An Analysis of the Mathematical Impacts of a Cross-Age Peer Tutoring Initiative

An Analysis of the Mathematical Impacts of a Cross-Age Peer Tutoring Initiative

Moshe Ohayon

A Thesis in the Field of Mathematics for Teaching

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

May 2018

Abstract

Cross-age peer tutoring is an instructional strategy in which older students serve as interventionists to improve the academic and attitudinal outcomes of their younger peers. Educational Justice, a nonprofit organization based in Louisville, Kentucky, is piloting an initiative called Educational Justice Activists (EJA) that implements this instructional strategy by pairing high-performing 9^{th}-12^{th} grade students with underserved 5^{th}-8^{th} grade students for weekly one-on-one tutoring sessions. By evaluating the EJA model's impact on tutees participating in the initiative, this study seeks to identify academic performance trends in mathematics sections of standardized exams as well as shifts in attitudes towards school and learning, and, in particular, towards mathematics. According to the study, the effect sizes (ES) indicate that it is possible to improve both academic and non-academic outcomes among low-income 5^{th}-8^{th} graders using the cross-age peer tutoring intervention model of EJA. The analysis of academic assessments included 56 tutees and revealed gains in addition/subtraction (ES = 0.13) and multiplication/division (ES = 0.30). Survey data analyzed for attitudinal shifts among 44 tutees also showed effect size gains ranging from 0.14 to 0.43. The study further explored trends along two mediating factors: tutoring dosage and tutee grade. It was found that, for addition/subtraction, higher dosage (more than 13 tutoring hours) resulted in greater gains (ES = 0.20) than lower dosage (ES = 0.04); however, for multiplication/division, lower dosage (13 or fewer tutoring hours) resulted in greater gains (ES = 0.52) than lower dosage (ES = 0.32). Similarly, it was found that, for addition/subtraction, older tutees (7^{th}-8^{th} graders) showed greater gains (ES = 0.19) than younger ones (ES = 0.08);

however, for multiplication/division, the younger students (5^{th}-6^{th} graders) displayed greater gains (ES = 0.42) than their older peers (ES = 0.21). The attitudinal survey results with respect to these two mediating factors were mixed.

Dedication

To my parents – Fany and Israel – whose words of inspiration and wisdom sustain me

And to my children – Sharōn, Avi, and Rachel – whose light continuously fills my world

two generations of unbounded love and support,

without which none of this would have been possible.

Acknowledgments

Many thanks to Dusty Grundmeier for his incredible patience with me throughout this project. Your tremendous insights and unwavering encouragement have made all the difference.

I also wish to thank Andy Engelward, who has been a trusted teacher, mentor, and friend throughout my time at the Extension School.

Finally, I owe a debt of gratitude to Hannah and Katherine at Educational Justice who are forever in my corner. I cannot thank you enough.

Table of Contents

List of Tables

Chapter 1 Introduction

When it comes to academic performance, the current income achievement gap is now nearly twice that of the black-white achievement gap: students from low-income families performing about 3-6 grade levels (1.25 standard deviations) below their peers from high-income backgrounds, and the current research indicates that this gap is widening (Reardon, 2013). The significance of the dilemma recently became even more apparent when data collected by the National Center for Education Statistics revealed that the growing socioeconomic divide among K-12 schoolchildren crossed an alarming milestone in 2013: now, a majority of the nation's public school students are low-income, according to the Southern Education Foundation (2015). Specifically, in 2013, more than 51 percent of students attending public school in the U.S. -- over 25 million children -- lived in low-income households (Southern Education Foundation, 2015). In a report issued by the Southern Education Foundation (2015), Steve Suitts writes:

> No longer can we consider the problems and needs of low income students simply a matter of fairness… Their success or failure in the public schools will determine the entire body of human capital and educational potential that the nation will possess in the future. Without improving the educational support that the nation provides its low income students – students with the largest needs and usually with the least support -- the trends of the last decade will be prologue for a nation not at risk, but a nation in decline…

Educational Justice (EJ), a nonprofit organization based in Louisville, Kentucky, is attempting to address the income achievement gap by leveraging the largely untapped resource of service-oriented, high achieving high school students. The nonprofit organization does so by attracting high-achieving high school students to apply to a new kind of honor society, called *Educational Justice Activists* (EJA), in which applicants are not only selected based on their academic records and accomplishments but, once accepted, are also required to apply their academic expertise as tutors for underserved younger peers. Selected high school students are first trained by EJ's educators – teachers and/or tutors with at least three years of teaching experience in various academic subjects, including mathematics – to serve as tutors called *EJ Activists*. At the same time, EJ identifies low-income 5^{th} to 8^{th} grade students in critical need of academic support and invite their parents/guardians to enroll those students in the program as *EJ Achievers*. Each EJ Achiever is then able to browse through the secure database of Activist profiles to select a single high school student to be his/her dedicated tutor for one-on-one tutoring sessions every week until the end of the school year. EJ Achievers may continue to be paired with their Activists until the end of 8^{th}-grade, i.e., for a maximum of four years.

This study examines academic performance trends in mathematics among Achievers participating in EJA. More specifically, it seeks to determine whether, among tutees in the EJA initiative, trends can be identified in test scores earned in mathematics sections on standardized exams. In addition, using entrance and exit survey data, the study seeks to determine whether trends arise among EJA students in attitudinal shifts towards school and learning, and, in particular, towards mathematics.

