, 45*(3), 195-203. doi:10.1177/0022219412442150

Fugate, M. H. (2003). Review of the group reading assessment and diagnostic evaluation. In B. S. Plak, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook.* Lincoln, Nebraska: The Buros Institute of Mental Measurements.

Fullan, M. (1982). *The meaning of educational change*. Toronto: Ontario Insitute for Studies in Education (OISE) Press.

Fullan, M. (2005). The meaning of educational change: A quarter of a century of learning. In A. Lieberman (Ed.), *The roots of educational change* (pp. 202-216). Netherlands: Springer.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Boston: Pearson.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*, 915-945. doi:10.3102/00028312038004915

Gibbs, C. (2003). Explaining effective teaching: self-efficacy and thought control of action. *Journal of Educational Enquiry, 4*(2), 1-14.

Girvan, C., Conneely, C., & Tangney, B. (2016). Extending experiential learning in teacher professional development. *Teaching and Teacher Education, 58*, 129-139. doi:http://dx.doi.org/10.1016/j.tate.2016.04.009

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *RASE: Remedial & Special Education, 7*, 6-10. doi:10.1177/074193258600700104

Granger, C. A., Morbey, M. L., Lotherington, H., Owston, R. D., & Wideman, H. H. (2002). Factors contributing to teachers' successful implementation of IT. *Journal of Computer Assisted Learning, 18*(4), 480-488. doi:10.1046/j.0266-4909.2002.00259.doc.x

Graves, E., & McConnell, T. (2014). *Response to Intervention: Providing Reading Intervention to Low Income and Minority Students.* Paper presented at the The Educational Forum.

Greenwood, C. R., Carta, J. J., Goldstein, H., Kaminski, R. A., McConnell, S. R., & Atwater, J. (2014). The center for response to intervention in early childhood: Developing evidence-based tools for a multi-tier approach to preschool language and early literacy instruction. *Journal of Early Intervention, 36*(4), 246-262. doi:10.1177/1053815115581209

Gresham, F. M. (2007). *Evolution of the response-to-intervention concept: Empirical foundations and recent developments*. New York: Springer.

Griffiths, Y., & Stuart, M. (2013). Reviewing evidence- based practice for pupils with dyslexia and literacy difficulties. *Journal of Research in Reading, 36*(1), 96-116. doi:10.1111/j.1467-9817.2011.01495.x

Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher, 15*, 5-12. doi:10.3102/0013189X015005005

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership, 59*(6), 45-51.

Hagans, K. S., & Good III, R. H. (2013). Decreasing reading differences in children from disadvantaged backgrounds: The effects of an early literacy intervention. *Contemporary School Psychology: Formerly" The California School Psychologist", 17*(1), 103-117.

Hart, B., & Risley, T. R. (1999). *The social world of children: Learning to talk*. Baltimore: P. H. Brookes.

Hatcher, P. J., Hulme, C., Miles, J. N. V., Carroll, J. M., Hatcher, J., Gibbs, S., . . . Snowling, M. J. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: A randomised controlled trial. *Journal of Child Psychology and Psychiatry, 47*, 820–827. doi:10.1111/j.1469-7610.2005.01559.x

Haynes, L., Service, O., Goldacre, B., & Torgerson, D. (2012). *Test, learn, adapt: Developing public policy with randomised control trials*. UK: Cabinet Office - Behavioural Insights

Team. Retrieved from https://www.gov.uk/government/publications/test-learn-adapt-developing-public-policy-with-randomised-controlled-trials.

Hermans, R., Tondeur, J., Van Braak, J., & Valcke, M. (2008). The impact of primary school teachers' educational beliefs on the classroom use of computers. *Computers & Education, 51*, 1499-1509. doi:10.1016/j.compedu.2008.02.001

Hew, K. F., & Brush, T. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational Technology Research and Development, 55*, 223-252. doi:10.1007/s11423-006-9022-5

Hill, D. R., King, S. A., Lemons, C. J., & Partanen, J. N. (2012). Fidelity of implementation and instructional alignment in response to intervention research. *Learning Disabilities Research & Practice, 27*(3), 116-124. doi:10.1111/j.1540-5826.2012.00357.x

Hipps, G., Abrami, P. C., Savage, R. S., Cerna, N., & Jorgensen, A. (2005). ABRACADARA: The research, design and development of web-based early literacy software. In S. Pierre (Ed.), *DIVA. Innovations et tendances en technologies de formation et d'apprentissage*. Montreal: Presses Internationales Polytechnique.

Hoekstra, A., & Korthagen, F. (2011). Teacher learning in a context of educational change: Informal learning versus systematically supported. *Journal of Teacher Education, 62*, 76-92. doi:10.1177/0022487110382917

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

Holmes, M. (1998). Change and tradition in education: The loss of community. In A. Hargreaves, M. Lieberman, M. Fullan, & D. Hopkins (Eds.), *International handbook of educational change* (pp. 242-260). Dordrecht, Netherlands: Kluwer.

Howell, D. C. (2013). *Statistical methods for physchology* (8th ed.). Belmont, CA: Wadsworth, Cengage Learning.

Hughes, J. (2005). The role of teacher knowledge and learning experiences in forming technology-integrated pedagogy. *Journal of Technology and Teacher Education, 13*(2), 277-302.

Hurford, D. P. (1990). Training phonemic segmentation ability with a phonemic discrimination intervention in second-and third-grade children with reading disabilities. *Journal of Learning Disabilities, 23*, 564-569. doi:10.1177/002221949002300906

Hurry, J., & Sylva, K. (2007). Long-term outcomes of early reading intervention. *Journal of Research in Reading, 30*, 227-248. doi:10.1111/j.1467-9817.2007.00338.x

Jadad, A. R., & Enkin, M. (2007). *Randomized controlled trials: questions, answers, and musings* (2nd ed.). Malden, MA: Blackwell.

Jamieson, D. G. (2006). Literacy in Canada. *Paediatrics & Child Health, 11*, 573-574.

Judson, E. (2006). How teachers integrate technology and their beliefs about learning: Is there a connection? *Journal of Technology and Teacher Education, 14*, 581-597.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Justice, L. M. (2006). Evidence-based practice, response to intervention, and the prevention of reading difficulties. *Language, Speech, and Hearing Services in Schools, 37*, 284-297. doi:10.1044/0161-1461(2006/033)

Kaasila, R., & Lauriala, A. (2010). Towards a collaborative, interactionist model of teacher change. *Teaching and Teacher Education, 26*, 854-862. doi:10.1016/j.tate.2009.10.023

Kamps, D., Abbott, M., Greenwood, C., Wills, H., Veerkamp, M., & Kaufman, J. (2008). Effects of small-group reading instruction and curriculum differences for students most at risk in kindergarten. *Journal of Learning Disabilities, 41*, 101-114. doi:10.1177/0022219407313412

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 0034654315626800. doi:10.3102/0034654315626800

Kim, W., Linan- Thompson, S., & Misquitta, R. (2012). Critical factors in reading comprehension instruction for students with learning disabilities: A research synthesis. *Learning Disabilities Research & Practice, 27*(2), 66-78.

Klingner, J. K., & Edwards, P. A. (2006). Cultural considerations with response to intervention models. *Reading Research Quarterly, 41*, 108-117. doi:10.1598/RRQ.41.1.6

Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and numeracy are more heritable than intelligence in primary school. *Psychological science, 24*(10), 2048-2056. doi:10.1177/0956797613486982

Kulik, J. A. (2003). *Effects of using instructional technology in elementary and secondary schools: What controlled evaluation studies say. SRI Project No. P10446.001,* Retrieved from Arlington, VA: http://downloads.kennisnet.nl/mediawijzer/Onderzoeken/Kulik_ITinK-12_Main_Report.pdf

Kurz, A., Elliott, S. N., & Roach, A. T. (2015). Addressing the missing instructional data problem using a teacher log to document tier 1 instruction. *Remedial and Special Education, 36*(6), 361-373. doi:10.1177/0741932514567365

Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing.* Thousand Oaks, CA: Sage Publications.

Lam, E. A., & McMaster, K. L. (2014). Predictors of responsiveness to early literacy intervention: A 10-year update. *Learning Disability Quarterly, 37*(3), 134-147. doi:10.1177/0731948714529772

Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment integrity: An essential, but often forgotten, component of school-based interventions. *Preventing School Failure, 48*(3), 36-43. doi:10.3200/PSFL.48.3.36-43

Lawless, K. A., & Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of Educational Research, 77*, 575-614. doi:10.3102/0034654307309921

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2015). *IBM SPSS for intermediate statistics: Use and interpretation* (5th ed.). New York: Routledge.

Linan-Thompson, S., Vaughn, S., Prater, K., & Cirino, P. T. (2006). The response to intervention of English language learners at risk for reading problems. *Journal of Learning Disabilities, 39*, 390-398. doi:10.1177/00222194060390050201

Locke, J. L., Hodgson, J., Macaruso, P., Roberts, J., Lambrecht-Smith, S., & Guttentag, C. (1997). The development of developmental dyslexia. In C. Hulme & M. J. Snowling (Eds.), *Dyslexia: Biology, cognition and intervention*. London: Whurr.

