Small Group Administration of a Mathematics Intervention in a School Setting

by

Danielle Raylee Mattson

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Education

in

SCHOOL AND CLINICAL CHILD PSYCHOLOGY

Department of Educational Psychology
University of Alberta

Abstract

Low math competence has been linked to negative outcomes in a number of areas of life. Adults with poorer math skills are more likely to report lower wages and rates of employment and more frequent health concerns and legal woes than those with higher math proficiency. As children's early math abilities are significantly predictive of their later math competence, it is crucial that effective interventions are implemented for those experiencing math difficulties. One such intervention is the Math Interactive Learning Experience (MILE). MILE was originally developed to help strengthen math abilities in children with fetal alcohol spectrum disorders (FASD), and although efficacious, the scope of this program was limited by the time and cost associated with delivering an individually-administered, lab-based intervention. The current pilot study ($N = 14$) was therefore carried out to determine whether the impact of the MILE program could be maintained when it was modified to increase feasibility in three key ways: (1) by shifting administration to educational professionals in a school setting, (2) by delivering the intervention in small groups, and (3) by including children with a variety of neurodevelopmental and learning difficulties. Results showed that the modified MILE program had a positive impact on participants' math scores. Children in the intervention group improved significantly more on a measure of math achievement than children in the waitlist control group, and importantly, these math score gains were also maintained at the time of 6-month follow up. Furthermore, nearly 86% of children who completed the intervention improved by a standard deviation or more in at least one math content area on the same measure. However, despite these positive findings, feedback from the educators who delivered the intervention also revealed issues with program feasibility that must be addressed in future studies.

Preface

This thesis is an original work by Danielle Mattson. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Health Research Ethics Board, Project Name "The Impact of the Math Interactive Learning Experience (MILE) program on math performance of young children with FASD/PAE and other neurodevelopmental difficulties: Evidence from classroom based settings," HREB No. Pro00052708, March 16, 2015.

Acknowledgements

To my supervisor Dr. Jacqueline Pei: thank you for your guidance, encouragement, and unwavering belief in me. The wisdom you have shared in our many long talks has left a lasting impression on me, and I am a better person, researcher, and clinician because of it.

To my supervisory committee: thank you Dr. Carmen Rasmussen for welcoming me into your research team and providing thoughtful feedback throughout this process. Thank you as well Dr. Phil Sevigny: your kindness and expertise has been greatly appreciated. I am lucky to have gotten the opportunity to learn from you both.

To the dedicated schools, educators, EAs, caregivers, and participants who made this study possible: thank you for being so generous with your time and spirit. A special thanks to the research coordinator for this study, Kathryn Kryska: your contributions (and friendship) were invaluable.

To my parents: thank you for the endless love and support you have always given me. It is because of you that I have even had the opportunity to pursue this dream, and I aim to make you proud every step of the way. You are my heroes. To the rest of my family, friends, the Dewey's crew, and Maybe: thank you for reminding me what matters most. To the six incredible women I was fortunate enough to walk alongside in this program: you have been a source of laughter, comfort, and inspiration, and my life is truly richer for knowing each of you. This is only the beginning.

Finally, to Ryan: thank you for *everything*. You have stood beside me through the darkest and lightest of times, and I will never be able to fully express my gratitude for that. You are the joy of my life.

Table of Contents

List of Tables

List of Figures

Introduction and Literature Review

Importance of Math Competence

Mathematical competence is crucial for everyday success. This is particularly true in modern Western societies where math is an essential part of many tasks of daily living: personal budgeting, time management, and various aspects of academic performance all rely heavily on the ability to understand and work with numbers. Over the course of an average day, an individual may need to use their knowledge of math concepts to determine what time to leave for work in the morning, how much fuel they will need to put in their car get there, the amount of cash they need to withdraw for their groceries, and how much change to put in the parking meter and leave as a tip at a restaurant. Less obvious are the secondary impacts of low math understanding on overall wellbeing. In fact, research has shown that math abilities may even have a greater impact on an individual's life chances than literacy abilities do. For example, a large study out of the UK that followed 17,000 babies from birth to present reported that at age 30, individuals with low numeracy skills are more likely to be unemployed, earn lower wages when they are employed, report poorer mental and physical health, and encounter more trouble with the law than those with low literacy skills (Parsons & Bynner, 2005). The level of math education a student attains in high school was also found to explain a large portion of the later differences in earnings between a child from a low-income family and a child from a medium-income family, even after controlling for overall GPA and other demographic, school, and family characteristics and proxies of ability and motivation (Rose & Betts, 2004). When you consider that even basic aspects of personal health care – such as the ability to understand medication schedules and interpret nutritional labels – require math knowledge and understanding (Ancker & Kaufman, 2007), the vital importance of building mathematical competence becomes clear.

Definition of Math Competence

Although the importance of math competence is widely accepted, a definition of the construct is not. For example, many definitions describe the concept in terms of the individual skills and cognitive processes required to demonstrate math competence. Both the number and nature of these skills differ in the literature (e.g., Niss, 2004; Turner, 2010), but at a minimum, most definitions acknowledge that math competence requires the successful application of communication, problem solving, modelling, and reasoning skills (Vorobjovs, 2017). Other definitions focus less on identifying the constituent skills and instead emphasize the ability to utilize them in a variety of different situations. Perrenoud (1997, as cited in Abrantes, 2001) argues that this capacity to 'improvise' distinguishes competence from *performance,* a distinction also implied by Domazet, Baranović, and Matić (2013) when they described math competence as "the mathematical ability that can be, and is, applied outside the strict mathematical factual knowledge reproduction tests" (p. 115).

Despite discussion about each of these elements of math competence separately in the literature, a concise definition that captured both was difficult to locate. Thus, for the purposes of this paper, "math competence" will refer to the ability to "apply mathematical thinking in order to solve a range of problems in everyday situations" (p. 6, European Communities, 2007).

Developing Math Competence

Learning trajectory. In order to possess the math competence defined above, an individual must first *learn* the math skills, concepts, and strategies that they will need to use to solve day to day to problems. Math learning typically follows a relatively defined trajectory as it is necessarily cumulative and hierarchical. For example, the ability to recognize and order numbers must come before operations like addition and subtraction can be understood, and proficiency with addition and subtraction must be reached before concepts like algebra and geometry can be learned. Thus, if a child fails to master a basic skill, they are likely to struggle with subsequent concepts that build upon it; Jordan and Levine (2009) refer to this snowball effect as a "cascade of mathematics failure."

Cognitive processes. Failure to master math skills can occur for a number of reasons. For many children with math difficulties, certain cognitive processes have been identified that contribute to their math learning challenges. For example, a meta-analysis of over 13,000 individuals with math difficulties found that deficits in processing speed and working memory were the strongest and most stable cognitive markers of such challenges (Peng, Wang, & Namkung, 2018). However, there are a number of other cognitive underpinnings important for successful math learning, and it is generally accepted that (1) poor working memory (e.g., Raghubar, Barnes, & Hecht, 2010), (2) reduced visual-spatial abilities (e.g., (Assel, Landry, Swank, Smith, & Steelman, 2003; Geary, 2004), and (3) difficulty performing executive functions such as inhibition, switching, and updating (e.g., Bull & Scerif, 2001; St Clair-Thompson & Gathercole, 2006; Yeniad, Malda, Mesman, van IJzendoorn, & Pieper, 2013) are all associated with math learning dysfunction.

Math Interventions

Because of the cumulative nature of math learning and the demonstrated importance of math competency on many markers of wellbeing in adulthood, early intervention is crucial for children who struggle with math learning. This is particularly true when you consider that early math abilities have been shown to be predictive of later math performance in grade 3

(Duncan, et al., 2007), grade 5 (Claessens, Duncan, & Engel, 2009), middle school (Geary, Hoard, Nugent, & Bailey, 2013), and at age 15 (Watts, Duncan, Siegler, & Davis-Kean, 2014), even after controlling for variables such as IQ, reading achievement, attentional control, and socio-emotional skills in many cases.

But how do math interventions actually help children who are not responding to traditional math instruction? One program, the Math Interactive Learning Experience (MILE), works by supporting the very cognitive processes that are linked with math learning.