It should be noted that the author of the present study, Moshe Ohayon, is also the founder and current executive director of Educational Justice, the nonprofit organization piloting the EJ Activists initiative. In light of this connection, the author has taken great care to maintain a neutral perspective in regards to the data collected and analyzed in this study.

Chapter 2 Background

A large body of research supports the effectiveness of peer tutoring interventions. Such programs are often described as those in which "student interventionists work to improve their peers' academic skills by providing one-on-one or small group academic instruction" (Collins, Hawkins, & Flowers, 2017). It is often impractical to expect a single teacher to lead and manage the instructional progress of an entire class of students throughout the school day while at the same time providing individual attention to one or more students who require an intervention (Collins et al., 2017). As such, and in view of the research findings on peer tutoring, programs that train students to implement such interventions for their peers may appeal to educators, school administrators, policy makers in the education sector, and others, but a number of variations on the peer tutoring model have been explored with a variety of differing objectives, results, and implications.

As a form of intervention, peer tutoring has increasingly been implemented as a preferred research-based approach for creating the sort of environment that assists students in mastering academic content as well the skills necessary to do so (Bowman-Perrott, deMarín, Mahadevan, & Etchells, 2016). In fact, a few of the key features of peer tutoring that have been identified across implemented variations of the intervention model include frequent opportunities to respond, increased time on-task, opportunities to practice academic content, and regular and immediate feedback (Bowman-Perrott at al., 2016). Maheady, Harper, and Sacca (1988) found that each of these features is correlated with an increase in academic achievement (Bowman-Perrott, et al., 2016). In addition, Greenwood, Terry, Arreaga-Mayer, and Finney (1992) identified a number of benefits that result from peer tutoring interventions, and these have been summarized by Bowman-Perrott et al. (2016) as including the following:

> (a) being paired with a peer partner for one-to-one instruction, (b) opportunities for error correction, (c) increased time spent on academic behaviors, (d) increased positive social interactions between students, (e) a decrease in off-task and disruptive behaviors, and (f) experiencing more success and report feeling more confident academically. (p. 360)

Peer tutoring interventions have been studied with respect to a host of mediating factors and settings. With respect to content areas, for example, peer tutoring interventions have yielded promising learning outcomes in mathematics (Fantuzzo, Polite, & Grayson, 1990; Harper, Mallette, Maheady, Bentley, & Moore, 1995; Robinson, Schofield, & Steers-Wentzell, 2005; Hawkins, Musti-Rao, Hughes, Berry, & McGuire, 2009), as well as across a number of other core academic subjects, including reading (Houghton & Bain, 1993; Elbaum, Vaughn, Hughes, & Watson Moody, 2000; Oddo, Barnett, Hawkins, & Musti-Rao, 2010), science (Bowman-Perrott, Greenwood, & Tapia, 2007; Kamps et al., 2008), and social diet it comes to student population types, studies demonstrate the effectiveness of peer tutoring on tutees of a various ages and grade levels from elementary through high school (Bowman-Perrott et al., 2007; Dietrichson, Bøg, Filges, & Klint Jørgensen, 2017; Kamps et al., 2008). In addition, studies reveal that peer tutoring is effective among tutees with learning differences or disabilities (Kunsch, Jitendra, & Sood, 2007; Okilwa & Shelby, 2010) as well as those who are non-native speakers of the language of instruction (Bowman-Perrott et al., 2016; Cole 2013, 2014).

Such interventions have even shown that they yield benefits for the tutors themselves (Cohen, Kulik, & Kulik, 1982; Collins et al., 2017; Galbraith & Winterbottom, 2011).

Overall, research on peer tutoring interventions is often assessed along one or both of two primary focuses: (1) academic achievement as measured by performance on examinations, typically in mathematics or reading; and (2) non-academic outcomes, which often include the evaluation of affective (attitudinal or emotional) parameters -- such as favorability towards the subject matter covered in tutoring sessions, shifts in attitudes towards school and learning, and modification to self-concept -- as well as social or behavioral ones. By aggregating and evaluating numerous research findings, meta-analytic studies on peer tutoring allow for the assessment of such interventions across a variety of mediating factors and settings. The majority of these studies, however, concentrate on studies of academic outcomes, with only a relatively small number of meta-analyses focusing on evaluating studies of nonacademic parameters (e.g., on-task behaviors, social interactions), self-concept, or attitudinal measures.

Quantitative Meta-analyses

Between 1982 and 2017, 12 major meta-analytic studies have been conducted on peer tutoring programs and interventions in which a range of tutoring formats -- including cross-age, class-wide, and small group arrangements -- were represented across a variety of academic subjects (Table 1). These meta-analyses attempted to evaluate and quantify the extent to which peer-mediated tutoring interventions affected learning outcomes among tutees involved in such programs in comparison to students assigned to control groups when a group research design was employed, or in comparison to each tutee's previous performance when a single-case research design was used. A review of the findings reveals that peer tutoring generally produced positive gains for participating students. However, even among meta-analytic studies, there are conflicting conclusions with regard to the effect of individual features of different peer tutoring intervention designs.