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

Low, P. B., & Siegel, L. S. (2005). A comparison of the cognitive processes underlying reading comprehension in native English and ESL speakers. *Written Language & Literacy, 8*, 207-231.

MacArthur, C. A., Ferretti, R. P., Okolo, C. M., & Cavalier, A. R. (2001). Technology applications for students with literacy problems: A critical review. *The Elementary School Journal, 101*(3), 273-301.

Macaruso, P., Hook, P. E., & McCabe, R. (2006). The efficacy of computer based supplementary phonics programs for advancing reading skills in at risk elementary students. *Journal of Research in Reading, 29*, 162-172. doi:10.1111/j.1467-9817.2006.00282.x

Maskit, D. (2011). Teachers' attitudes toward pedagogical changes during various stages of professional development. *Teaching and Teacher Education, 27*, 851-860. doi:10.1016/j.tate.2011.01.009

Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-assisted learning strategies for first-grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly, 33*, 62-94.

Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology, 101*(3), 689-704. doi:10.1037/a0014240

Matzen, N. J., & Edmunds, J. A. (2007). Technology as a catalyst for change: The role of professional development. *Journal of Research on Technology in Education, 39*, 417-430. doi:10.1080/15391523.2007.10782490

McArthur, G., Eve, P. M., Jones, K., Banales, E., Kohnen, S., Anandakumar, T., . . . Castles, A. (2012). Phonics training for English- speaking poor readers. *Cochrane Database of Systematic Reviews*, Issue 12. Art. No.: CD009115. doi:10.1002/14651858.CD009115.pub2.

McBride, J. R., Ysseldyke, J., Milone, M., & Stickney, E. (2010). Technical adequacy and cost benefit of four measures of early literacy. *Canadian Journal of School Psychology, 25*, 189-204. doi:10.1177/0829573510363796

McIntosh, K., MacKay, L. D., Andreou, T., Brown, J. A., Mathews, S., Gietz, C., & Bennett, J. L. (2011). Response to intervention in Canada: Definitions, the evidence base, and future directions. *Canadian Journal of School Psychology, 26*, 18-43. doi:10.1177/0829573511400857

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sources book* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Mueller, J., Wood, E., & Willoughby, T. (2008). The integration of computer technology in the classroom. In T. Willoughby & E. Wood (Eds.), *Children's Learning in a Digital World* (pp. 272-297). Oxford, UK: Blackwell Publishing.

Mueller, J., Wood, E., Willoughby, T., Ross, C., & Specht, J. (2008). Identifying discriminating variables between teachers who fully integrate computers and teachers with limited integration. *Computers & Education, 51*, 1523-1537. doi:10.1016/j.compedu.2008.02.003

Mumtaz, S. (2006). Factors affecting teachers' use of information and communications technology: A review of the literature. *Journal of Information Technology for Teacher Education, 9*, 319-342. doi:10.1080/14759390000200096

Musti-Rao, S., Cartledge, G., Bennett, J. G., & Council, M. (2015). Literacy instruction using technology with primary-age culturally and linguistically diverse learners. *Intervention in School and Clinic, 50*(4), 195-202. doi:10.1177/1053451214546404

National Early Literacy Panel. (2008). *Developing early literacy*. Washington, DC: National Institute for Literacy. Retrieved from lincs.ed.gov/publications/pdf/NELPReport09.pdf

National Institute of Child Health and Human Development (NICHD). (2006). *Report of the National Reading Panel: Teaching children to read: Findings and determinations of the National Reading Panel by topic areas.* Retrieved from http://www.nichd.nih.gov/publications/nrp/findings.cfm

National Reading Panel. (2000). *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.* Bethesda, MD: National Institute of Child Health and Human Development. Retrieved from https://www.nichd.nih.gov/publications/pubs/nrp/documents/report.pdf

Nelson, J. R., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta analytic review. *Learning Disabilities Research & Practice, 18*, 255-267. doi:10.1111/1540-5826.00080

Neumann, M. M., Hood, M., & Neumann, D. L. (2009). The scaffolding of emergent literacy environment: A case study. *Early Childhood Education Journal, 36*, 313-319. doi:10.1007/s10643-008-0291-y

O'Connor, R. E. (2000). Increasing the intensity of intervention in kindergarten and first grade. *Learning Disabilities Research & Practice, 15*, 43-54. doi:10.1207/SLDRP1501_5

O'Connor, R. E., & Klingner, J. (2010). Poor responders in RTI. *Theory Into Practice, 49*(4), 297-304. doi:10.1080/00405841.2 010.510758

Opfer, V. D., Pedder, D. G., & Lavicza, Z. (2011). The role of teachers' orientation to learning in professional development and change: A national study of teachers in England. *Teaching and Teacher Education, 27*(2), 443-453. doi:10.1016/j.tate.2010.09.014

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal, 44*, 921-958. doi:10.3102/0002831207308221

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual Differences in Reading Development A Review of 25 Years of Empirical Research on Matthew Effects in Reading. *Review of Educational Research, 84*, 203-244. doi:10.3102/0034654313509492

Piquette, N. A., Savage, R. S., & Abrami, P. C. (2014). A cluster randomized control field trial of the ABRACADABRA web-based reading technology: Replication and extension of basic findings. *Frontiers in Psychology, 5, Article 1413.* doi:10.3389/fpsyg.2014.01413

Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on head start teachers and children. *Journal of Educational Psychology, 102*, 299-312. doi:10.1037/a0017763

Pressley, M. (2002). *Reading instruction that works.* (2nd ed.). New York: Guilford Press.

Preston, A. I., Wood, C. L., & Stecker, P. M. (2016). Response to intervention: Where it came from and where it's going. *Preventing School Failure: Alternative Education for Children and Youth, 60*(3), 173-182. doi:10.1080/1045988X.2015.1065399

Pullen, P. C., Tuckwiller, E. D., Konold, T. R., Maynard, K. L., & Coyne, M. D. (2010). A tiered intervention model for early vocabulary instruction: The effects of tiered instruction for young students at risk for reading disability. *Learning Disabilities Research & Practice, 25*, 110-123. doi:10.1111/j.1540-5826.2010.00309.x

Pyle, A. (2011). Considering coherence: Teacher perceptions of the competing agendas of RTI and an existing special education model. *Exceptionality Education International, 21*(3), 66-81.

Rathvon, N. (2004). *Early reading assessment: A practicioner's handbook*. New York: Guildford Press.

Reinhart, A. L., Haring, S. H., Levin, J. R., Patall, E. A., & Robinson, D. H. (2013). Models of not-so-good behavior: Yet another way to squeeze causality and recommendations for practice out of correlational data. *Journal of Educational Psychology, 105*, 241-247. doi:10.1037/a0030368

Richards, C., Pavri, S., Golez, F., Canges, R., & Murphy, J. (2007). Response to intervention: Building the capacity of teachers to serve students with learning difficulties. *Issues in Teacher Education, 16*(2), 55-64.

Ricketts, J., Bishop, D. V., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading, 15*, 47-70. doi:10.1080/10888438.2011.536129

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school examining direct cognitive assessments and teacher ratings. *American Educational Research Journal, 48*(2), 268-302. doi:10.3102/0002831210372249

Rogers, E. M. (2003). *Diffusion of innovations*. New York: Free Press.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*, 479-530.

Ross, J. A., & Bruce, C. D. (2007). Teacher self-assessment: A mechanism for facilitating professional growth. *Teaching and Teacher Education, 23*, 146-159. doi:10.1016/j.tate.2006.04.035

Russell, M., Bebell, D., O'Dwyer, L., & O'Connor, K. (2003). Examining teacher technology use. *Journal of Teacher Education, 54*, 297-310. doi:10.1177/0022487103255985

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods, 15*(1), 85-109. doi:10.1177/1525822X02239569

Sammons, P., & Smees, R. (1998). Measuring pupil progress at key stage 1: Using baseline assessment to investigate value added. *School Leadership & Management, 18*, 389-407. doi:10.1080/13632439869574

Sandholtz, J. H., Ringstaff, C., & Dwyer, D. C. (1997). *Teaching with technology: Creating student-centred classrooms.* New York: Teachers College Press.

Savage, R. S. (2008). Evidence-informed teaching strategies for improved literacy skills. *National Strategy for Early Literacy (NSEL) Policy Research Papers*, 1-49.

Savage, R. S. (2012). Evidence-based reading interventions: Implementation issues for the 21st century. In B. Kelly & D. F. Perkins (Eds.), *Handbook of implementation science for psychology in Education* (pp. 277-297). New York: Cambridge University Press.