MILE

The MILE program was developed at Emory University in Atlanta, Georgia by Dr. Claire Coles, Dr. Julie Kable, and Dr. Elles Taddeo. It was designed to improve math abilities in children affected by prenatal alcohol exposure by not only teaching math skills directly but by doing so while supporting the underlying neurocognitive difficulties that children with fetal alcohol spectrum disorders (FASD) experience (Kable, Coles, & Taddeo, 2007). For example, as processing speeds are often lower in children within this population, instruction is delivered at a slower pace to accommodate this. As visual-spatial abilities are also reduced, MILE uses simple tools like graph paper to help children align number columns and tangible manipulatives to help them better understand concepts such as sorting and ordering, measuring, and computing simple operations. These concrete objects are also used to help children acquire mental representations of numeracy and support deficits in working memory as they reduce the cognitive load that a child must hold to successfully perform a task.

In addition to these compensatory tactics, an important component of the MILE program is the "FAR" teaching strategy. FAR, which stands for *Focus/Plan*, *Act*, and *Reflect*, was adapted from the High/Scope Perry Preschool (HSPP) Project's "plan-do-review" approach (Kable, Coles, & Taddeo, 2007). The study, which aimed to reduce long-term school failure rates for pre-school children living in poverty, based this approach on Jean Piaget's theory that children are intentional learners who benefit from planning, carrying out, and then reviewing their own activities (Schweinhart, 2006). Under this framework, adults support and encourage children's learning by providing opportunities for self-directed exploration and engaging in thoughtful process-driven (rather than content-driven) questioning.

MILE's FAR strategy emphasizes interactive learning using the same guiding principles as the HSPP Project. Metacognitive questioning, scaffolded prompting, and specific praise are used to facilitate engagement, reinforce progress and positive behaviors, and help children become more aware of their own thinking and more reflective in their problem solving. Metacognitive questioning focuses on the learning process and involves using statements that foster reflection such as, "Tell me how you did that." Scaffolded prompting involves giving children time and space to respond before moving incrementally 'closer' with prompts in an effort to help them solve problems and draw conclusions themselves. Finally, specific praise is offered to reinforce the positive elements of a child's actions and behaviour and reward and encourage their efforts and successes.

The evolution of MILE. The first MILE study was carried out in Atlanta in 2007 by program developers Kable, Coles, and Taddeo (2007). The intervention was comprised of three main components: (1) the math intervention, which was administered individually in a clinical laboratory setting; (2) psychoeducational training sessions for caregivers; and (3) socioemotional and behavioural support services (e.g., psychiatric consultations) to help address any outstanding concerns in these areas. Participants (N = 56) were aged 3 to 10 and had either a pre-existing diagnosis of fetal alcohol syndrome (FAS) or partial FAS (pFAS) or exhibited significant alcohol-related dysmorphic features as measured by a pediatric geneticist. The study found that the children who participated in the 16 hour, 6-week math intervention demonstrated significantly higher math achievement gains than those who did not. Children in the MILE group were also significantly more likely to demonstrate an improvement of one standard deviation or more (i.e., a clinically significant gain) on the math outcome measures used by the researchers.

In 2009, a second report was published by the program developers examining whether the math gains observed in the initial study had been maintained six months after completing the intervention (Coles, Kable, & Taddeo, 2009). Follow-up testing was conducted with 54 of the original participants, and results showed that not only were the math gains maintained, but the children who participated in the MILE program also still demonstrated significantly greater math gains than the control group from pre-testing to follow up.

Given the success of the initial MILE study, the Atlanta-based research team ran a second study ($N = 52$) in 2015 with the goal of determining whether the MILE intervention was effective when administered by a trained non-clinical instructor in a community setting. The time frame for this intervention was also extended to 15 weeks which was more consistent with local tutoring programs (Kable, Taddeo, Strickland, & Coles, 2015). The sites were primarily private schools for children with learning disabilities, although one site, a community tutoring service, was located within a university. The instructors had a variety of backgrounds ranging from college students with no teaching experience to licensed teachers. Again, the research team found that children who received the MILE intervention improved significantly more than those who did not. However, this time the researchers also determined that the intervention was equally as effective when administered in the community as in a clinical setting, opening the door to increased accessibility and public translation of the MILE program.

Building on the promising results of the prior two studies in Atlanta, a research team in Edmonton, Alberta replicated and extended the MILE program ($N = 28$) to examine whether the math gains exhibited in prior studies could be retained when modifications were made to further enhance community accessibility (Kully-Martens, et al., 2018). These changes were significant, and included eliminating the parent training component, administering the intervention in an ecological setting (i.e., a participant's school or home), and including an alternate intervention for the control group (a social skills intervention). This study also used more flexible inclusion criteria, incorporating participants who had confirmed prenatal alcohol exposure (PAE) but did not have a clinical diagnosis of FASD. Although smaller than the studies conducted in Atlanta, the researchers again found that the children in the MILE group saw significantly higher math score improvements than the children who received only a social skills intervention. The intervention group also still showed greater gains at the time of 6-month follow up than the control group. Interestingly, different participant characteristics were found to be associated with an increased response to the MILE intervention in the studies in Atlanta and Edmonton; while Coles et al. (2009) found that younger children were more likely to see math score gains, Kully-Martens et al. (2018) found that older age was associated with greater overall math improvement.

Present Study

This pilot study aims to build upon previous studies and further extend the MILE program. Specifically, the adaptations detailed here were chosen with the goal of increasing program feasibility to allow a greater number of children who may benefit from the intervention to access it. Although Kully-Martens et al. (2018) found very promising results by removing what may be significant barriers for some – the caregiver training component and lab-based administration – the reach of the original MILE program was still limited by the time, cost, and manpower required to deliver the intervention to children one-on-one in an extracurricular setting (i.e., outside of school). To address these limitations, three key modifications were made. First, the intervention was to be administered by trained educators or educational assistants and carried out entirely in a school setting. Second, the program was shifted from individual administration to small-group delivery. Finally, the inclusion criteria were expanded to allow children with other neurodevelopmental difficulties (including, but not limited to, FASD and PAE) to participate in the intervention as the cognitive difficulties experienced by children with FASD (and supported by MILE) are also common to other disorders.

Research Questions

Broadly, the present study aims to examine whether the modified MILE intervention is effective when delivered by educators to small groups of children with various neurodevelopmental difficulties in a school setting. As this study is part of a larger research project examining the impact of the program, the current paper will define “effectiveness” as whether math score improvements are observed and whether educators find the program feasible to deliver, as a program that is not practical or reasonable to employ is not likely to be utilized. This overarching question will be explored in four ways.

Research Question 1a

Does the immediate intervention group exhibit greater math score gains than the waitlist control group?

Hypothesis. Children who participate in the modified MILE intervention will exhibit greater math score gains than children in the waitlist control group (who receive math instruction as usual).

Research Question 1b

Do children's math scores improve from pre- to post-intervention when all participants who complete the intervention are considered together?

Hypothesis. Children who participate in the modified MILE intervention will score higher on math achievement measures after completing the program than before they began.

Research Question 2

Are math score gains maintained 6 months after completing the intervention?

Hypothesis. When children are tested 6 months after completing the modified MILE intervention, their math scores will be as high as or higher than at the time of post-testing.

Research Question 3

Are any participant characteristics correlated with greater math achievement gains?

Hypothesis. Based on the results of the study by Coles et al. (2009), children who are younger at the time of baseline testing will see greater math score improvements than children who are older. Neither sex, IQ, nor diagnostic status will be significantly correlated with the magnitude of math score changes.

Research Question 4

Do interventionists find this program feasible to implement with the current modifications in place?

Hypothesis. Interventionists will find the program feasible to implement in schools based on the small group size and the formal training they receive.

Methodology

Participants

Twenty-eight children ages 5 to 8 were initially recruited to participate in this pilot study. These children had been identified by their current teacher as students who they felt would benefit from additional math instruction. From the teacher-referred group, participants were considered for this study if they also (1) experienced neurodevelopmental difficulties that may impact their learning, and/or (2) had been identified as experiencing some or significant difficulty in at least one of the following domains on the *Early Years Evaluation – Teacher Assessment* (*EYE-TA*): Awareness of Self and Environment, Social Skills and Approaches to Learning, and/or Cognitive Skills. The *EYE-TA* is a developmental screening tool that provides a framework for measuring and recording observations about a child's performance in five areas (The Learning Bar, 2019). The 55-item structured checklist is un-normed but rather uses a 4-point scale to measure the developmental strengths and weaknesses of students to identify those who may require further support or evaluation. The *EYE-TA* is utilized in many school districts throughout the country (including a number in the Edmonton area), and is suitable for use with children aged 4 to 6 in kindergarten or grade 1.