Table 1. Major Meta-Analyses on Group Tutoring Interventions

Meta-Analysis	Year of Publication	Number of Studies
Cohen et al.	1982	65
Cook et al.	1985	19
Rohrbeck et al.	2003	90
Ginsburg-Block et al.	2006	36
Jun et al.	2010	12
Bowman-Perrott at al.	2013	26
Cole et al.	2013	28
Bowman-Perrott at al.	2014	20
Cole et al.	2014	28
Leung	2014	72
Rees et al.	2015	10
Zaneli et al.	2016	49
Dietrichson et al.	2017	101

Academic Performance

When evaluating studies of different interventions and their possible effect on academic performance among students in the general K-12 population, five of the twelve meta-analyses reported, to some extent, effect sizes for two central mediating factors: (1) dosage, which quantifies the total amount of tutoring hours received by the tutee, commonly defined formulaically as follows: Dosage = intervention duration (in weeks) X frequency (sessions/week) X session length (hours/session); and (2) the grade and/or age of the tutees receiving the intervention (Table 2).

With respect to dosage, the meta-analyses appear to be in agreement that, perhaps contrary to a greater dosage of tutoring does not necessarily equate to improved results. In fact, interventions logging fewer overall tutoring hours were equally or slightly more effective in producing improved academic results than those offering a somewhat greater dosage of tutoring time. Even for the meta-analysis performed by Cohen et al. (1982), for which dosage could not be calculated (as it did not include information about tutoring frequency but rather only provided effect sizes for interventions studies spanning three different ranges of durations in weeks), the trend appears to indicate that dosage and effect size are inversely correlated -- that is, at least over the span evaluated (interventions spanning no more 36 weeks).

The meta-analyses, however, displayed less agreement with respect to which ages or grades of tutees responded more favorably to peer tutoring interventions. Two of the meta-analyses -- Cohen et al. (1982) and Rohrbeck et al. (2003) -- indicated that younger elementary school students (grades 1-3) achieved greater gains while the other studies -- Jun et al. (2010), Bowman-Perrott (2013), and Leung (2014) -- pointed to greater gains among older students of middle to high school ages.The meta-analyses, however, displayed less agreement with respect to which ages or grades of tutees responded more favorably to peer tutoring interventions.

Table 2. Reported Effects Sizes (ES) of Academic Performance by Tutoring Dosage and Tutee Grade/Age

Meta-analysis	No. Studies	Tutoring dosage	ES	Tutee grade/age	ES
Cohen et al. (1982)	65	Low (0-4 weeks)*	0.95	Grade 1-3	0.45
		Medium (5-18 weeks)*	0.42	Grade 4-6	0.25
		High (19-36 weeks)*	0.16	Grade 7-9	0.33
Rohrbeck et al. (2003)	90	Low (≤ 19 hours)	0.38	Grade 1-3	0.37
		High (> 19 hours)	0.32	Grade 4-6	0.28
Jun et al. (2010)	12	Low (≤ 7 hours)	0.24	Grade 6-7	0.60
		Medium (8-15 hours)	0.20	Grade 8-9	0.18
		High (≥ 16 hours)	0.66	Grade 10+	0.90
Bowman-Perrott et al. (2013)	26	Low (≤ 6 hours)	0.75	Elementary	0.69
		High (> 6 hours)	0.75	Middle/high	0.74
Leung (2014)	72	Low (≤ 16.25 hours)	0.44	Elementary	0.34
		High (> 16.25 hours)	0.40	Middle/high	0.52

* frequency unavailable   ** average of adult tutoring (0.70) and computer-based tutoring (0.19)

Two of the meta-analyses -- Cohen et al. (1982) and Rohrbeck et al. (2003) -- indicated that younger elementary school students (grades 1-3) achieved greater gains while the other three studies -- Jun et al. (2010), Bowman-Perrott (2013), and Leung (2014) -- pointed to greater gains among older students of middle to high school ages. In the meta-analysis conducted by Jun et al. (2010), the reasons behind the precipitous drop in effect size to 0.18 among students of grade 8-9 were unclear. The researchers believe it may result from the format and insufficient sample size of the data (of the twelve studies in the meta-analysis, the bulk of the data analyzed for students in grades 8-9 originated from only two of them, both of which employed computer-based tutoring) but cite the "maturational and psycho-emotional" challenges faced particularly by students in this age range as a less likely explanation for the apparent temporary decline in the effect of the interventions.

The number of studies included in each meta-analysis have been identified in Table 1 above. The apparent disagreement among meta-analyses may be at least partially attributed to size: large meta-analytic studies (those including data from 50 studies or more) have the advantage of including more data but are generally less selective than smaller ones that often implement rigorous inclusion criteria, such as allowing only peer-reviewed publications.

In addition, all five meta-analyses described a number of mediating factors (such as format, incentives provided, etc.) in an effort to evaluate the effect each variable's effect on academic gains among tutees participating in peer tutoring interventions. These variables included determining the effect when (1) the tutoring is conducted in a one-on-one format, in which one tutor is paired with one tutee, as opposed to small-group or classroom-wide tutoring settings; (2) the tutoring is cross-age, in which the tutor is at least one school grade above that of the tutee but is still a student (often just a few years older than the tutee), as opposed to same-age tutoring or tutoring provided by adult tutors; (3) the instructor is static, in which the tutee always receives instruction from the same tutor; (4) the tutoring is structured, in which tutoring is provided on a regular schedule, as opposed to on occasion as needed (e.g., help room tutoring); (5) different academic content areas, such as math and reading, are compared; (6) training is provided to the tutor before and/or throughout the peer tutoring intervention; and (7) the tutee is provided with some kind of reward to incentivize learning in the intervention.