Savage, R. S., Abrami, P. C., Hipps, G., & Deault, L. (2009). A randomized controlled trial study of the ABRACADABRA reading intervention program in grade 1. *Journal of Educational Psychology, 101*, 590-604. doi:10.1037/a0014700

Savage, R. S., Abrami, P. C., Piquette, N. A., Wood, E., Sanghera-Sidhu, S., Deleveaux, G., & Burgos, G. (2013). A (Pan-Canadian) cluster randomised control effectiveness trial of the ABRACADABRA web-based literacy program. *Journal of Educational Psychology, 105*(2), 310-328. doi:10.1037/a0031025

Savage, R. S., Abrami, P. C., Piquette-Tomei, N., Wood, E., & Deleveaux, G. (2009). *ABRACADABRA: A study in the development, implementation and effectiveness of a web-*

based literacy resource. A research progress report submitted to the Canadian Council on Learning. Retrieved from http://www.ccl-cca.ca/pdfs/OtherReports/Abrami-ABRA-ReportE.pdf

Savage, R. S., Burgos, G., Wood, E., & Piquette, N. A. (2015). The simple view of reading as a framework for national literacy initiatives: A hierarchical model of public-level and classroom-level factors. *British Educational Research Journal (paper in press).* doi:10.1002/berj.3177

Savage, R. S., & Carless, S. (2004). Predicting curriculum and test performance at age 7 years from pupil background, baseline skills and phonological awareness at age 5. *British Journal of Educational Psychology, 74*(2), 155-171. doi:10.1348/000709904773839815

Savage, R. S., & Carless, S. (2008). The impact of early reading interventions delivered by classroom assistants on attainment at the end of year 2. *British Educational Research Journal, 34*, 363-385. doi:10.1080/01411920701609315

Savage, R. S., Carless, S., & Erten, O. (2009). The longer-term effects of reading interventions delivered by experienced teaching assistants. *Support for Learning, 24*, 95-100. doi:10.1111/j.1467-9604.2009.01405.x

Savage, R. S., Carless, S., & Ferraro, V. (2007). Predicting curriculum and test performance at age 11 years from pupil background, baseline skills and phonological awareness at age 5 years. *Journal of Child Psychology and Psychiatry, 48*, 732-739. doi:10.1111/j.1469-7610.2007.01746.x

Savage, R. S., & Cloutier, E. (2016). Early reading interventions: The state of the practice, and some new directions in building causal theoretical models. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (Paper accepted).

Savage, R. S., & Deault, L. (2010). Understanding and supporting children experiencing Dyslexia and ADHD: Constitutional and classroom influences. In K. Littleton, C. Wood, & J. K. Staarman (Eds.), *International Handbook of Psychology in Education* (pp. 569-608). Bingley, UK: Emerald Group Publishing.

Savage, R. S., Erten, O., Abrami, P. C., Hipps, G., Comaskey, E., & van Lierop, D. (2010). ABRACADABRA in the hands of teachers: The effectiveness of a web-based literacy intervention in grade 1 language arts programs. *Computers & Education, 55*, 911-922. doi:10.1016/j.compedu.2010.04.002

Savage, R. S., & Pompey, Y. (2008). What does the evidence really say about effective literacy teaching? [The development of literacy: Implications of urrent understanding for applied psychologists and educationalists]. *Education and Child Psychology, 25*(3), 17-26.

Scardamalia, M., & Bereiter, C. (1996). Adaptation and understanding: A case for new cultures of schooling. In S. Vosniadou, E. DeCorte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the design of technology-supported learning environments* (pp. 149-163). Mahwah, NJ: Erlbaum.

Scheltinga, F., van der Leij, A., & Struiksma, C. (2010). Predictors of response to intervention of word reading fluency in dutch. *Journal of Learning Disabilities, 43*(3), 212-228. doi:10.1177/0022219409345015

Seethaler, P. M., & Fuchs, L. S. (2005). A drop in the bucket: Randomized controlled trials testing reading and math interventions. *Learning Disabilities Research & Practice, 20*, 98-102. doi:10.1111/j.1540-5826.2005.00125.x

Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly, 30*, 958-996.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151-218. doi:10.1016/0010-0277(94)00645-2

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*, 95-129.

Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accomodating individual differences into a model of acquisition. *Issues in Education, 1*, 1-57.

Shelley, M. C., Yore, L. D., & Hand, B. (2009). Education research meets the 'Gold Standard': Evaluation, research methods, and statistics after No Child Left Behind. *Quality Research in Literacy and Science Education*, 3-15. doi:10.1007/978-1-4020-8427-0_1

Siegel, L. S. (2006). Perspectives on dyslexia. *Paediatrics & Child Health, 11*, 581-587.

Simmons, D. C., Coyne, M. D., Kwok, O., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities, 41*(2), 158-173. doi:10.1177/0022219407313587

Singleton, C., & Thomas, K. (1999). Computerised baseline assessment of literacy. *Journal of Research in Reading., 22*(1), 67-80.

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly, 43*, 290-322. doi:10.1598/RRQ.43.3.4

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research, 79*, 1391. doi:10.3102/0034654309341374

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review, 6*, 1-26. doi:10.1016/j.edurev.2010.07.002

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, D.C.: National Academy Press.

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development, 74*(2), 358-373.

Snowling, M. J., & Hulme, C. (2011). Evidence based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology, 81*, 1-23. doi:10.1111/j.2044-8279.2010.02014.x

Stage, S. A., Abbott, R. D., Jenkins, J. R., & Berninger, V. W. (2003). Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal -- word reading discrepancy: Failure to validate discrepancy method. *Journal of Learning Disabilities, 36*, 24-33. doi:10.1177/00222194030360010401

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

Statistics Canada (2011). *Visible minority population and population group reference guide, national household survey, 2011*. (Catalogue number 99-010-XWE2011009). Statistics Canada. Retrieved from http://www12.statcan.gc.ca/nhs-enm/2011/ref/guides/99-010-x/99-010-x2011009-eng.cfm.

Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Sáenz, L., Yen, L., . . . Compton, D. L. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational Evaluation and Policy Analysis, 30*, 368-388. doi:10.3102/0162373708322738

Stuart, M. (2004). Getting ready for reading: A follow up study of inner city second language learners at the end of Key Stage 1. *British Journal of Educational Psychology, 74*, 15-36. doi:10.1348/000709904322848806

Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal, 39*, 469-518. doi:10.3102/00028312039002469

Suggate, S. P. (2014). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, 1-20. doi:10.1177/0022219414528540

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks a meta-analysis. *Review of Educational Research*, 698-739. doi:10.3102/0034654314566989

Taylor, J., Roehrig, A. D., Hensler, B. S., Connor, C. M., & Schatschneider, C. (2010). Teacher quality moderates the genetic effects on early reading. *Science, 328*, 512-514. doi:10.1126/science.1186149

Thomas, M. K., Barab, S. A., & Tuzun, H. (2009). Developing critical implementations of technology-rich innovations: A cross-case study of the implementation of Quest Atlantis. *Journal of Educational Computing Research, 41*, 125-153. doi:10.2190/EC.41.2.a

Torgerson, C. J. (2003). *Systematic reviews*. London, UK: Continuum International Publishing Group.

Torgerson, C. J. (2007). The quality of systematic reviews of effectiveness in literacy learning in English: A "tertiary" review. *Journal of Research in Reading, 30*, 287-315. doi:10.1111/j.1467-9817.2006.00318.x

Torgerson, C. J., Brooks, G., & Hall, J. (2006). *A systematic review of the research literature on the use of phonics in the teaching of reading and spelling.* London, UK: Department for Education and Skills [Report No. 1844786595]. Retrieved from http://www.dfes.gov.uk/research/data/uploadfiles/RR711_.pdf.

Torgerson, C. J., & Torgerson, D. (2008). *Designing and running randomised trials in health and the social sciences.* New York: Palgrave Macmillan.

Torgerson, C. J., & Zhu, D. (2003). A systematic review and meta-analysis of the effectiveness of ICT on literacy learning in English *Research Evidence in Education Library*. London, UK: EPPI-Centre, Social Sciences Research Unit Institute of Education. pp. 5-16.

Torgesen, J. K. (1998). Catch them before they fall: Identification and assessment to prevent reading failure in young children. *American Educator, 22*, 32–39.

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice, 5*, 55-64.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities. *Journal of Learning Disabilities, 34*, 33-58. doi:10.1177/002221940103400104

Tran, L., Sanchez, T., Arellano, B., & Lee Swanson, H. (2011). A meta-analysis of the RTI literature for children at risk for reading disabilities. *Journal of Learning Disabilities, 44*, 283-295. doi:10.1177/0022219410378447

Van Daal, V., & Sandvik, J. M. (2013). The effects of multimedia on early literacy development of children at risk: A meta-analysis. In A. Shamir & O. Korat (Eds.), *Technology as a support for literacy achievements for children at risk* (pp. 73-119). Dordrecht (Netherlands): Springer.