Of the 28 children who originally consented to participate in the study, 14 students from three schools completed the MILE intervention. This substantial attrition occurred for a number of reasons. Ten participants were lost after pre-testing but before beginning the program as their interventionists no longer wished to participate. Two more participants withdrew from the study individually. In one case, the participant changed schools; in the other case, the parents withdrew consent but did not provide a reason for doing so. The final two participants were lost as their interventionist encountered significant health difficulties and was not able to complete the program. Of the 14 participants who completed the intervention, six children from two schools were assigned to an "immediate intervention" group and eight children from one school were assigned to a "waitlist control" group. This study design was employed to allow all students to receive the intervention regardless of treatment group. To minimize disruption to teachers and classrooms, the groups were created based on convenience sampling; that is, all children participating from the same school belonged to the same treatment group.

Inclusion criteria demographics. Within the sample, every participant whose *EYE-TA* data was available (N = 13) had been identified as experiencing some or significant difficulty in at least one of the three areas of interest for inclusion, although many had difficulties in the two other areas reported as well (mean number of areas of difficulty = 3.15). In addition, 11 participants (78.6%) had also been diagnosed with a neurodevelopmental, mood, or behavioural disorder listed in the *Diagnostic and Statistical Manual of Mental Disorders,* Fifth Edition (*DSM-5*; American Psychiatric Association, 2013) that was impacting their learning. The most common diagnosis was attention-deficit/hyperactivity disorder (50.0%), although other diagnoses included oppositional defiant disorder (28.6%), specific learning disorder (14.3%), undisclosed anxiety disorders (14.3%), fetal alcohol spectrum disorder (7.1%), and autism spectrum disorder (7.1%), with many children presenting with more than one diagnosis.

Other demographics. Table 1 provides further details about participant characteristics both by group and at the broader sample level. Following comparative analysis, only one significant group difference was found, and that was in classroom type (p = .015, Fisher's exact test). However, this was to be expected as treatment groups were formed based on school participation, and all eight students in the waitlist control group were members of the same classroom (a Behaviour and Learning Assistance Program classroom; BLAP).

Table 1. *Participant Characteristics at Baseline Testing (Time 1)*

Participant characteristic	Group		p	All participants (N = 14)
	Immediate intervention (n = 6)	Waitlist control (n = 8)		
Age in years [*M* (range)]	7.2 (7-8)	6.8 (5-8)	.309[b]	6.9 (5-8)
Sex [*n* male (%)]	4 (66.7%)	4 (50.0%)	.627[c]	8 (57.1%)
Grade in school			.473[c]	
1	0 (0.0%)	2 (25.0%)		2 (14.3%)
2	6 (100.0%)	6 (75.0%)		12 (85.7%)
Classroom type			.015[c]*	
Regular	4 (66.7%)	0 (0.0%)		4 (28.6%)
BLAP[a]	2 (33.3%)	8 (100.0%)		10 (71.4%)
Psychological diagnosis [yes (%)]	3 (50.0%)	8 (100.0%)	.055[c]	11 (78.6%)
WRIT General IQ score [*M* (range)]	90.0 (76-103)	99.1 (83-125)	.219[b]	95.2 (76-125)
Keymath-3 DA standard score [*M* (range)]	75.8 (55-96)	85.0 (65-111)	.370[b]	81.1 (55-111)
Presence of *EYE-TA* difficulties [yes (%)]				
Awareness of Self/Environment	2 (33.3%)	3 (37.5%)	1.000[c]	5 (35.7%)
Cognitive Skills	4 (66.7%)	5 (62.5%)	1.000[c]	9 (64.3%)
Social Skills/Approaches to Learning	4 (66.7%)	7 (87.5%)	.192[c]	11 (78.6%)
Language/Communication	3 (50.0%)	4 (50.0%)	1.000[c]	7 (50.0%)
Physical Development	5 (83.3%)	4 (50.0%)	.559[c]	9 (64.3%)
Annual household income [*n* > $50k (%)]	4 (66.7%)	4 (50.0%)	1.000[c]	8 (57.1%)

Note. Missing *EYE-TA* and household income data for one participant in the waitlist control group.
[a] Behaviour and Learning Assistance Program, [b] analyzed using independent-samples T-test, [c] analyzed using Fisher's exact test
*$p < 0.05$

Ethical Considerations

Before undertaking this study, we first sought and received approval from the University of Alberta's Health Research Ethics Board (HREB) and the Cooperative Activities Program (CAP), a team who oversees research projects between the university and local school districts. We were then provided with a list of schools who were open to participating in school-based research, and principals from these schools were contacted with information about the current study. A research assistant then met with interested teachers from these schools to discuss the study and provided information sheets outlining the risks, benefits, terms of confidentiality, and voluntary nature of the study before obtaining written consent.

Next, when students were identified by a participating teacher as potential candidates for the intervention, information sheets and consent forms were sent to parents via the school. These forms were similar to those provided to the educators. Finally, and importantly, before beginning the baseline testing process, the same information was provided to the children using age-appropriate language to facilitate understanding as best as possible. Verbal and written assent was then obtained from the participants and was closely monitored and adhered to throughout the course of the intervention.

Procedure

All participants underwent approximately 1 to 2 hours of math achievement and/or cognitive functioning (IQ) testing at three time points during the study period (Figure 1). Testing took place in a quiet room at the child's school and was completed by either a graduate student or experienced research assistant, all of whom held a psychology degree and received additional training on the measures used. Testing was sometimes spread across multiple days to respect teacher and classroom schedules and accommodate students who had difficulty completing the assessment in a single session. Children were provided small toys and prizes at the end of each testing session, but no other remuneration was provided. The entire study period lasted just over one calendar year: the immediate intervention group (two schools) completed all testing from March to December 2016, and the waitlist control group (one school) completed all testing from December 2016 to May 2017.

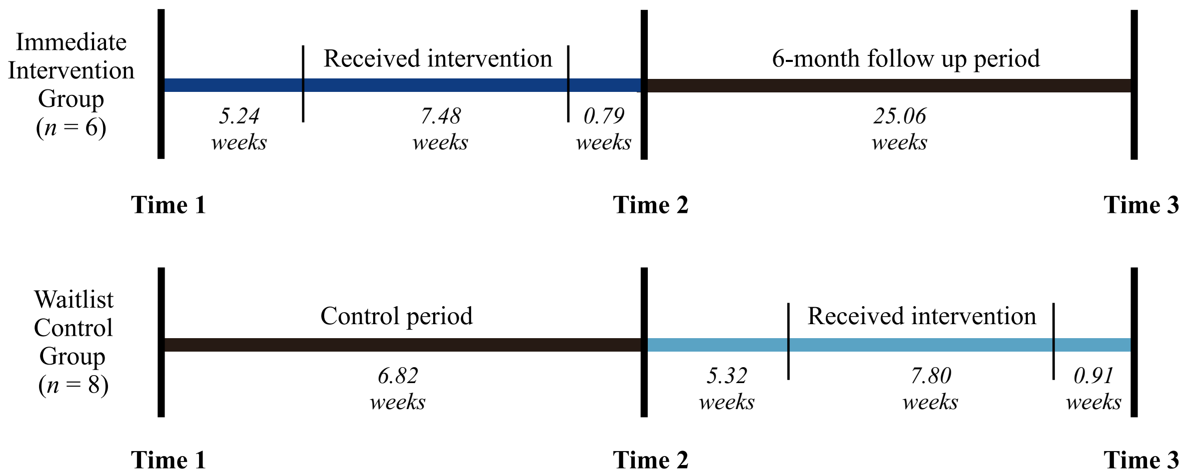

Figure 1. Testing procedure and timeline in weeks.