Non-Academic Outcomes

In evaluating non-academic outcomes, three of the twelve meta-analytic studies reported effect sizes (see Table 3a) for non-academic outcomes among tutees in peer tutoring interventions. The meta-analyses evaluated non-academic outcomes of peer tutoring based on multiple studies focusing on one or more of the following parameters: (1) social (2) behavioral (3) self-concept and (4) attitudinal/emotional. Social outcomes included assessment of social competencies, such as frequency and ease of interaction with peers and adults, cooperative and conflict resolution skills, sociometrics, and the like. Behavioral outcomes included, for example, extent of participation and effort, compliance with rules, and ability to remain on-task. Self-concept refers to the tutee's feelings about herself/himself, which includes general self-confidence as well as academic self-perceptions, such as competence in a particular academic content area. Lastly, other attitudinal/emotional responses (such as attitudes toward learning, academic achievement, and the academic topics covered in the intervention's tutoring sessions) were excluded from consideration in recent meta-analyses due to their small base rate but were reported briefly in Cohen et al. (1982) for positive attitude shifts regarding the subject matter tutored. Data was gathered using various approaches depending on the study within each meta-analysis but generally relied on informants (the students themselves, peers, teachers, parents, and/or researchers) to produce outcome measures via direct observations, questionnaires, and surveys with ratings scales. Table 3a summarizes the results of these meta-analyses -- Cohen et al. (1982); Ginsburg-Block, Rohrbeck, & Fantuzzo (2006); and Bowman-Perrott, Burke, Zhang, & Zaini (2014) -- along reported non-academic parameters and the number of studies upon which the calculated effect size relied for the given parameter.

Table 3a. Overall Reported Effects Sizes (ES) for Non-academic Outcomes

Meta-analysis	No. Studies	Non-Academic Outcome	ES
Cohen et. al (1982)	8	Attitude toward Subject Matter	0.29
	9	Self-concept	0.09
Ginsberg-Block et al. (2006)	30	Social	0.52
	15	Self-concept	0.40
	12	Behavioral	0.65
Bowman-Perrott et al. (2014)	20	Overall behavioral and social	0.62
	11	- improving social interaction/skills	0.69
	10	- reducing disruptive/off-task behavior	0.60
	6	- improving academic engagement	0.38

Bowman-Perrott et al. (2014) specifically analyzed 20 studies that employed single-case research designs and included tutees in grades preK-12. The meta-analysis reported an overall TauU weighted effect size of 0.62 on behavioral and social outcomes among tutees in peer tutoring interventions. With respect to outcomes in specific subcategories, Bowman-Perrott et al. (2014) reported that peer tutoring yielded greater effects when it came to improving skills related to social situations and and interactions (ES = 0.69) and or curtailing off-task or disruptive behaviors (ES = 0.60), while the effect on improving a tutee's academic engagement in school (ES = 0.38) was more moderate.

Of the three meta-analyses, only Ginsberg-Block et al. (2006) provided a breakdown along the two central mediating factors identified above for academic performance: (1) dosage, which quantifies the total amount of tutoring hours received by the tutee, commonly defined formulaically as: Dosage = intervention duration (in weeks) X frequency (sessions/week) X session length (hours/session); and (2) the grade and/or age of the tutees receiving the intervention (Table 3b).

Table 3b. Effects Sizes (ES) of Non-academic Outcomes by Tutoring Dosage and Tutee Grade/Age as Reported by Ginsberg-Block et al. (2006)

Outcome	Tutoring dosage	No. Studies	ES	Tutee grade/age	No. Studies	ES
Social	Low (≤ 15 hours)	12	0.38	Grade 1-3	6	0.35
	High (> 15 hours)	11	0.26	Grade 4-6	24	0.28
Self-concept	Low (≤ 15 hours)	7	0.35	Grade 1-3	1	1.20
	High (> 15 hours)	3	0.27	Grade 4-6	14	0.17
Behavioral	Low (≤ 15 hours)	3	0.99	Grade 1-3	4	0.49
	High (> 15 hours)	4	0.68	Grade 4-6	8	0.43

Chapter 3 Methods

This study explores the effects of *Educational Justice Activists* (EJA), a one-on-one cross-age peer tutoring program being piloted by the nonprofit organization Educational Justice in Louisville, Kentucky. The program is intended to address the educational inequity behind the opportunity gap between students from low-income backgrounds and their more privileged peers. By evaluating whether the the EJA model produces positive learning outcomes among student participants, and if so, the extent thereof, this study is intended to contribute to the body of knowledge in the field of education, specifically in the area of cross-age peer tutoring interventions, and in particular, with respect to academic interventions offered to students from low socioeconomic backgrounds.

Program Description

In the EJA program, academically high achieving high school students in grades 9-12 apply to participate as tutors. Each applicant accepted is trained to serve as a tutor and academic coach ("Activist") for a younger peer ("Achiever"). All Achievers are students in grades 5-8 from low-income households. Activist-Achiever pairs meet weekly for a minimum of one hour of academic tutoring until the end of the school year. The program allows each Achiever to continue in the program for multiple school years until he/she completes the eighth grade (i.e., up to four school years). Every attempt is made to keep an Achiever paired with the same Activist over multiple years. The pilot program is currently in its third academic year, and a number of Activists and Achievers have participated over multiple years, with the central goal of the program being to maintain the original Activist-Achiever pairing across multiple years whenever possible. As part of the EJA program, weekly tutoring sessions are required during the academic year and may even continue, albeit are optional, during the summer months.

Activists (tutors)

Academically high-performing high school students submit an online application to become “EJ Activists.” The program, structured as a selective honor society, requires that each applicant submit a number of items in order to be considered for admission into the program, including a full academic transcript, a letter of recommendation from a school teacher/counselor, and brief essays describing their motivation for applying to the program. Educational Justice staff members review each application and invite the most promising candidates to an interview. The staff members then admit those applicants they deem most qualified into the program as EJ Activists.