Van der Kooy-Hofland, V. A., Bus, A. G., & Roskos, K. (2012). Effects of a brief but intensive remedial computer intervention in a sub-sample of kindergartners with early literacy delays. *Reading and Writing, 25*(7), 1479-1497. doi:10.1007/s11145-011-9328-5

Van der Kooy-Hofland, V. A., Van der Kooy, J., Bus, A. G., Van IJzendoorn, M. H., & Bonsel, G. J. (2012). Differntial susceptibility to early literacy intervention in children with mild perinatal adversities: Short- and long-term effects of a randomized control trial. *Journal of Educational Psychology, 104*(2), 337-349. doi:10.1037/a0026984

Vanderwood, M., & Nam, J. (2007). Response to intervention for English language learners: Current development and future directions. In S. R. Jimerson, M. K. Burns, & A. M.

VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 408-417). New York, NY: Springer.

Vaughn, S., & Swanson, E. A. (2015). Special education research advances knowledge in education. *Exceptional Children, 82*(1), 11-24. doi:10.1177/0014402915598781

Vaughn, S., Wanzek, J., Murray, C. S., Scammacca, N., Linan-Thompson, S., & Woodruff, A. L. (2009). Response to early reading intervention examining higher and lower responders. *Exceptional Children, 75*, 165-183.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*, 601-638.

Vellutino, F. R., Scanlon, D. M., Small, S. G., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities. *Journal of Learning Disabilities, 39*, 157-169. doi:10.1177/00222194060390020401

Vellutino, F. R., Scanlon, D. M., Zhang, H., & Schatschneider, C. (2008). Using response to kindergarten and first grade intervention to identify children at-risk for long-term reading difficulties. *Reading and Writing, 21*(4), 437-480. doi:10.1007/s11145-007-9098-2

Wagner, B. D., & French, L. (2010). Motivation, work satisfaction, and teacher change among early childhood teachers. *Journal of Research in Childhood Education, 24*, 152-171. doi:10.1080/02568541003635268

Wanzek, J., & Vaughn, S. (2008). Response to varying amounts of time in reading intervention for students with low response to intervention. *Journal of Learning Disabilities, 41*, 126-142. doi:10.1177/0022219407313426

Waterman, B. B. (2003). Review of the group reading assessment and diagnostic evaluation. In B. S. Plak, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook.* Lincoln, Nebraska: The Buros Institute of Mental Measurements.

Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education, 21*(7), 747-766. doi:10.1016/j.tate.2005.05.009

Whitworth, B. A., & Chiu, J. L. (2015). Professional development and teacher change: The missing leadership link. *Journal of Science Teacher Education, 26*(2), 121-137. doi:10.1007/s10972-014-9411-2

Williams, K. T. (2001). *Group reading assessment and diagnostic evaluation: Teacher's scoring & interpretive manual.* Circle Pines, MN: American Guidance Service.

Wolgemuth, J. R., Savage, R. S., Helmer, J., Bottrell, C., Emmett, S., Lea, T., . . . Abrami, P. C. (2011). Using computer-based instruction to improve indigenous early literacy in Northern Australia: A quasi-experimental study. *Australasian Journal of Educational Technology., 27*(4), 727-750.

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

Wood, E., Specht, J., Willoughby, T., & Mueller, J. (2008). Integrating Computer Technology in Early Childhood Education Environments: Issues Raised by Early Childhood Educators. *Alberta Journal of Educational Research, 54*, 210-226.

Wozney, L., Venkatesh, V., & Abrami, P. C. (2006). Implementing computer technologies: Teachers' perceptions and practices. *Journal of Technology and Teacher Education, 14*, 173-207.

Yin, R. K. (2016). *Qualitative research from start to finish* (2nd ed.). New York: Guilford Press.

Zhang, L., & Han, Z. (2008). A Teaching Attitude Adjusting Model for College Teachers Based on Attitude Theories. *International Education Studies, 1*, 64-68.

Appendices

List of Appendices:

Appendix A

Follow-up Principal Information Letter

McGill

McGill University. 3700 McTavish Street, Montreal, Quebec, H3A 1Y2
Faculty of Education, Department of Educational and Counselling Psychology
(514) 398-7162

Researcher: Sukhbinder 'Bindy' Sanghera-Sidhu, sukhbinder.sanghera@mail.mcgill.ca / (514) 778-7345
Supervisor: Dr. Robert Savage, robert.savage@mcgill.ca / (514) 398-3435

Title of Research: *Pan-Canadian ABRACADABRA RCT follow-up. What do we know three years later about students and teachers' response to intervention (RtI)?*

Spring 2012

Information Letter for Principals

Dear Principal,

Background: During the academic years of 2007-2009 your school participated in a collaborative project involving the ABRACADABRA (ABRA) reading software program. The research for this pan-Canadian study was conducted by researchers in Alberta (Dr. Noella Piquette), Ontario (Dr. Eileen Wood) and Quebec (Dr. Robert Savage, and the Center for the Study of Learning and Performance (CSLP)). This study, along with previous studies examining the effectiveness of ABRA, have provided us with great insights into how the integration of computer-based innovations into existing Language Arts programs can assist in improving student learning outcomes in beginning reading skills. Our explorations have also given birth to new questions with regard to student-level and teacher-level factors that may best facilitate the long-term effectiveness of early literacy interventions such as the ABRA program.

What We Are Asking From You: For my doctoral research project I am hoping to follow-up with the original participating teachers and students from your school as they were an integral part of the success of our pan-Canadian study. Therefore, I am writing to you to ask for the involvement of your school in this follow-up study that will examine the long-term effects on the participants who were involved in the original 2007-2009 ABRA intervention study.

We will respect the demanding time frame of your school staff, by carrying out all research activities according to your and your staff's decisions and wishes. Our aim will be to conduct our research activities as quick and efficient as possible. Further, we will work around your teachers' classroom schedules in order to respect their routines and minimize any interruptions that may be caused by our presence in the class and school.

Purpose & Value of Present Research Study: One key objective of this proposed study is to examine if the short-term reading gains obtained by students, who received the ABRA instruction, were maintained three years later. This will add to existing literature by exploring interactions between ABRA and children's initial reading levels in order to better inform us about potential factors influencing the effectiveness (fit) between child and intervention, and consequently their response-to-intervention.

Another objective will be to examine the long-term response of teachers who have been involved in a whole classroom level intervention study. Specifically, with teachers from the original ABRA study, we want to examine if they are currently using ABRA, in what capacity, and what factors lead to their present use (or lack of use) of the ABRA program. This examination will provide us with greater insight on potential mediating factors that may facilitate a teacher's long-term use of an innovation deemed to be effective. Such information is of value as it assists future educational researchers and practitioners in improving their delivery of professional development training sessions. Furthermore, effective professional development opportunities facilitate a greater likelihood of teachers adopting innovative practices that may enhance their existing teaching practice and improve students' learning outcomes as a result.

Confidentiality: All information collected for this study will be kept completely confidential and anonymous. Data may be used for publication purposes in academic journals and will be presented in an anonymous form at all times, in order to ensure the confidentiality and anonymity of the data. General information about what is learned from the project can be shared with interested schools that wish to use this data in support of ongoing professional practice.

Voluntary Participation: Your participation in the study is completely voluntary, and you are under no obligation to agree in allowing our research team to collect follow-up data from your staff and students in your school. A decision not to volunteer will not influence the nature of the ongoing relationship the participant has with the researcher and/or the nature of their relationship with McGill University either now, or in the future.

Withdrawal From the Study: Any of the participants at your school can withdraw from the study at any time, or decline to participate in the data collection process of this follow-up study for any reason. The decision to terminate participation on any grounds will not affect any relationships with the researcher or McGill University. If you decide to withdraw your school from the study, all data generated as a consequence of the participation of your school shall be destroyed immediately.

Questions About the Research: I would be happy to discuss the details of the proposal and timeframe of the study in more depth. I will be in contact within the next week to set up a meeting, if you so desire, to discuss specific aspects of the project. If you have any questions about my research project in general or about your teachers and students' role in this study, please contact me, or my supervisor, as we would be happy to answer any questions you may have about this project:

Sukhbinder 'Bindy' Sanghera, sukhbinder.sanghera@mail.mcgill.ca / (514) 778-7345
Supervisor: Dr. Robert Savage, robert.savage@mcgill.ca / (514) 398-3435

Further, should you have any ethical concerns regarding this research project, 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,

If you wish your school to part of this follow-up study please carefully read, answer and sign below:

I) Do you freely provide us with permission to conduct this follow-up study in your school? *(Please select either YES or NO)* **YES or NO**

Name of School (please print) ____

Name of Principal (please print) ____

Signature_______________________________________ **Date**_____________________

We hope that your school will be part of this valuable ongoing research. Looking forward to hearing from you.