Intervention

Interventionists. The intervention was administered to small groups of two participants each by an educator, educational assistant (EA), or research assistant (RA; this acronym will only be used when referring to the individual in the role of interventionist). Initially, the program supposed to be delivered by school staff only (i.e., educators or EAs), however, when two interventionists determined it would not be feasible for them to administer the program to four groups from one classroom, a study RA agreed to take on two groups to enable the children who had consented to still receive the intervention.

All interventionists received formal training in the MILE program at the University of Alberta over the course of approximately six hours. Eleven educators and EAs (and one RA) completed this training; five (and one RA) completed the intervention. Instruction was provided by study investigators, all of whom received training from the original program developers. The training workshop included an overview of the study, educational instruction, and videos outlining and explaining the key facets of the MILE program (e.g., the FAR methodology). Time was also allotted for attendees to familiarize themselves with the materials, practice with fellow educators, and discuss any questions with the trainers. Interventionists received a $25 gift card for their participation in the study.

MILE intervention. To accommodate the shift from individual to small-group administration, the total number of sessions in the intervention was increased from 10 to 14. The intervention took place approximately twice a week until each participant had completed all 14 sessions. The average length of the intervention was 7.66 weeks (SD = 1.08) for the entire sample; the length of the intervention was not significantly different between groups with a mean difference of only 0.33 weeks, $p > .05$.

Interventionist duties. Each intervention session had three core duties: planning, teaching, and reflecting. Before beginning the MILE program, interventionists were provided with each child's *KeyMath-3 DA* pre-test scores so they could determine which skills to target to first. Then, before each subsequent session, interventionists chose a goal or curriculum page to work on based on these scores and their notes from the previous session, and devised a lesson plan. Finally, following the session, instructors noted on the planning sheet any comments about what did or did not work, whether or not the learning objectives were met, and anything that needed to be carried over into the next session.

Session details. Each session was carried out in three phases according to FAR methodology. Sessions always began by co-creating the day's schedule with the participants (*Focus/Plan*). This was done on a sheet of paper with blank spaces to fill in for (1) the day's goal ("Today we will ___"), (2) three possible activities that would be used to achieve this goals, and (3) two reminders relevant to the participants or the session. Children were encouraged to contribute their ideas to this sheet and although the instructor created a lesson plan with the session's objectives prior to beginning, children were given as much choice as possible over how they would like to accomplish these goals (e.g., by choosing the manipulatives used). The children then carried out the activities aligned with the learning objectives (*Act*) as the instructor supported them and facilitated engagement through questioning, praise, and prompting. Finally, at the end of the session, children and instructors reviewed the plan they had created together and discussed what they had learned in the day's session (*Reflect*).

Measures

Demographic questionnaire. All caregivers completed a demographic questionnaire before the intervention began to provide information about their child's age, grade, sex, ethnicity, medical and educational history as well as caregiver/household factors such as annual income range, relationship to child, and educational attainment.

MILE Instructor Satisfaction Evaluation Form. Immediately following completion of the intervention, interventionists completed a questionnaire providing information about their experience delivering the modified MILE program and their opinions about its strengths, weaknesses, and feasibility. The questionnaire contained a total of 17 items including seven short answer questions, five multiple choice questions, and five questions posed on a 5-point Likert scale ranging from strongly agree to strongly disagree. An example of the full questionnaire can be found in Appendix A.

Wide-Range Intelligence Test (WRIT). The *WRIT* was administered at the time of pre-testing to provide an estimate of participants' baseline intellectual abilities. The *WRIT* is a brief, standardized test of cognitive functioning suitable for use with individuals aged 4-85 (Glutting, Adams, & Shelow, 2000). The test is comprised of four subtests that combine to generate two composite scores (Verbal IQ and Visual IQ) and one overall score (General IQ). The General IQ composite, which was the area of interest for this study, is typically reported as a standard score with a mean of 100 and standard deviation of 15. This composite has high internal consistency, test-retest reliability, and interrater reliability (all above .90) and adequate construct, concurrent, and predictive validity.

KeyMath-3 Diagnostic Assessment – Canadian Edition (KeyMath-3 DA). The *KeyMath-3 DA* was administered at all three time points to determine current math achievement scores. The *KeyMath-3 DA* is a standardized tool used to measure math abilities in children aged 5 to 17 (Connolly, 2008). It is available in two parallel forms, Form A and Form B, which were alternated at consecutive time points to minimize any practice effect.

The *KeyMath-3 DA* is comprised of 10 subtests that combine to create three composite scores and one total score. The 372-item instrument covers a wide variety of math concepts that span the math learning continuum from number and shape identification to factoring polynomials. The first composite, Basic Concepts, encompasses five subtests that measure foundational procedural and computation skills (Numeration, Algebra, Geometry, Measurement, and Data Analysis and Probability). The second composite, Operations, consists of three subtests that assess both mental and written calculation abilities (Mental Computation and Estimation, Addition and Subtraction, and Multiplication and Division). The third composite, Applications of Problem Solving, includes two subtests that measure the application of conceptual knowledge and operational skills (Foundations of Problem Solving and Applied Problem Solving). All subtest scores are converted to scaled scores with a mean of 10 and standard deviation of 3; all composite scores (including the Total index score) are converted to standard scores with a mean of 100 and standard deviation of 15.

Norms. The *KeyMath-3 DA* was normed based on 2006 Canadian census data and was closely matched on province, urbanization, and sex statistics. The 71 schools chosen for standardization testing were also selected to represent the cultural and socioeconomic diversity of Canada. Standardized scores are not available for younger students on some subtests; for example, Multiplication and Division scores are not available until age 7. For consistency based on the age of our participants, scores on this subtest are not reported in the current study.

Reliability. All reliability statistics reported for the *KeyMath-3 DA* were computed using an American sample of children aged 5 to 21 (the age range considered in the American-normed instrument); Canadian statistics were not uniquely available. Three types of reliability were reported: internal-consistency, alternate-form and test-retest reliability. The vast majority of coefficients in each of these areas were above .80, indicating that they are adequate or better by most standards (e.g., Erford, 2013).

Validity. Two types of validity were reported in the *KeyMath-3 DA* manual: content validity and construct validity. Evidence for content validity was provided by detailing how the test items were developed for the Canadian edition of the instrument. This involved connecting each item to a curriculum expectation in at least one of the four major math curricula in Canada and conducting an independent content review with a leading Canadian math educator. Evidence for construct validity was provided through subtest intercorrelations as well as correlations with other math achievement measures. All coefficients accompanying these procedures generally exceeded .60, again indicating that they are adequate or better by most standards (e.g., Drummond, Sheperis, & Jones, 2016, as cited in Swank & Mullen, 2017).

Enhancing Reliability and Validity

Program fidelity. A 15-point fidelity to intervention checklist (Kable, Taddeo, Strickland, & Coles, 2015) was also completed by a research assistant familiar with the MILE program at two time points during the study for each interventionist. The interventionists were scored *yes* (2 points), *sometimes*/*partial* (1 point) or *no* (0 points) on the presence of important MILE teaching methodology such as "the instructor allowed the child to choose some aspect of the math fun work" and "the instructor praised the child's efforts" as well as logistical details such as "the instructor's materials were readily available."

The fidelity scores ranged from 26 to 30 out of 30, with an average score of 27.42 out of 30. Only three instances of *no* were recorded, each for separate instructors. Two of these occurred because the interventionist did not co-create the lesson plan with the children at the beginning of the session; in these cases, the interventionists reported that this required too much instructional time, and so instead they made the day's plan prior to the session and reviewed it with the participants before beginning. The third instance of *no* was scored as such because the instructor did not discuss breaks or behavioural contracts with the participants.

Testing and scoring. To minimize practice effects on the measure of math achievement, alternate forms of the *Keymath-3 DA* (Form A and Form B) were used at consecutive testing time points. In addition, to ensure that testing materials were correctly scored and data was entered into SPSS without errors, all instrument protocols were double-scored and the data file was reviewed by both an experienced research assistant and a psychology graduate student.

Data Analysis

Data analysis took place using IBM SPSS 25. To begin, descriptive analyses were run to determine the spread and nature of the sample demographics, and the database was manually searched for missing data.