Before tutoring begins, new Activists must attend a preliminary five-hour training, in which Educational Justice staff members who are experienced educators (professional tutors and certified teachers) cover four central topics: (1) the scope of educational inequity on the local and national levels and to connection to the history and mission of EJA; (2) program requirements and guidelines for Activists regarding appropriate conduct and sensitivity when tutoring an Achiever; (3) guidance on how to effectively tutor and academically mentor an Achiever; and (4) program requirements, protocols, and guidelines for Activists regarding scheduling sessions and communicating with Achievers, their families, and EJA staff members (including appropriate use of EJA's secure web application).

Once new Activists complete the preliminary training, they are provided access to a secure online account where they (1) create a profile, which includes a photo and some personalized information (e.g., favorite and least favorite subject in school, other likes and dislikes); and (2) define their tutoring availability, which includes selecting one-hour slots from a list of times, days, and pre-approved sites (e.g., public libraries and community centers throughout the city) where tutoring may be conducted. To address safety and liability concerns, the pre-approved tutoring sites include an open, shared space where all Activist-Achiever pairs meet under staff supervision. After submitting their online availability, Activists must then wait to be selected by Achievers, who are granted access to all of the unpaired Activist profiles with matching availability. Once an Achiever chooses an Activist, the pair begin to meet at the selected day, time, and location for a minimum of one hour each week. In accordance with the design of the program, the pairings are static and exclusive: Activists may not tutor more than one Achiever, and Achievers may not be tutored by more than one Activist.

Once paired, each Activist is required to attend a minimum of four additional additional training sessions with EJA staff. In these small-group (limited to a maximum of 12 Activists) development workshops, trainers guide and support Activists with any challenges they face in tutoring their Achievers (e.g., by suggesting educational enrichment activities, resources, and lesson plans that can be implemented by Activists during tutoring sessions with their Achievers or techniques for helping an Achiever stay engaged during sessions) as well as gather feedback from Activists on the progress of each pairings.

Achievers (tutees)

As the EJA model was designed to provide educational support to students from low socioeconomic backgrounds, there is no cost to participate as an Achiever. Parents/guardians interested in having their children become Achievers in the program are required to apply for one of a limited number of scholarships available each year. To qualify for a full scholarship and, thus, for participation in EJA, each parent/guardian must (1) demonstrate financial need (i.e., low-income status) by providing appropriate documentation (e.g., designation of free/reduced lunch status by the local school district, receipt of government aid benefits); (2) provide documentation that the child is currently in 5th, 6th, 7th, or 8th grade; and (3) must agree to the terms of participation in the program (e.g., commit to weekly tutoring sessions, reschedule any missed sessions, arrange for the child's transportation to/from tutoring site every week).

A parent/guardian who applies and meets the three qualifications above is then invited to attend a registration session, along with his/her child, in order to learn about the requirements of participation and complete enrollment paperwork. In addition, at the registration session, the new Achiever is given the opportunity to browse through profiles of unpaired Activists and to choose his/her tutor. One-on-one weekly tutoring sessions between the pair typically begin within two weeks. Every pairing is exclusive: once paired, the Activist cannot tutor any other Achiever, and the Achiever may not be tutored by any other Activist.

EJA program design

Tutoring is always one-on-one but must always occur in an open space and during predetermined hours and locations, where an EJA staff member is present to supervise and support the sessions. If no tutors are available, the Achiever is placed on a waitlist, and the parent/guardian is notified once additional tutors become available. In the meantime, waitlisted Achievers are provided access to a group tutoring help room offered every weekend by Educational Justice.

To evaluate the effectiveness of the EJA program, Educational Justice collects and analyzes data from five sources: (1) global metrics from the EJA software platform (e.g., number of paired/unpaired students in the program, attendance, punctuality); (2) weekly meeting and progress reports that every Activist completes following each tutoring session; (3) attitudinal and feedback surveys that Activists, Achievers, and Achiever parents/guardians complete upon entrance and exit from the program (intended to gauge potential shifts in attitudes towards school and learning, satisfaction with the program, and perceived progress to date); (4) data, though a partnership with the local school district, on the performance of Achievers and Activists in school, including grades, standardized test scores, and attendance; and (5) an internal administration, conducted by EJA staff members, of nationally-recognized standardized assessments in mathematics and reading to Achievers before pairing and at the end of the school year. The math and reading assessments used are, respectively, the *KeyMath-3* developed by Pearson and the *TORC-3* produced by Pro-Ed.

It should be noted that no set curriculum is employed by the EJA program. Instead, to take full advantage of the one-on-one format of the model, Activists are trained to respond and cater to the individual academic needs of their respective Achievers by (1) supporting their academic obligations and performance in school; and (2) providing academic remediation or enrichment in appropriate subject areas. For the former, Activists gather information from the Achiever, the Achiever's parent/guardian, and, in some cases, even the Achiever's school teacher; for the latter, EJA provides Activists as well as Achievers and their parents/guardians with the results of the math and reading assessments as a means to guide instruction. The objectives, course, and content of the tutoring sessions is, therefore, largely left to each Activist's discretion and is often decided on together with the Achiever's input, as well as, in some cases, with that of the parent/guardian. Activists, however, are provided with resources and support, such as access to a database of recommended and popular learning tools (e.g., worksheets, games, reading passages); regular training workshops, as previously described; and regular access to the EJA educational staff for guidance.