Appendix B

Follow-up Parent Consent Form

McGill
McGill University. 3700 McTavish Street, Montreal, Quebec, H3A 1Y2
Faculty of Education, Department of Educational and Counselling Psychology
(514) 398-7162

Researcher: Sukhbinder 'Bindy' Sanghera, sukhbinder.sanghera@mail.mcgill.ca / (514) 778-7345
Supervisor: Dr. Robert Savage, robert.savage@mcgill.ca / (514) 398-3435

Spring 2012

Parent Informed Consent Form

Dear Parents:

I am writing to tell you about a research project that our research team at McGill University is doing together with your child's school board. You may remember that in 2007-2009 you gave permission for your child to take part in the ABRACADABRA (ABRA) reading study. The purpose of the 2007-2009 study was to look at how teachers could use the ABRA program in their reading lessons with your children. The findings from that study showed us that teachers could effectively use a computer-based reading program as part of their literacy program to improve students' early reading skills.

The reason of this letter is to tell you about our follow-up project on the 2007-2009 ABRA study. The purpose of this study is to find out if the early exposure to the ABRA reading program had any positive long-term effects on students' reading skills. We would like to see how well the children who participated in the original ABRA study are reading three to four years after being involved in the ABRA study. We are asking for your permission to collect information on your child's current reading skills. Participating children will be asked to complete reading activities that will be completed over two separate 45 minutes sessions. Arrangements will be made to administer the reading activities with your child's classroom teacher. All efforts will be made to limit the amount of disruption to your child's regular classroom schedule. Upon request, a summary of your child's results on the reading activities can also be provided to you.

All information that is collected in this study is confidential. Research presentations and publications about this study will not contain any personal or identifying information about the children who participate. All data will be kept in locked offices at McGill University.

You are free to refuse permission for your child at any point in the project. Further, your child's participation is completely voluntary and she or he has the right to withdraw from the study at any time without giving any reason. The decision not to participate will not affect the relationships you or you child have with your child's teacher and school, the researcher or McGill University.

If you consent, we are asking you to also fill out the attached parent questionnaire. We are asking for information about your child's language background, reading experiences and general demographic background in order to better understand the experiences, needs and background of all of the students in the class. However, it is important for you to know that we will not communicate information from this questionnaire to the school without your permission. To keep the information confidential please return the questionnaires back in the sealed envelope provided.

However, with your permission, we would like to share your child's reading assessment scores with your child's teacher to help in learning and planning of your child's current reading program. We also would like permission to ask your child's teacher general questions about your child's current reading level and if she or he is currently receiving any additional literacy support.

If you would like your child to participate in this follow-up project on the effectiveness of ABRA on students' long-term reading skills, please read, complete and sign this consent form:

1. In order to aid in the learning and planning of future reading lessons, do you give consent for the researcher to share your child's results on the reading activities with your child's teacher?
   (Please select either YES or NO): ***Yes______ or No _____***

2. Do you give consent for your child's classroom teacher to share information regarding your child's current reading skills ability, any reading difficulties he or she may have and any additional support he or she may be receiving?
   (Please select either YES or NO): ***Yes______ or No _____***

3. Do you give consent to share information from the parent questionnaire with your child's teacher? *(Please select either YES or NO):* ***Yes______ or No _____***

Parent's Name (please print) ____

Signature_______________________________________ Date__________________

Child's name (please print)____

Child's birth date (month, date, year)____

School ____ Grade _______________

Current Teacher (please print) ____

Teacher who taught ABRA to your child 3-4 years ago _________________________

If you have any concerns about your child's rights or welfare as a participant in this research project, please feel free to contact Lynda McNeil at the McGill Research Ethics Office by phone (514) 398-6831 or by email at lynda.mcneil@mcgill.ca.

Further, if you have any questions about your child's role in the study, please contact me, or my supervisor, as we would be happy to answer any questions you may have about this project:
Sukhbinder 'Bindy' Sanghera, sukhbinder.sanghera@mail.mcgill.ca / (514) 778-7345
Supervisor: Dr. Robert Savage, robert.savage@mcgill.ca / (514) 398-3435

We hope that you will be part of this valuable research. Please return the signed letter and completed questionnaire back to your child's classroom teacher.

Appendix C

ABRA Study Follow-up Parent Questionnaire

This questionnaire usually takes less than 5 minutes! It is an important part of the ABRACADABRA research program that your child is taking part in, and your time to complete this is much appreciated. This information will better help us understand how children' reading skills develop. You are of course entirely free to not complete this questionnaire.

Your Full Name: ______________________________

Your Child's Full Name: ______________________________

Your Child's Birth Date: ______________________________

Please circle either **YES** or **NO**.

1) Has your child ever missed school for significant periods of time? YES / NO

How long were they absent in total? Number of weeks________

2) Does your child have normal (or corrected to normal) hearing? YES / NO

Give details here:

____

____

3) Does your child have normal (or corrected to normal) vision? YES / NO

Give details here:

____

____

4) Has your child been diagnosed with a problem in motor skills? YES / NO

If YES, when was it first identified?

____

____

5) Has your child been diagnosed with any -learning problems? YES / NO

If YES, when was it first identified?

____

____

6) How often do you read to your child in:

English?	French?	Other?
☐ Everyday	☐ Everyday	☐ Everyday
☐ 2-3 times per week	☐ 2-3 times per week	☐ 2-3 times per week
☐ Once a week	☐ Once a week	☐ Once a week
☐ Once a month or less	☐ Once a month or less	☐ Once a month or less
☐ Never	☐ Never	☐ Never

Please specify language(s) read: ___________________________________

7) What is/are Mother's first language________________________________

8) What is/are Father's – first language ________________________________

9) What is the language used at home between:

a) mother and father__________________________ (write language(s) used here)

b) mother and child __________________________ (write language(s) used here)

c) father and child __________________________ (write language(s) used

10) Child's ethnic origin (select all that apply):

☐ White
☐ Aboriginal (North American Indian, Métis or Inuit (Eskimo))
☐ Chinese
☐ South Asian (e.g., East Indian, Pakistani, Sri Lankan, etc.)
☐ Southeast Asian (e.g., Vietnamese, Cambodian, Malaysian, Laotian, etc.)
☐ Latin American
☐ Black
☐ Arab
☐ Korean
☐ Japanese
☐ West Asian (e.g., Iranian, Afghan, etc.)
☐ Filipino
☐ Other (i.e., if multiple ethnicities apply, specify)______________________________

11) Mother's education experiences

☐ Elementary school only
☐ Left school with graduation diploma
☐ College/CGEP
☐ University higher degree
☐ Did not receive school graduation diploma
☐ Technical training
☐ University Bachelor's degree

12) Father's education experiences

☐ Elementary school only
☐ Left school with graduation diploma
☐ College/CGEP
☐ University higher degree
☐ Did not receive school graduation diploma
☐ Technical training
☐ University Bachelor's degree

13) What is your postal code? ____

Thank you again for your time in completing this questionnaire.

Bindy Sanghera-Sidhu
PhD Candidate, Graduate Studies
Faculty of Education, McGill University

Dr. Robert Savage
Professor & Graduate Advisor
Faculty of Education, McGill University

Appendix D

Verbal Assent Script for Student Participants

Prior to administrating the group reading assessment measure (the *GRADE*) to the student participants I will deliver the following verbal assent script:

"Hello class, my name is Bindy Sanghera, and I am a student from McGill University. I am very happy to have the opportunity to spend some time with you this morning (or afternoon). I am collecting information for my research project, which is about the computer reading program ABRACADABRA and students reading skills. By raising your hand, who remembers working with the ABRA reading program about three years ago?"

Solicit a few responses from the group of students to see what they remember…and if no hands raised, provide them with a brief reminder of the ABRA program and show them a few illustrations of the main characters from the ABRA program.

Then continue by saying, "*Today, I'm here to work with some of the students who were involved in the original ABRA study. You are all here today with me because your parents said it was okay for you to complete a few reading activities for me which will help me with my research project. Only if your parents have provided permission, I will share your completed work on these reading activities with your classroom teacher. However, otherwise all of the work you do for me today will be private (confidential) and only used by me to get a better idea of how well students your age are reading.''*

"If you have any questions or problems with the reading activities you have been asked to complete please let me know. You are free to withdraw from these activities at any time without giving any reasons. Would you like to participate in my research project?" (Note, if the group is larger than 5 students I will ask students to show their assent by raising their hand).

Appendix E

Follow-up Teacher Informed Consent Form

McGill

McGill University. 3700 McTavish Street, Montreal, Quebec, H3A 1Y2
Faculty of Education, Department of Educational and Counselling Psychology
(514) 398-7162

Researcher: Sukhbinder 'Bindy' Sanghera, sukhbinder.sanghera@mail.mcgill.ca / (514) 778-7345
Supervisor: Dr. Robert Savage, robert.savage@mcgill.ca / (514) 398-3435

Title of Research: *Pan-Canadian ABRACADABRA RCT follow-up. What do we know three years later about students and teachers' response to intervention (RtI)?*

Spring, 2012

Teacher Informed Consent Form

Dear Teacher,

Background: During the academic years of 2007-2009 you participated in a collaborative project involving the ABRACADABRA (ABRA) reading software program. You may recall that your involvement was part of a pan-Canadian study that was conducted by researchers in Alberta (Dr. Noella Piquette), Ontario (Dr. Eileen Wood) and Quebec (Dr. Robert Savage, and the Center for the Study of Learning and Performance (CSLP)). Due to the involvement of teachers like you, this study provided us with great insights of how the integration of computer-based innovations into existing Language Arts programs have the potential to assist in improving the learning outcomes of students' beginning reading skills. Our explorations during the original pan-Canadian study have lead to new questions with regard to student-level and teacher-level factors that may best facilitate the long-term effectiveness of early literacy intervention such as the ABRA program. The new questions that arose from the original ABRA study are the reason for this current follow-up study. I am conducting this follow-up study as part of my doctoral research, and I hope you will consider being a part of.