Missing data. An annual household income bracket was found to be missing for one participant as caregivers did not provide this information. *EYE-TA* scores were also not available for one participant; inclusionary criteria were met for this child through a neurodevelopmental disorder diagnosis. Two participants from the waitlist control group were also noted to be missing some Time 2 math achievement scores. In the case of the first participant, a subtest and its respective composite could not be calculated due to administrator error. In the case of the second participant, the child requested to discontinue testing on both occasions it was attempted, and as a result, data could not be collected for a number of subtests and composites. These two participants were therefore excluded from analyses that required *KeyMath-3 DA* Total index scores as it was not possible to derive them. Finally, one participant in the immediate intervention group could not be reached at the time of 6-month follow up as they had moved away, so Time 3 data is not available for this child.

Research question 1a. In order to determine if those in the immediate intervention group achieved greater math score gains than those in the waitlist control group, independent sample T-tests were conducted to compare the mean *KeyMath-3 DA* Total score change from Time 1 to Time 2. Raw scores were used for this analysis as they were more sensitive to change than standardized scores. For example, although one participant gained 24 raw score points over the intervention period, this equated to only a 3-point standard score gain which masked the true magnitude of this student's improvement.

Research question 1b. To further examine the impact of the modified MILE program, the treatment groups were combined and paired-samples T-tests were run to determine if the mean post-intervention scores were significantly higher than the mean pre-intervention scores on each subtest and composite when calculated across all participants. Again, this was done using raw scores so as to better detect changes that may be meaningful to individuals. In addition, to better understand math achievement score changes at the individual level, subtest and composite standard score gains were categorized by for each participant by the number of areas they improved by at least one standard deviation in and summarized at the sample level.

Research question 2. To answer the second research question, paired-sample T-tests were used to compare the immediate intervention group's *KeyMath-3 DA* raw scores on each subtest and composite at post-test and 6 months after completing the intervention.

Research question 3. To determine if any participant characteristics were associated with higher math score gains in response to the intervention, biserial and point-biserial Pearson product-moment correlations were calculated between demographic variables and raw *KeyMath-3 DA* change scores at the composite level.

Research question 4. Finally, as the modifications made to this iteration of the MILE program were made to increase feasibility (and thereby accessibility), interventionist feedback forms were collected and analyzed based on instructor responses to scaled, multiple choice, and short answer questions. Answers on the Likert scale and multiple choice questions were tallied across interventionists to better understand overall trends in responses, and short answers were considered descriptively in the context of the feedback provided and then grouped at the larger sample level.

Results

Research Question 1a

Immediate intervention vs. waitlist control group. We hypothesized that the immediate intervention group would improve more overall than the waitlist control group as a result of participating in the modified MILE intervention. Two participants from the waitlist control group were not included in this analysis as their *KeyMath-3 DA* Total index score could not be calculated at Time 2 due to missing subtest scores. Still, as shown in Figure 2, independent-samples T-test results demonstrated that the children who received the intervention achieved significantly greater raw score gains (M = 10.50, SD = 4.09) than the children who did not receive any additional math instruction (M = 3.33, SD = 3.33), $t(10) = 3.33$, $p = 0.008$. Significant differences were not found between groups on any other subtests or composites, all ps > .05.

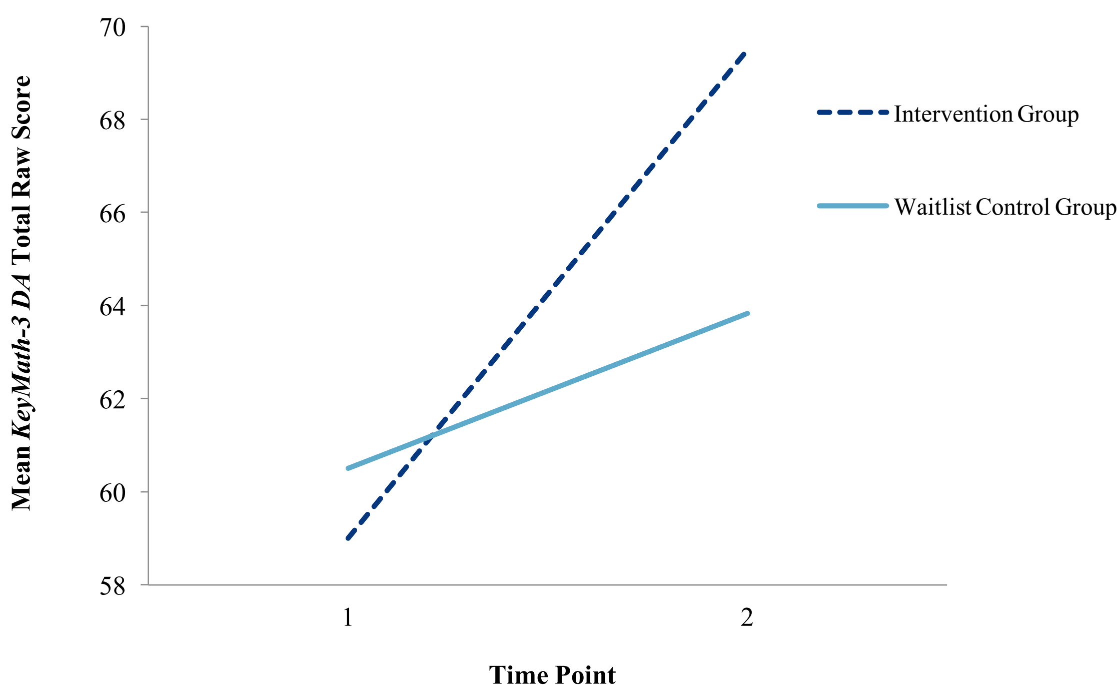

Figure 2. Mean *KeyMath-3 DA* Total raw score change from Time 1 to Time 2 by treatment status.

Research Question 1b

Sample-wide intervention gains. We hypothesized that mean math scores on each subtest and composite would be higher post-intervention than pre-intervention when calculated across all children who received the intervention. Two participants had some missing data in this analysis as well; however, the scores that were available for these children were included where possible (see *n* column in Table 2). Paired-samples T-test results showed that statistically significant raw score growth occurred on more than half of the *KeyMath-3 DA* subtests and composites, namely Basic Concepts, Algebra, Measurement, Data Analysis and Probability, Applications of Problem Solving, Foundations of Problem Solving, Applied Problem Solving, and the Total index.

Table 2. *Raw Score Changes Pre- and Post-Intervention for All Participants*

KeyMath-3 DA content area	Mean Raw Score			*N*	*p*	Effect size[a]
	Pre-intervention	Post-intervention	Change			
Basic Concepts	42.54	51.31	8.77	13	.001**	1.281
Numeration	10.86	11.36	0.50	14	.680	0.113
Algebra	5.64	7.50	1.86	14	.016*	0.741
Geometry	11.57	12.64	1.07	14	.158	0.401
Measurement	5.77	8.54	2.77	13	.004**	0.968
Data Analysis/Probability	7.46	10.00	2.54	13	.000***	1.441
Operations	11.17	13.17	2.00	12	.155	0.441
Mental Computation/Estimation	4.85	6.39	1.53	13	.129	0.452
Addition/Subtraction	6.77	7.46	0.69	13	.239	0.343
Applications of Problem Solving (PS)	11.54	15.77	4.23	13	.000***	1.828
Foundations of PS	5.69	7.39	1.69	13	.001**	1.286
Applied PS	5.85	8.39	2.54	13	.003**	1.042
Total	61.42	76.33	14.92	12	.000***	1.583

[a] Effect size calculated using Cohen's $d = t / \sqrt{N}$
***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

The effect of the intervention was also analyzed by tallying each participant's standard score change on the *KeyMath-3 DA* subtests and composites and organizing them by the number of participants who saw meaningful improvement (i.e., by at least one standard deviation) in one or more (85.7%), two or more (57.1%), three or more (28.6%), and four or more (14.3%) math content areas (Figure 3). Only two participants (14.3%) did not improve by at least one standard deviation in any math content areas; however, it is worth noting that one of these participants was missing a composite and Total index score at Time 2 which resulted in less data available to analyze for this child overall.