Present Study

No experimental research has yet been conducted on the educational impact of the EJA model. The present study is intended to examine trends among Achievers in their academic performance specifically in mathematics and in their attitudes toward school and learning after having participated in EJA. To do so, the academic performance of Achievers was evaluated by analyzing the data gathered by Educational Justice for Achievers who completed both the pre-assessment and post-assessment of the *KeyMath-3* exam. In addition, specific survey responses relating to attitudes toward math, learning, tutoring, and school was selected and analyzed for Achievers in the program who completed both the entrance and exit surveys in an effort to identify trends in non-academic outcomes.

Experimental Group

The evaluation sample consists of as many as 127 Achievers, each of whom is paired with an Activist for weekly tutoring sessions in the 2016-2017 academic year. Being that they are elementary and middle school students enrolled in grades 5-8, Achievers are generally 10-14 years of age, while Activists -- high school students enrolled in grades 9-12 -- are generally 14-18 years of age.

Control group

The expected control sample consists of 12 Achievers, all of whom are on the waitlist and, as such, are not paired with Activists. As such, Achievers in the control sample do not receive weekly one-on-one tutoring from an Activist. However, since they may attend the weekly group tutoring sessions offered by Educational Justice to all students on the waitlist, the frequency and duration of group tutoring attendance was tracked for Achievers in the control sample.

For clarification, EJA already gathers data on its participants for non-research purposes. The present study proposes to analyze the pre-existing data for research purposes. The division of participants into experimental and control groups only takes place within the data analysis portion of the study. In other words, participants are not physically divided for the purposes of the study, and, in fact, the EJA program is not altered in any way for the study. For the purposes of the study, it is not EJA students who are physically grouped or removed but rather only the pre-existing data relating to those students is manipulated. For example, in its analysis, the study does not consider the results of participants who do not meet the parameters of the research pre-analysis plan (e.g., participants who opted out of participation in the present research study). However, as means of assessing its own intervention programs, Educational Justice already administers the assessments and surveys described above and acquires its own consent forms from all of the students' parents/guardians.

For the present study, Educational Justice has agreed to share the data it collects on students whose parents/guardians also agree to participate in the study. In addition, written permission to access and use this data for the present study has been granted by Educational Justice.

Data Sources

To elaborate on the two primary data sources mentioned above, the present study examines data from the following:

1) Standardized Math Assessment

Educational Justice administers two content area tests of the *KeyMath-3 Diagnostic Assessment*, which is a norm-referenced measure of mathematical skills and is produced by Pearson, to all Achievers on two occasions throughout their participation in EJA: (i) first test date (pre-assessment) -- during predetermined testing days in January and February 2017 and (ii) second test date (post-assessment) -- near the end of the school year (Connolly, 2007).

The content areas of the two tests are (i) written computation in addition and subtraction and (ii) written computation in multiplication and division. These tests are available in two forms: Form A and Form B; the former was administered on the first occasion and the latter on the second occasion.

To make testing feasible for Achiever families, multiple testing dates were made available during the following testing windows:

First Test Date Options:	January 24, 2017 – February 28, 2017
Second Test Date Options:	May 1, 2017 – June 11, 2017

Similar testing procedures were implemented for the first and second rounds of testing. Likewise, the same testing environment was utilized for administration of the first and second rounds of testing.

2) Attitudinal Surveys

Educational Justice administers attitudinal and feedback surveys that Achievers complete on two occasions. Educational Justice uses these surveys in an attempt to gauge changes in attitudes among Achievers towards school and learning, satisfaction with the program, and progress to date. The surveys are administered as follows: (i) first survey date (entrance survey) -- during their registrations sessions (before tutoring commences) and (ii) second survey date (exit survey) -- near the end of the school year (For the present study, the second survey date coincided with the second test date window, i.e., between May 1st and June 11th of 2017).

Items in the survey provide Achievers with various statements, and in completing the survey, each Achiever is requested to select the extent to which he/she identifies (agrees/disagrees) with the statement provided. To meet the focus of the present study, only items 2, 3, 5, 7 of the Achiever Attitudinal Survey were examined. Specifically, those items are as follows:

Item #2: *I like learning.*

Item #3: *When I'm having trouble with homework, I have someone who can help me.*

Item #5: *I am good at math.*

Item #7: *Overall, I get good grades.*

For the attitudinal surveys, in order to quantify the results, each response was assigned an numerical value in accordance with Table 4 below.

Table 4. Numerical Values Assigned to Attitudinal Survey Responses

Survey Option	Assigned Value
Strongly Disagree	1
Disagree	2
Slightly Agree	3
Agree	4
Strongly Agree	5

Hypotheses

In the present study, three hypotheses are posed, as follows:

- Hypothesis A: The data will indicate a significant improvement in performance on the standardized math assessment between the first and second test dates among Achievers in the experimental group relative to those in the control group;

- Hypothesis B: The data will indicate a direct relationship between x, the number of hours of one-on-one tutoring completed with in the period from the first to the second test date, and y, the extent of improvement on the standardized math assessment from the first to the second test date, among Achievers in the experimental group relative to those in the control group; and
- Hypothesis C: The data will indicate an improvement in the average equivalent numerical value for responses to items # 2, 3, 5, and 7 on the attitudinal surveys between the first and second survey dates among Achievers in the experimental group relative to those in the control group.