Purpose & Value of Present Research Study: One key objective of this study is to examine the long-term response of teachers who have been involved in a whole classroom level intervention study. Specifically, with teachers from the original ABRA study, we want to examine if they are currently using ABRA, in what capacity, and what factors lead to their present use (or lack of use). This examination will provide us with greater insight on potential mediating factors that may facilitate a teacher's long-term use of an innovation deemed to be effective. Such information is of value as it assists future educational researchers and practitioners in improving their delivery of professional development training sessions, and in providing appropriate support during the implementation phase. The delivery of more effective professional development opportunities will in turn facilitate a greater likelihood that future teachers may adopt innovative practices that could enhance their existing teaching practice which could lead to improve students' learning outcomes.

What We Are Asking From You: Your participation in this follow-up study will be of great value in our exploration of how educational researchers can facilitate an environment most conducive for supporting teachers who have thoughtfully agreed to participate in research intervention studies.
If you choose to participate you, we are asking you to complete the attached ***follow-up questionnaire*** and return it in the preaddressed stamped envelope provided. Or if you rather, the option of completing an ***online version*** of the questionnaire is also available. The follow-up questionnaire will have you recall your experience of being part of the pan-Canadian study and ask you to provide honest feedback about how your involvement in the ABRA study may have influenced your current language arts program.

You could also consent to participating in a brief (15-25 minute) audiotaped interview with the principal researcher, Bindy Sanghera, in order to explore in depth the themes addressed in the follow-up questionnaire. The interviews will be audiotaped in order to accurately capture your experiences in a transcription of your discussion. You will have an opportunity to review the transcript once it is completed so that you can make any revisions you consider necessary. Your feedback will assist in gaining insight on how professional development opportunities can be best facilitated to support teachers learning. Further, if you are using the ABRA tool in any capacity in your current literacy program, we would like your consent to come observe how ABRA is being used in your class.

Confidentiality and Participation: Information collected will be kept confidential and anonymous. No identifiable data will be shared with your school; only summary data from all questionnaires and interviews will be reported. Data used for publication purposes will be presented in an anonymous form in order to ensure the confidentiality and anonymity of your involvement. Further, your participation is completely voluntary. You are under no obligation to agree to participate. You also have the right to withdraw from the study at any time without indicating any reason. A decision not to volunteer or to withdraw will not influence the nature of the ongoing relationship you have with the researcher and/or the nature of your relationship with McGill University either now, or in the future.

Questions About the Research: I would be happy to discuss the details of the proposal and timeframe of the study in more depth. If you have any questions about my research project in general or about your role in this study, please contact me, or my supervisor, as we would be happy to answer any questions you may have about this follow-up project:

Sukhbinder 'Bindy' Sanghera, sukhbinder.sanghera@mail.mcgill.ca / (514) 778-7345
Supervisor: Dr. Robert Savage, robert.savage@mcgill.ca / (514) 398-3435

Further, should you have any ethical concerns regarding this research project, 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.

If you wish to be part of this follow-up study please carefully read, answer and sign below:

I) Do you freely consent and agree to participate in the following research activities of this study?

(Please select either YES or NO)

a) To complete the written follow-up questionnaire. **Yes_______ or No ______**

b) To complete the **online** follow-up questionnaire. **Yes_______ or No ______**

* If yes to (b) please provide your email address: ______________________________________

c) To voluntarily take part in an audiotaped interview to discuss your experiences of being part of the 2007-2009 ABRA intervention study? **Yes_______ or No ______**

d) To provide permission for an observation of your language arts lesson when ABRA is being used.

Yes_____ No_______ or Not Applicable, as I do not use ABRA _______

Grade level taught (please print) ____

Name of Teacher (please print) ____

Signature____ Date_________________________

We hope that you will continue to be part of this valuable ongoing research. Looking forward to hearing from you.

Appendix F

Teacher Follow-up Questionnaire

McGill University. 3700 McTavish Street, Montreal, Quebec, H3A 1Y2
Faculty of Education, Department of Educational and Counselling Psychology
(514) 398-7162

Researcher: Sukhbinder 'Bindy' Sanghera-Sidhu, sukhbinder.sanghera@mail.mcgill.ca / (514) 778-7345
Supervisor: Dr. Robert Savage, robert.savage@mcgill.ca / (514) 398-3435

Title of Research: *Pan-Canadian ABRACADABRA RCT follow-up. What do we know three years later about students and teachers' response to intervention (RtI)?*

Teacher's Name: ______________________________

A) GENERAL BACKGROUND

1. Age: _____ **2. Sex:** ☐ Male ☐ Female

3. What is your ethnic origin (select all that apply):

☐ White
☐ Aboriginal (North American Indian, Métis or Inuit (Eskimo))
☐ Chinese
☐ South Asian (e.g., East Indian, Pakistani, Sri Lankan, etc.)
☐ Southeast Asian (e.g., Vietnamese, Cambodian, Malaysian, Laotian, etc.)
☐ Latin American
☐ Black
☐ Arab
☐ Korean
☐ Japanese
☐ West Asian (e.g., Iranian, Afghan, etc.)
☐ Filipino
☐ Other (i.e., if multiple ethnicities apply, specify)____________________________

4. How many years have you been teaching: _____

5. How many years have you been teaching early literacy: _____

B) ABRA BACKGROUND

1. Name of School during 2007-2009 ABRA study:____________________________

2. Did you deliver the ABRA intervention _______ or the non-ABRA lessons ________

3A) Indicate year(s) you formally used ABRA and /or took part in an ABRA study as the non-ABRA comparison group teacher:

_____ 2005-2006 _____ 2006-2007 _____ 2007-2008 _____ 2008-2009

3B) What Grades did you teach during these years (Kindergarten 1 or 2 or Grade 1-2 split)

_____ Kindergarten _____Grade 1 _____ Grade 2 _____ Grade 1-2 split

3C) Which year(s) did you attend a training session on the ABRA program

_____ 2005-2006 _____ 2006-2007 _____ 2007-2008 _____ 2008-2009

C) ABRA TRAINING and SUPPORT

For the 2007-2009 ABRA study our research team tried to provide the following for all of the intervention teachers using the ABRA tool.

Training and Support BEFORE the Intervention Phase	**Training and Support DURING the Intervention Phase**

• Full day of training
• Background training and hands-on time with the ABRA software program
• Provided ABRA resources on planning balanced literacy lessons:
  ○ ***"Teacher's Zone"*** an online resource
  ○ Hardcopy of a ***"ABRA Teacher's Manual"***
• Small group collaboration time to plan introductory ABRA lessons
	• ***In-class support*** for the initial 3-5 ABRA lessons
	• Opportunities to discuss use of ABRA with ABRA team members during ***classroom observation visits***
	• Opportunity to take part in a ***focus group*** feedback session (mid-way through intervention phase)
	• Opportunity to participate in ***an additional hands-on training session*** (mid-way through intervention phase)

1) With reference to the above, was your training in any way different (for example, length of in-class support provided, feedback given, involvement in additional training sessions or participation in focus group sessions, etc.). Discuss.

____

____

____

____

2) During your involvement in the ABRA study, can you describe what key factors over the course of the training and support phases (as described in C above) influenced your use of ABRA in your classroom?

____

____

____

____

3) Can you describe what key factors over the course of the training and support phases influenced your CURRENT professional use of the ABRA software?

____
____
____
____

4) What were your attitudes about using information and computer technology (ICT) in the classroom BEFORE your involvement in the ABRACADABRA literacy study? For example, did you think technology was appropriate? Discuss.

____
____
____
____

5) Has involvement in the ABRA study changed your attitudes and beliefs toward using computer technology in your language arts (LA) lessons with your students? Explain.

____
____
____
____

6) What role did the training, the ongoing support, and your observations about student learning play in your attitudes towards implementing ABRA into your LA program? Explain.

____
____
____
____

7a) Using the 5-point scale provided, **please rate** how important the following factors are in positively influencing your decision to incorporate a new teaching strategy or innovation into your teaching practice.