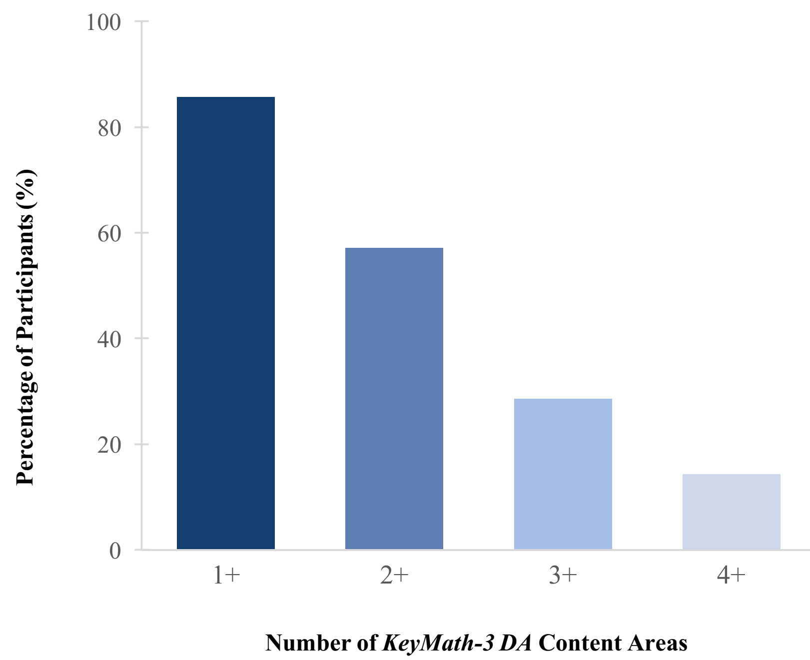

Figure 3. Percentage of children that improved by one standard deviation or more in each number of *KeyMath-3 DA* content areas.

Research Question 2

Maintenance of math score changes. We hypothesized that math gains achieved in each content area as a result of the intervention would be maintained at the time of 6-month follow up. Time 3 data was only available for 5 of the 6 immediate intervention group participants as one student moved away and was not reachable at the time of testing. Paired-samples T-test results showed that the math gains achieved following the intervention were maintained after six months and that no participants' raw scores had changed significantly in either direction (i.e., improved or regressed), *p*s > .05 (Table 3).

Table 3. *Raw Score Changes from Post-Intervention to 6-Month Follow Up for the Immediate Intervention Group*

KeyMath-3 DA content area	Mean raw score: Post-intervention	Mean raw score: 6-month follow up	Mean raw score: Change	*p*
Basic Concepts	47.00	47.20	0.20	.965
Numeration	11.40	10.60	–0.80	.670
Algebra	6.80	7.00	0.20	.871
Geometry	11.60	11.00	–0.60	.736
Measurement	7.80	8.20	0.40	.648
Data Analysis/Probability	9.40	10.40	1.00	.230
Operations	10.40	11.60	1.20	.178
Mental Computation/Estimation	4.80	6.20	1.40	.052
Addition/Subtraction	5.60	5.40	–0.20	.838
Applications of Problem Solving (PS)	14.00	13.80	–0.20	.919
Foundations of PS	7.00	6.80	–0.20	.893
Applied PS	7.00	7.20	0.20	.749
Total	71.40	72.60	1.20	.798

Note. All calculations performed using *n* = 5.

Research Question 3

Correlations with participant characteristics. We hypothesized that older participants would see greater math gains based on the results of the previous MILE study that took place in Edmonton (Kully-Martens, et al., 2018). We did not hypothesize there would be a relationship between any other demographic variables and the magnitude of math score improvement. Results of the Pearson product-moment correlation analysis showed a significant and strong positive relationship between participants' *WRIT* General IQ score and *KeyMath-3 DA* Total raw score change (Table 4). This suggested that children with higher scores on the *WRIT* General IQ composite showed greater positive change in math achievement. A significant and strong negative relationship was also found between participant age at baseline testing (Time 1) and raw score change on the Basic Concepts composite, $r(11) = -.76$, $p < .01$. This suggested that younger children tended to show larger improvements on this foundational composite. All other correlations between participant characteristics and composite scores were non-significant.

Table 4. *Correlations Between Participant Characteristics and KeyMath-3 DA Subtest and Composite Raw Score Changes*

	Math composite			
Participant characteristic	Basic Concepts	Operations	Applications of PS	Total
Sex	.01	.15	.32	.18
Age (years)	$-.76^{**}$	–.08	.41	–.51
Presence of diagnosis	.06	.36	–.11	.19
WRIT General IQ score	.36	.32	.40	$.61^{*}$

**correlation significant at 0.01 level, *correlation significant at 0.05 level

Research Question 4

Program feasibility. We hypothesized that instructors would find the intervention feasible to deliver with the current modifications in place because of the small group size and supportive formal training provided. Of the 6 instructors who completed the study, 5 returned satisfaction evaluation forms; the RA who administered the intervention did not complete a form as the questions did not apply. Feedback on the forms was mixed. Most instructors reported observing some positive outcomes as a result of the intervention; 80% stated they learned something new from the MILE program and the same proportion felt that the intervention had benefitted their students (Table 5). Qualitatively, instructors reported liking the "hands-on exploration" and individualized nature of the MILE sessions. As one instructor commented, "[the students] seemed to be able to verbalize their learning with more ease" as a result of participating in the intervention.

Table 5. *Summary of Responses to Scaled Items on the MILE Instructor Satisfaction Evaluation Form*

	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I enjoyed completing the MILE program		20%		60%	20%
I learned something new from the MILE program[a]				20%	60%
The intervention was beneficial for the students[a]				20%	60%
I would be willing to complete the program again as part of a research study	20%	40%	20%		20%
I would be willing to complete the program again as part of my regular duties at work	20%	20%		60%	
The intervention program would be beneficial if it was included in the regular curriculum	20%		20%		60%
I was able to choose activities/curriculum pages that applied to the needs/skill level of all children in my group				20%	80%

[a] One interventionist did not provide a response

Notably, however, instructors also reported a number of concerns with the program's delivery and feasibility. The most frequent written comments involved trouble fitting the sessions into their schedules (particularly in the case of the primary classroom educators) and difficulty balancing group needs and managing student behaviours. In addition, nearly one in two interventionists (40%) also reported that the amount of time spent preparing was a burden to their schedule (Figure 4), with one instructor stating that the time associated with planning and filling out forms was too stressful on top of other teaching duties. As a result, 40% of instructors somewhat or strongly disagreed that they would be willing to complete the program again as part of their regular teaching duties.

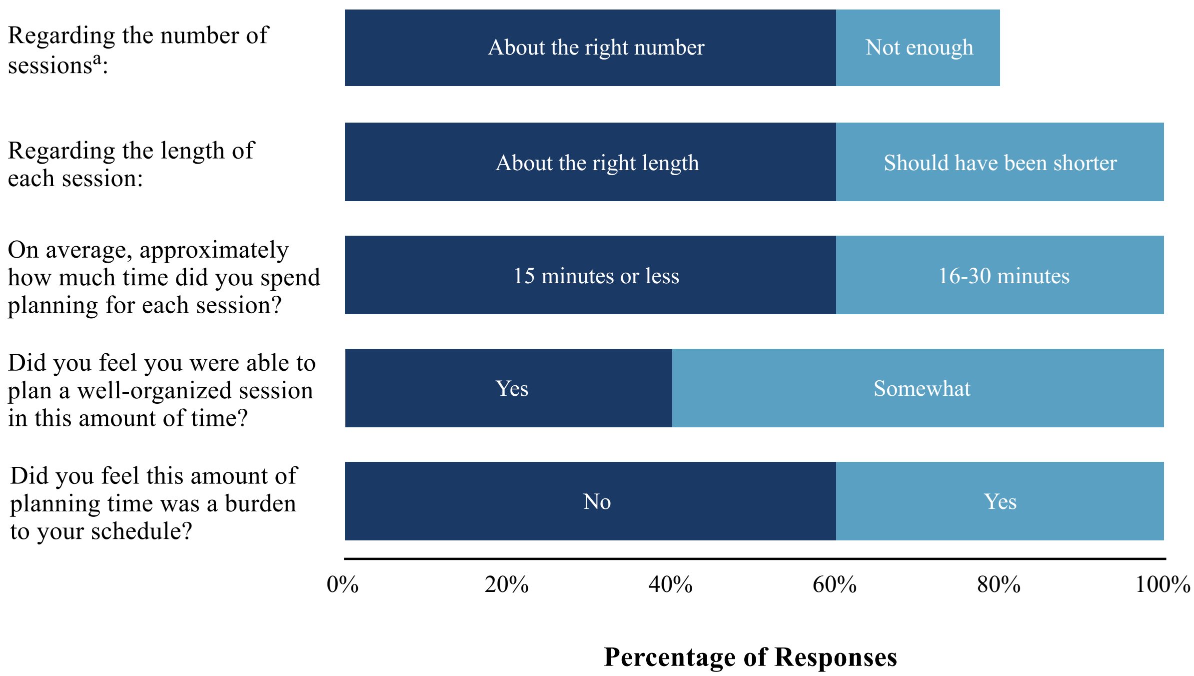

Figure 4. Summary of responses to multiple choice items on the MILE Instructor Satisfaction Evaluation Form.
[a] One interventionist did not provide a response

Discussion

Research indicates that an individual's level of math competence is linked to a number of important financial, psychological, and even physical health outcomes. Early intervention is especially important for young children who struggle with math learning because of both the importance of attaining math competence and the cumulative nature of math education. However, many promising and efficacious interventions do not end up having a 'real world' impact because of difficulty translating them from a research setting to the community.