Estimation methodology

The following exclusion criteria apply to assessment and survey data analyzed for this study:

Standardized Math Assessment Data

All testing data associated with any Achievers in the following cases was not included in the evaluation sample:

(1) Achievers whose parents/guardians opted out of participation in the present study.

(2) Achievers who are unable to complete both assessments: Form A and Form B on the first and second test dates, respectively.

(3) Achievers in the experimental group who do not complete a minimum of 12 hours of one-on-one tutoring with their Activists between the first and second test dates.

(4) Achievers who do not test within the testing windows specified above.

(5) Achievers in the experimental group who change Activists at any point between the first and second testing dates.

(6) Achievers in the control group who were paired at any point between the first and second testing dates.

Attitudinal Survey Data

All survey data associated with any Achiever in the following cases was not included in the evaluation sample:

(1) Achievers whose parents/guardians opted out of participation in the present study.

(2) Achievers who are unable to complete the first survey during a registration session occurring no later than February 28, 2017.

(3) Achievers who are unable to complete the second survey between May 1, 2017 and May 31, 2017.

(4) Achievers in the experimental group who do not complete a minimum of 12 hours of one-on-one tutoring with their Activists between the first and second test dates.

(5) Achievers who do not test within the testing windows specified above.

(6) Achievers who change Activists at any point between the first and second testing dates.

Chapter 4 Results

By the conclusion of the second testing window (May 31, 2017), 11 of the 12 students in the control group had become paired with an Activist and, therefore, had to be excluded from analysis, in accordance with the present study's pre-analysis plan. With no significant control group sample size remaining, none of the three hypotheses could be evaluated as described above. However, in order to nevertheless provide a potentially useful analysis of the collected data, the results of the present study were instead evaluated as follows: (1) the approach was shifted to a more single-case research style of analysis (as in some of the studies discussed in the background), in which each student's raw scores from the post-assessment and exit survey were compared to that same student's raw score results from the pre-assessment and entrance survey, respectively, from which effect sizes were calculated; and (2) since the *KeyMath-3 Diagnostic Assessment* is a normalized exam, effect sizes were also calculated based on the average changes in percentile rank for all students (Connolly, 2007).

In addition, rather than using the customary *Cohen's d* to calculate effect size, in which the differences between the post-assessment and pre-assessment scores are divided by the standard deviation of the pooled scores, more conservative approach was employed using, instead, the standard deviation of only the pre-assessment scores. More specifically, to calculate each Achiever's effect size (ES) for each measure (i.e., raw score, percentile rank, or attitudinal shift), the student's result on the pre-assessment (or entrance survey) was subtracted from the result on the post-assessment (or exit survey), and the difference was then divided by the standard deviation of all of the Achievers' results on the pre-assessment (or entrance survey) for that measure. In other words, the following formula was employed in the present study to calculate the effect size for each student:

$$ES = \frac{x_{post} - x_{pre}}{s_{pre}}$$

The effect sizes for all Achievers were then averaged for each measures, and the resulting mean effect sizes are presented in the results and data analysis provided below.

Overview

Tables 5 and 6 summarize the results for Achievers' change in performance on the standardized math assessments and any change in responses on the attitudinal surveys.

With regard to academic performance, while the original potential sample size included 127 Achievers, data for 71 students were excluded for not meeting all six of the inclusion criteria outlined above and in the pre-analysis plan for analysis of the standardized math assessment data, leaving an overall sample size of 56 for the experimental group. Of the 71 excluded, the vast majority did not two criteria: (1) they either were not able to complete the post-assessment in the specified testing window, if at all (criterion #2); or (2) they did not manage to complete at least 12 hours of tutoring between the pre-assessment and post-assessment (criterion #3).

Table 5 reveals that of the 56 tutees who met the inclusion criteria, tutees showed greater gains in the "Written Computation in Multiplication and Division" section of the

KeyMath-3 Diagnostic Assessment than in the "Written Computation in Addition and Subtraction" sections. Specifically, when it comes to performance in terms of raw score (number of exam items for which the Achiever provided a correct response) in multiplication/division, tutees showed a positive gain of 30% of a standard deviation (ES = 0.30) on Form B (the post-assessment) in comparison to Form A (the pre-assessment). However, performance on the addition/subtraction portion of the exam revealed a smaller gain (ES = 0.13). In terms of percentiles, the effect sizes, while still positive, were considerably smaller -- 0.14 in multiplication/division and only 0.02 in addition/subtraction -- implying that the gains by Achievers are less substantial in comparison to the general population of test-taking peers on the *KeyMath-3 Diagnostic Assessment* (Connolly, 2007). Since the original intention of the research design to conduct a comparison to a control group could not be realized, the percentile effect size calculations are provided here as a means for comparing the experimental group to a outside sample of students (i.e., the normalized sample group of students in the general population who completed the *KeyMath-3 Diagnostic Assessment*, as per Pearson's Diagnostic Assessment Manual) (Connolly, 2007).

In addition, because Pearson's Diagnostic Assessment Manual also provides normative tables for an estimated equivalent age (the age -- in years and months -- at which a given raw score is the median score) for each student's score, the average change in age equivalency for the experimental group has also been calculated and displayed alongside the average amount of time that transpired between administrations of Forms A and B for all tutees in the experimental group (Connolly, 2007). Thus, while the average amount of time between test dates for all Achievers was 3.26 months, the results indicate that tutees experienced a 4.46-month growth in addition/subtraction and an improvement equivalent to 8.70 months of growth in multiplication/division during that time.