The new teaching strategy or innovation…

		Very Important	Moderately Important	Important	Not Very Important	Not Important at All
		4	3	2	1	0
1.	Engages the students interest					
2.	Has empirical research support as an effective strategy					
3.	Meets the provincial curriculum objectives					
4.	Can be used flexibly with your existing practice					
5.	Has been shown to have positive effects on student learning outcomes					
6.	Has been used and praised by your teaching colleagues for being an effective teaching tool					
7.	Fits well with your existing pedagogical beliefs of effective teaching strategies					
8.	Has generated positive feedback from students and/or their parents					
9.	The use of the new innovation leads to an increased satisfaction with respect to your teaching					
10.	Was adaptable and simple to integrate into your existing practice					
11.	Corresponds with your existing pedagogical knowledge, belief and value system					
12.	Was well supported (i.e., through training and/or ongoing consultation support, etc.) making implementation easier					
13.	Is being used by a number of your colleagues with whom you can share your experiences of implementing the new strategy					
14.	Has been promoted by the school administration as a valuable tool to integrate into your practice					
15.	Has been promoted by the school board as a district-wide initiative					

7b) From the 15 factors above, can you isolate the single most important factor that would influence you to integrate a new teaching strategy into your existing practice.

D) PRESENT USE of ABRA

1) Are you currently using ABRA?

_____No

_____Yes

2) If not, could you explain what factors have influenced your decision NOT to use the ABRA program? Explain.

____

____

3) If yes, what factors have influenced your decision to continue using ABRA to support your current literacy program? Explain.

____

____

4) Reflecting back to the professional development you received to be part of the 2007-2009 ABRA literacy intervention study, can you identify factors which have influenced your PRESENT professional use (or lack of use) of ABRA? Explain.

____

____

5) How often do your students work with ABRA (circle the closest response)?

A) Not at all
B) Infrequently (an odd lesson once in a while)
C) Occasionally (less than 30 minutes / week)
D) Regularly (30 – 60 minutes / week)
E) Frequently (1 -2 hours /week)

6) Could you explain the ways in which you are currently using the ABRA program?

____

____

7) Could you explain in what ways the ABRA program fit or did not fit with your teaching philosophy and existing LA program:

____

____

8) When you have used ABRA professionally, did you notice that some students more than others appeared to benefit more from its use? If so, could you describe the students who were successfully using ABRA? And which activities work best?

____

____

9a) Please indicate if any of the following factors may have **NEGATIVELY** influenced your current use of the ABRA program.

		Significant negative	Some negative influence	Very little negative	No influence	Not Applicable
		4	3	2	1	0
1.	Students interest in using the program					
2.	Access to computers in school					
3.	Reliable internet access					
4.	Access to technical support					
5.	Personal comfort level with the program					
6.	Adaptability of the program to be integrated into your existing LA practice					
7.	Observed effects on student learning outcomes					
8.	Personal belief of the appropriateness of using computers as a tool in early literacy programs					
9.	Other colleagues comments concerning effectiveness of the program					
10.	Fit with your existing pedagogical beliefs of effective teaching strategies					
11.	Feedback from students and/or their parents					
12.	Existing curriculum demands					
13.	School administration's support of the program					
14.	School district's support of the program					

9b) If other factors may have played a role in negatively influencing your use of the ABRA program please discuss:

____

____

10) Overall how would you describe the fit between the ABRA program with your existing LA curriculum? Please CIRCLE the option that fits best:

Not a good fit at all	A poor fit	Moderately a good fit	A good fit	A very good fit
0	1	2	3	4

Thank you for taking the time to complete this questionnaire. Please place completed questionnaire in the pre-addressed and pre-stamped envelope provided. Seal it and return it back to the principal investigator of this study, Bindy Sanghera-Sidhu.

Appendix G

Teacher Interview Guide & Additional Interview Questions

The following is a draft of a verbal guide I will use with teachers who have consented to be interviewed.

Thank-you again for agreeing to be interviewed for this follow-up study that is examining teachers' short and long-term experiences of being involved with the 2007-2009 ABRACADABRA intervention study. As mentioned in the consent form you signed, in order to accurately capture your responses I would like to audiotape our interview today. I will only use the recordings in order to transcribe our discussion. You will have a chance to see the transcript once it is completed so that you can review and make revisions to it if you wish. Once all of the teacher interviews have been completed, I use the reviewed transcripts to draw out themes of teachers' experiences of being part of the original ABRA intervention study and their current use of technology in their classrooms.

Once again all the information collected for this study will be kept completely confidential and anonymous, so your responses will not be traced back to you. Identifying information will not be used; rather, as mentioned, themes from our discussion will only be reported on. You may also stop the interview at any point and are free to withdraw from the study at any time, without giving any reasons, and with no negative consequences. Now, understanding all of this, are you comfortable to begin our interview…

In addition to reviewing their responses to the teacher questionnaires, I have included an example of some of the general types of questions I may be asking teacher participants to respond to, and which will guide our interview discussion:

CURRENT USE OF ABRA

1. Are you still teaching primary students, if so which grade?
2. And if so, are you currently using the intervention (ABRA) in language arts program?
3. If yes, how? If not, why?

REFLECTIVE QUESTIONS ABOUT THEIR EXPERIENCE IN THE ABRA STUDY

4. Going back a few years, what can you remember (if anything) about being involved in the ABRA study?
5. What do you remember about the training on the ABRA program (if anything)?
6. What do you remember about the flexibility of the ABRA program?
7. Do you remember if you were able to customize ABRA lessons to support your existing LA curriculum? If yes, how?
8. Would you say that the ABRA activities were a 'good fit' with your existing Language Arts curriculum? If yes, could you explain how? If no, could you explain why?

VALUES AND BELIEFS ABOUT TECHNOLOGY (computers) IN CLASSROOMS

9. How do you feel about using technology / computers in the classroom? In what ways do you believe computers should be used in the classroom (if at all)?
10. Do your students use computers during class-time? If so, how would you describe how their time is spent on the computers? What is their reaction when they get to work on the computers?
11. Would you say that the environment of your school supports the use of technology? Could you explain?
12. What other factors can you think of that may influence your decision to use a new computer program with your students?

Appendix H

Teacher Change Models Vignettes & Model Creation Process Guidelines

The following vignettes were used to describe the two linear teacher models of interest. These vignettes were introduced after the teacher interview portion was completed. Upon completion of the interviews, the teachers were all asked if they would be willing to participate in one final activity, which would be the drawing of their own teacher change process models.

Model A Vignette – "*The first model describes the teacher change process as follows:* ***Model A*** *predicts that if you go to a ProD workshop and get good training and information on, lets say, a new technology-based innovation like the ABRA program (changes in your* ***external domain****), this new information may change your opinion on the value of technology in the classroom and how you think about delivering a particular language arts lesson (changes in your* ***personal domain****).* ***Model A*** *suggests that such changes in your attitudes and perceptions about the usefulness of technology and computer-based programs, like ABRA, and this change in beliefs/attitudes will lead to specific changes in your classroom behaviours and practices (i.e. the permanent implementation of the ABRA program into your existing literacy practice and the increase use of technology) (changes in your* ***domain of practice****) which, in turn, will lead to improved student learning (changes in your* ***domain of consequence****)*."

Model B Vignette – "*Whereas the next model,* ***Model B*** *predicts that you first go to a ProD workshop and learn about and get to play around on a new technology-based innovation like the ABRA program (changes in your* ***external domain****). But you have not yet really changed your beliefs and attitudes about the value of technology use in your classroom (changes in your* ***personal domain****). Rather you decide to first try it out or 'field-test' the program in your classroom for a short while (changes in your* ***domain of practice****). After a few days/weeks of trying it out you observe improved changes in student learning outcomes (changes in your* ***domain of consequence****), and then after personally seeing positive gains in your students outcomes this is when there are changes in your beliefs and attitudes around the value of using computer-based programs like ABRA in your literacy program (changes in your* ***personal domain****).*"