The present study aimed to determine the MILE program was effective when modified to increase feasibility in three key ways: (1) by delivering it in a school setting; (2) by administering the program in small groups; and (3) by broadening inclusion criteria to allow children with other neurodevelopmental and learning difficulties (including, but not limited to, FASD or PAE) to participate. This entailed a marked extension from the original MILE program which was developed specifically for children with FASD and delivered individually in a clinical laboratory alongside a parent training component (Kable, Coles, & Taddeo, 2007). Although previous studies have successfully adapted the MILE program while retaining the math achievement benefits by shifting it out of the laboratory and into an ecological setting (i.e., a child's home or school; Kully-Martens et al., 2018), this was the first study to attempt to enable educator-led, small-group administration and to expand the participant catchment beyond only those with FASD.

Intervention Effectiveness

The first two research questions in the current study focused on the math score changes associated with the intervention. The results showed that children who participated in the modified MILE intervention achieved significantly greater math gains on the *KeyMath-3 DA* Total index than children who received math instruction as usual. When all children who completed the intervention were then considered as a larger group, the findings showed that mean math scores on a number of subtests and composites also improved significantly after participating in the intervention. At an individual level, 85.7% of children improved by at least one standard deviation or more in at least one math content area; more than half (57.1%) improved by the same margin in two or more math content areas, 28.6% improved in three or more areas, and at the highest end, two participants (14.3%) improved by at least one standard deviation in four math content areas. These findings were extremely promising given that this was a small pilot study conducted with a very diverse group. Although the results need to be considered in the context of the limited power and generalizability that the small sample size provides, the findings do suggest that the modified MILE program affected meaningful change on these participants. In addition to this, results showed that the math score improvements that had been observed were maintained six months after completing the intervention. This was another very positive finding as a math intervention that imparts only short-term change would provide somewhat misleading results about the program's true effectiveness and would likely not be worth a considerable investment of the limited resources available to most schools.

Characteristics Associated with Response to Intervention

The third research question posed in this study explored the relationship between participant characteristics and the magnitude of math improvement on the four main *KeyMath-3 DA* composites. Findings from a previous MILE study in Atlanta indicated that younger children were more likely to experience math score gains after receiving the intervention

(Coles et al., 2009), and in line with this, the current study found a strong, significant negative correlation between age and math score gains on the Basic Concepts composite.

This makes sense given the content emphasized in this composite. Basic Concepts covers the foundational knowledge (such as number awareness) that is essential to solving and understanding later math problems. If a student has already mastered these basic skills, they may not experience substantial additional growth in this domain. As younger children are less likely to have mastered these abilities, the MILE curriculum would likely suggest they spend more time building understanding in this area and therefore their progress may be more noticeable.

When looking at the relationship between age and the other composite score gains, an interesting pattern emerged that further supported the logic of this finding. Although no other correlations were significant, the coefficient transformed from strongly negative to moderately positive as the math learning continuum progressed. That is, as the skills became more advanced, greater gains were associated with older age. Again, this makes sense as older children are more likely to have already mastered foundational math skills and therefore spent more time developing their abilities in more complex domains.

The second significant correlation found in this study existed between math score change and baseline cognitive functioning. The moderate positive correlation indicated that higher math score gains were positively correlated with *WRIT* General IQ scores. There are two possible reasons for this.

First, it is very possible that behaviour contributed (both positively and negatively) to performance on the IQ and math measures *and* learning during the intervention sessions. Recall that 10 out of 14 participants were enrolled in a BLAP classroom, which by eligibility criteria necessitates that children display “severe emotional and behavioural challenges” (Edmonton Public Schools, 2013). For the participants experiencing the most significant behavioural concerns, their ability to remain attentive and compliant during testing and learning sessions was certainly compromised; this may have caused either additional ‘noise’ on the testing protocols that obscured their scores or a reduced ability to benefit from the intervention. The opposite of this is also true: the children whose behaviour interfered less with their testing also likely interfered less with their learning, enabling them to score higher in both areas and ultimately leading to this correlation.

Second, it is also possible that the children who saw greater math gains actually did have a higher cognitive capacity that was accessed through the intervention. For these students, traditional classroom math instruction may not have been sufficient and the individualized support provided by the modified MILE intervention may have allowed them to ‘bridge the gap’ and achieve greater gains that children with less cognitive capacity to learn the content and benefit from the strategies taught.

Program Feasibility

Given that this was a pilot study exploring new ways of implementing the MILE intervention, the fourth and final research question centered around whether the program was feasible with these modifications in place. Although most instructors identified notable program strengths and felt they learned something new from taking part in this study, they also largely reported that the intervention was difficult to manage on top of their already full schedules. Almost half of interventionists indicated they would not be willing to complete the program again as part of their regular teaching duties without additional support (e.g., another classroom EA).

Additionally, some of the negative feedback stemmed from the difficulties interventionists experienced engaging students with severe behavioural concerns. This is a very important obstacle to find a way to navigate as research suggests that even outside of BLAP classrooms behavioural difficulties are not uncommon in children with math learning challenges. Studies show that children as young as second and third grade who are struggling with math learning display higher levels of attentional difficulty and domain-specific math anxiety (Wu, Willcutt, Escovar, & Menon, 2014) and are at higher risk of experiencing other internalizing symptomology such as depression (Graefen, Kohn, Wyschkon, & Esser, 2015). In young children in particular, such symptoms can manifest themselves through a variety of behaviours that interfere with learning, and therefore it is likely that at least some children participating in a math intervention will exhibit additional behavioural challenges that put a strain on educators. It is clear that in order for children to benefit from positive math gains noted in this study, feasibility for the educators who are to administer the program will need to continue to be addressed.

Limitations

Although the academic results of this intervention appear very promising, the study also had a number of limitations. First and foremost was the small sample size. Initially, exactly twice as many participants consented to participate in the study; however, for various reasons, only 14 participants completed the intervention. This severely limited the power of the study and the generalizability of the results, and thus any conclusions made are inevitably tentative. This is particularly true when attempting to interpret the results of the intervention vs. waitlist and post-testing vs. 6-month follow up comparisons as each group in these scenarios had eight or fewer participants.

Next, given the small sample size, it is possible that the results were impacted by the lack of Time 2 *KeyMath-3 DA* data for two waitlist control group participants. This was a particularly unfortunate time point to be missing data as it meant that neither of the change scores we were interested in could be calculated for this group; neither Time 1 to Time 2 control period changes nor Time 2 to Time 3 intervention effect changes were available. Some discussion took place within the research team about potentially using these students' Time 1 data as a rough estimate of their baseline score in order to allow us to still calculate an *approximate* intervention effect (i.e., use Time 1 data as pseudo-Time 2 data) as the mean score change during the control period was non-significant. However, as we couldn't completely rule out the possibility that significant growth could have occurred in these participants during the control period as a result of classroom instruction or maturation, we opted not to do so which further reduced our sample size for these analyses.