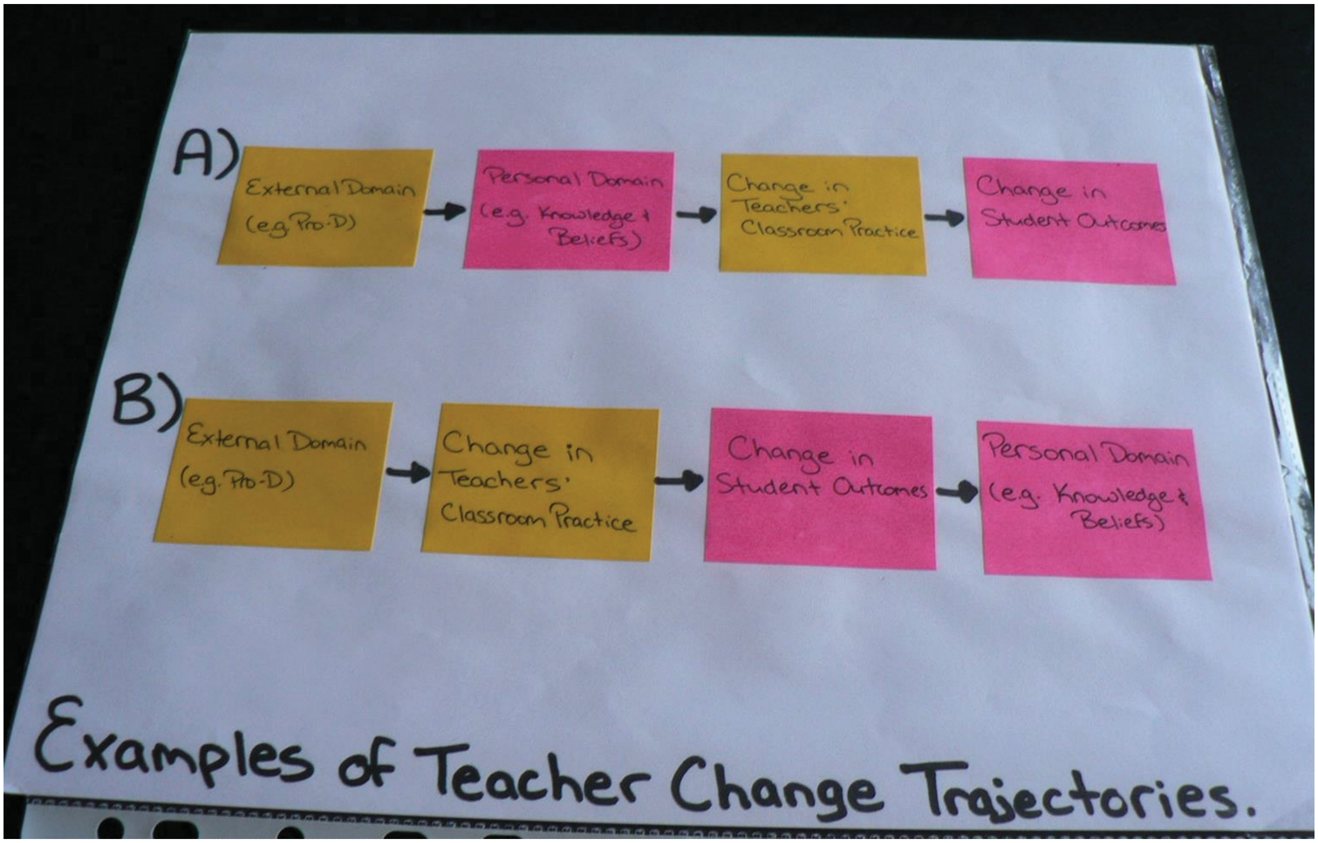
Selecting their Models from Linear Model A or Model B - After hearing the descriptions of the two teacher change models, teachers were asked to select which model they felt best represented their own process of deciding to change their practice to incorporate a new innovation into their literacy practice. To assist in this process, the interviewer provided a premade illustration of the two models of interest (See Appendix I)

Creating their own Process of Change Model - After selecting either Model A or B, the teachers were then asked to create their own personal model of their teacher change process. The interviewer stressed that even though she would like them to try to incorporate the four domains of change (which were pre-written out onto post-its) their personal process models did not have to resemble neither Model A or B.

Using the supplies provided by the interviewer (blank paper, different colored post-it notes, different colored pens, markers, highlighters and pencils) the teachers were encouraged to add any additional descriptions of variables they believe influence their decision making process of incorporating a new resource into their teaching practice. They were also encouraged to think about the direction (flow) of their change process, and then illustrate where their change process usually initiates from, and indicate by drawing in links between the domains to illustrate how their process of change moves from one domain to the next.

Appendix I

Illustration Examples of Teacher Change Model Trajectories



Appendix J

Pupil-level Variables of Interest Used in SPSS Analyses

DV or IV	Variable	SPSS Variable Name	Variable Description & Values
DV	Reading 2	BWBinary	Blending Words score percentile (binary) @ T2 0 = Nonresponders ≤ 30% on BW2 1 = Responders > 30% on BW2 *(Reference group)*
DV	Reading 3	AtRisk3Cpl	Total Comprehension percentile score (binary) @ T3 0 = At risk, ≤ 30% on Comprehension 1 = Not at risk >30% on Comprehension *(Reference group)*
IV	Reading 1	AtRisk1	At risk if bottom 30% on latent variable regression factor score for reading measures (LSK, FRY & BW). **Binary** @ T1. 0 = At risk (bottom 30%) 1 = Not at risk (top 70%) *(reference category)* ** Note: -.53 z-score is cutoff*
IV	LC1	LCrw1	Listening Comp (raw score @ T1)
IV	Sex	Gender	0 = Female 1 = Male (*entered as reference category*)
IV	Ethnicity	Ethnic	1 = White *(reference category)* 2 = Asian (includes Chinese, South Asian, Southeast Asian, Arab, Korean, Japanese, West Asian & Filipino) 3 = Other (includes Aboriginal, Latin American, Black & Other)
IV	SES	MomEd	0 = No postsecondary *(reference)*(includes, Elementary school only; Did not receive school graduation diploma & Left school with graduation diploma) 1 = Technical or College 2 = University
IV	Intervention	IFM_C	Intervention group (Combined) 0 = Control Group *(reference category)* 1 = Entry Level (Poor) 2 = Adoption/Adaption of intervention (Good)
IV	Reading 1 (NB)	ORegFac	**NOT** Binary Split @ T1. Are the calculated, ***continuous*** latent scores on baseline reading measure (LSK, FRY & BW)
IV	Ethnicity (minimize)	Ethnic2	1 = White *(reference category)* 2 = Other
IV	SES (minimize)	MomEd2	0 = No Post Secondary *(reference category)* 1 = Post Secondary
IV	Intervention (minimize)	IFM_CvsG	0 = Control *(reference category)* 1 = Good Implementation ** Note – Poor implementation removed completely from this analysis (loss of 10 participants at time of follow-up & at risk)*

Appendix K
ABRA Implementation Rubric
(used during ABRA 2007-2009 Study)

ABRA – IMPLEMENTATION RUBRIC

CONTROL (0)	ENTRY (1)	ADOPTION (2)	ADAPTATION (3)	DIFFERENTIATED ADAPTATION (4)
♦ No implementation ♦ No aspect of ABRA was used	♦ Little to No evidence of teacher planning of ABRA ELA lessons ♦ Little to No evidence of teacher instructional guidance ♦ Little to No evidence of teacher monitoring students' use of ABRA ♦ No evidence of teacher's awareness of zone of proximal development as students instructed to all work on same activity level ♦ Minimal student exposure to ABRA activities ♦ ABRA exposure – mainly unstructured lessons where students choose own activities (play-time /free-time) ♦ Occasional disruption and off-task behaviour of students ♦ Unclear about teacher and student's navigational comfort level with ABRA **OR** ♦ Teacher frustration/discomfort with technology evident	♦ Basic evidence of teacher planning of ABRA lessons ♦ Basic evidence of teacher instructional guidance of ABRA ♦ Basic evidence of teacher monitoring students' use of ABRA ♦ Some evidence of teacher's awareness of zone of proximal development as students are instructed to move up task levels if too easy / completed or move back if too hard ♦ ABRA exposure – evidence of structured lessons ♦ Little off-task behaviour of students. ♦ Teacher and students appear comfortable with navigating through ABRA activities ♦ Some evidence of differentiated use of ABRA activities, but mainly within one skill level (i.e. Phonics / Word Level activities).	♦ Clear evidence of teacher planning of ABRA lessons. Teacher links planning and target setting according to students ability level ♦ Clear evidence of teacher providing appropriate instructional guidance / feedback while students on ABRA ♦ Clear evidence of teacher monitoring students' use of ABRA ♦ Clear evidence of teacher's awareness of zone of proximal development as students are instructed to move up task levels if too easy / completed or move back if too hard ♦ ABRA exposure – evidence of structured lessons ♦ Students are clearly engaged in the lesson ♦ Teacher & students are comfortable with navigating through ABRA activities ♦ Extension of ABRA – Some evidence of entry-level activities that extend skills explored in one domain of ABRA, usually Word level (i.e., Rhyme matching game; spelling words or simple sentences; playing BINGO, etc). ♦ Evidence of differentiated use of ABRA activities, at more than one skill level (i.e. Word Level, Text Level or Writing activities). ♦ Evidence of collaborative work & use of collaborative learning opportunities	**Criteria as per level four with the addition of the following:** ♦ **Extension of ABRA – Clear evidence of extension activities that incorporate higher-level skills (i.e. Comprehension) that extend beyond simple WORD level activities. Examples such as, writing alternate story endings; journal entry reflections on ABRA story; creating a drama skit/puppet show based on ABRA story, etc.** ♦ **Teacher clearly differentiated use of ABRA across ALL FOUR suggested levels of implementation (i.e. Word Level, Text Level, Collaborative Work & Extension Activities)** ♦ **Teacher uses all collaborative learning opportunities – peer supported dialogues, different roles, reciprocal tutoring, etc.**

Original Observer / Class / Date: ______________________________

RATER: ______________________ **DATE:** ______________________ **SCORE:** ☐