Finally, as mentioned in other areas, behaviour considerably inhibited testing and learning for a number of participants involved in this study. Although any test scores that were invalidated were noted as such and not used in any analyses, it was sometimes difficult to determine how much a child's scores may have been impacted by their emotional and behavioural difficulties on any given day. In such cases, although these participants were well within the inclusion criteria we originally intended, their ability to benefit from a *math* intervention may have been overshadowed by more urgent needs. For these students, it may have been more beneficial to address overall learning readiness and wellbeing before focusing on math abilities.

Implications

The current pilot study contributes important information to the literature on early math intervention as it demonstrates that students quantifiably benefitted from participating in a modified MILE intervention. Not only this, but it contributes to the body of research on the MILE program itself by demonstrating that the modified program can affect substantial change in the math achievement scores of children with a wide range of neurodevelopmental and learning difficulties, a diverse population in which its impact had not previously been established. However, a third outcome of this study was the finding that educators and EAs had significant reservations about the feasibility of the current intervention. Given the cascading effect of math difficulties and the unfortunate outcomes associated with poor math competence, it is clear that effective early interventions must be developed and implemented for children who are struggling with math learning. As the modified MILE program showed a number of otherwise very promising results, ways to further adapt the program to make delivery more feasible should be explored in future studies.

References

Abrantes, P. (2001). Mathematical competence for all: Options, implications and obstacles. *Educational Studies in Mathematics, 47*(2), 125-143.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Ancker, J. S., & Kaufman, D. (2007). Rethinking Health Numeracy: A Multidisciplinary Literature Review . *Journal of the American Medical Informatics Association, 14*(6), 713-721.

Assel, M. A., Landry, S. H., Swank, P., Smith, K. E., & Steelman, L. M. (2003). Precursors to mathematical skills: Examining the roles of visual-spatial skills, executive processes, and parenting factors. *Applied Developmental Science, 7*(1), 27-38.

Bertrand, J. (2009). Interventions for children with fetal alcohol spectrum disorders (FASDs): Overview of findings for five innovative research projects. *Research in Developmental Disabilities, 30*(5), 986-1006.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*(3), 273-293.

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review, 28*(4), 415-427.

Coles, C., Kable, J., & Taddeo, E. (2009). Math performance and behavior problems in children affected by prenatal alcohol exposure: Intervention and follow-up. *Journal of Developmental and Behavioral Pediatrics, 30*, 7-15.

Connolly, A. (2008). *KeyMath-3 Diagnostic Assessment: Manual Forms A and B*. Toronto, ON: Pearson Canada Assessment.

Domazet, M., Baranović, B., & Matić, J. (2013). Mathematics competence and international mathematics testing: Croatian starting point. *Journal for Spatial and Socio-Cultural Development Studies, 51*(1), 109-131.

Drummond, R. J., Sheperis, C. J., & Jones, K. D. (2016). *Assessment procedures for counselors and helping professionals* (8th ed.). Boston, MA: Pearson.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . & Sexton, H. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1466.

Edmonton Public Schools. (2013). *Behaviour and Learning Assistance: Program Description*. Retrieved from Edmonton Public Schools: https://www.epsb.ca/programs/special education/behaviourandlearningassistance/

Erford, B. T. (2013). *Assessment for counselors* (2nd ed.). Belmont, CA: Cengage Wadsworth.

European Communities. (2007). *Key Competences for Lifelong Learning: European Reference Framework.* Luxembourg: Office for Official Publications of the European Communities.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*(1), 4-15.

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PloS One, 8*(1), e54651.

Glutting, J., Adams, W., & Shelow, D. (2000). *Wide Range Intelligence Test.* Lutz, FL: Psychological Assessment Resources Inc.

Graefen, J., Kohn, J., Wyschkon, A., & Esser, G. (2015). Internalizing Problems in Children and Adolescents With Math Disability. *Zeitschrift für Psychologie, 223*(2), 93-101.

Jordan, N. C., & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews, 15*(1), 60-68.

Kable, J. A., Taddeo, E., Strickland, D., & Coles, C. D. (2015). Community translation of the Math Interactive Learning Experience Program for children with FASD. *Research in Developmental Disabilities, 39*, 1-11.

Kable, J., Coles, C., & Taddeo, E. (2007). Socio-cognitive habilitation using the math interactive learning experience program for alcohol-affected children. *Alcoholism - Clinical and Experimental Research, 31*, 1425-1434.

Kully-Martens, K., Pei, J., Kable, J., Coles, C. D., Andrew, G., & Rasmussen, C. (2018). Mathematics intervention for children with fetal alcohol spectrum disorder: A replication and extension of the math interactive learning experience (MILE) program. *Research in Developmental Disabilities, 78*, 55-65.

Niss, M. (2004). The Danish KOM project and possible consequences for teacher education. In R. Strässer, G. Brandell, B. Grevholm, & O. Helenius (Ed.), *Educating for the Future: Proceedings of an International Symposium on Mathematics Teacher Education* (pp. 179-190). Göteborg, Sweden: The Royal Swedish Academy of Science.

Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?* London: National Research and Development Centre for Adult Literacy.

Peng, P., Wang, C., & Namkung, J. (2018). Understanding the cognition related to mathematics difficulties: A meta-analysis on the cognitive deficit profiles and the bottleneck theory. *Review of Educational Research, 88*(3), 434-476.

Perrenoud, P. (1997). *Construire des compétences dès lécole.* Paris, FR: ESF éditeur.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*(2), 110-122.

Rose, H., & Betts, J. R. (2004). The effect of high school courses on earnings. *Review of Economics and Statistics, 86*(2), 497-513.

Schweinhart, L. J. (2006). The High/Scope approach: Evidence that participatory learning in early childhood contributes to human development. In N. F. Watt, C. Ayoub, R. H. Bradley, J. E. Puma, & W. A. LeBoeuf (Eds.), *The crisis in youth mental health: Critical issues and effective programs* (Vol. 4, pp. 207-228). Westport, CT: Praeger Publishers/Greenwood Publishing Group.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59*(4), 745-759.

Swank, J. M., & Mullen, P. R. (2017). Evaluating evidence for conceptually related constructs using bivariate correlations. *Measurement and Evaluation in Counseling and Development, 50*(4), 270-274.

The Learning Bar. (2019). *Early Years Evaluation: Teacher Assessment (EYE-TA)*. Retrieved from Early Years Evaluation: https://www.earlyyearsevaluation.com/index.php/en/products/eye-ta

Turner, R. (2010). Exploring mathematical competencies. *Research Develoments, 24*(5), 2-7.

Vorobjovs, A. (2017). Adolescents' mathematical competence formation influencing factors. *Rural Environment, Education, Personality (REEP): Proceedings of the 10th International Scientific Conference.* Jelgava, Latvia: Latvia University of Agriculture.

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. . *Educational Researcher, 43*(7), 352-360.

Wu, S. S., Willcutt, E. G., Escovar, E., & Menon, V. (2014). Mathematics achievement and anxiety and their relation to internalizing and externalizing behaviors. *Journal of Learning Disabilities, 47*(6), 503-514.

Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. . *Learning and Individual Differences, 23*, 1-9.

Appendix A

MILE Instructor Satisfaction Evaluation Form
Part I

	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I enjoyed completing the MILE program					
I learned something new from the MILE program					
The intervention was beneficial for the students					
I would be willing to complete the program again as part of a research study					
I would be willing to complete the program again as part of my regular duties at work					
The intervention program would be beneficial if it was included in the regular curriculum					
I was able to choose activities/curriculum pages that applied to the needs/skill level of all children in my group					

Part II

1. Regarding the number of sessions in the intervention:
   - There were not enough sessions
   - The number of sessions was about right
   - There were too many sessions
2. Regarding the length of each session:
   - The sessions should have been shorter
   - The sessions were the right length
   - The sessions should have been longer
3. On average, approximately how much time did you spend planning for each session?
   - 15 minutes or less
   - 16-30 minutes
   - 31-45 minutes
   - 46-60 minutes
   - Over an hour

Consider your answer to #3 above for questions 4 and 5.

4. Did you feel that you were able to plan a well-organized session in this amount of time?
   - Yes
   - Somewhat
   - No
5. Did you feel that this amount of planning time was a burden to your schedule?
   - Yes
   - Somewhat
   - No

Part III

How many children were in your intervention group?

What did you like about the intervention?

What did you dislike about the intervention?

If MILE were to be implemented into the curriculum, are there any changes you would like to make to the intervention?

Other comments?