Table 5. Overall Results in Academic Performance Assessment

Academic Performance Assessment	**Mathematics Subject Area**	
Measure (N = 56)	**Addition and Subtraction**	**Multiplication and Division**
Average Effect Size (Raw Score)	0.13	0.30
Average Effect Size (Percentile)	0.02	0.14
Average Time between Tests (in months)	3.26	3.26
Average Change in Age Equivalency (in months)	4.46	8.70

When it comes to the assessment of non-academic outcomes (Table 6), Achievers' responses showed gains in all four survey items, suggesting that their attitudes about school and learning may have undergone some small-to-moderate shifts during the EJA intervention. The sample size (N = 44) for these results is somewhat smaller than for the academic outcomes, primarily due to the third inclusion criterion for the attitudinal survey data; namely, 12 Achievers who met the criteria for inclusion in the analysis of the standardized math assessment data were unable to complete the second/exit survey within the specified window and, as such, had to be excluded.

Perhaps not surprisingly, the greatest effect size (ES = 0.43) occurred with item #3, which asks Achievers the extent to which they agree that they have someone who can help them with homework. Achiever responses associated with overall grade performance in school (item #7) was also significant (ES = 0.36), although a smaller attitudinal shift (ES = 0.28) was associated specifically with perceived ability in mathematics (item #5). The survey item designed to assess changes in attitude regarding a student's general love of learning (item #2) showed the least amount of gain (ES = 0.14).

Table 6. Overall Results of Attitudinal Shifts

Non-Academic Outcome Assessment (N = 44)		
Item	**Survey Statement**	**Average ES**
2	*I like learning.*	0.14
3	*When I'm having trouble with homework, I have someone who can help me.*	0.43
5	*I am good at math.*	0.28
7	*Overall, I get good grades.*	0.36

Tutoring Dosage

Tables 7, 8, and 9 show the results for both sample groups (academic and non-academic outcomes) according to tutoring dosage (in hours). As is customary in the research literature and in the studies cited in the background, each sample was divided into two subgroups along the median number of tutoring hours received, which was 13. Thus, those receiving a lower dosage of tutoring completed 13 or fewer one-on-one tutoring hours with their Activists, while those receiving a higher dosage of tutoring completed more than 13 hours of one-on-one tutoring. For the lower-dosage group, which included 25 Achievers, all of the Achievers completed either 12 or 13 hours of tutoring (those who received less than 12 tutoring hours did not meet the inclusion criteria). For the higher-dosage group, which included 31 Achievers, all of the students received between 13.5 and 20 hours of tutoring, with the exception of two Achievers who managed to complete 28 and 30 hours of tutoring with their Activists between the pre-assessment and post-assessment.

Table 7: Academic Performance Assessment Results by Dosage for Written Computation in Addition and Subtraction

Academic Performance: Addition/Subtraction	**Tutoring Dosage**	
Measure	**Lower (≤ 13 hours) (N = 25)**	**Higher (> 13 hours) (N = 31)**
Average Effect Size (Raw Score)	0.04	0.20
Average Effect Size (Percentile)	-0.01	0.03
Average Time between Tests (in months)	3.18	3.33
Average Change in Age Equivalency (in months)	3.92	4.90

Table 7 summarizes the results according to dosage for the "Written Computation in Addition and Subtraction" portion of the *KeyMath-3 Diagnostic Assessment*. Overall, students in the higher-dosage group showed greater gains. Specifically, Achievers in the higher-dosage group showed a greater improvement in raw scores (ES = 0.20) between the pre-assessment and post-assessment in comparison to Achievers in the lower-dosage group (ES = 0.04). The difference between the two groups was less pronounced with respect to percentile change, where the higher-dosage group (ES = 0.03) showed only small gains in comparison to the lower-dosage group, which showed a slightly negative result (ES = -0.01). Despite the miniscule changes in percentile rank, the change in age equivalency showed that in 3.18 months on average between test dates, students in the lower-dosage group achieved a 3.92-month improvement in performance. Achievers in the higher-dosage group displayed an even greater improvement: a gain of 4.90 months in the 3.33 months, on average, between the assessment dates.

Table 8. Academic Performance Assessment Results by Dosage
for Written Computation in Multiplication and Division

Academic Performance: Multiplication/Division	**Tutoring Dosage**	
Measure	**Lower (≤ 13 hours) (N = 25)**	**Higher (> 13 hours) (N = 31)**
Average Effect Size (Raw Score)	0.52	0.32
Average Effect Size (Percentile)	0.42	0.16
Average Time between Tests (in months)	3.18	3.33
Average Change in Age Equivalency (in months)	15.64	3.10

Table 8 summarizes the results according to tutoring dosage for the "Written Computation in Multiplication and Division" portion of the *KeyMath-3 Diagnostic Assessment*. Overall, in contrast with the results for addition/subtraction above, students in the lower-dosage group showed greater gains than those in the higher-dosage group. Specifically, while Achievers in both groups showed significant gains in raw scores, those in the lower-dosage group showed a greater improvement (ES = 0.52) between the pre-assessment and post-assessment than that of the higher-dosage group (ES = 0.32). The lower-dosage group even showed significant gains with respect to percentile change (ES = 0.42), whereas the lower-dosage group's improvement was less pronounced (ES = 0.16). With respect to change in age equivalency, the results indicated that in 3.18 months on average between test dates, students in the lower-dosage group achieved a considerable 15.64-month improvement in performance. Achievers in the higher-dosage group, however, showed only a 3.10-month improvement, despite 3.33 months, on average, traspiring between test dates.

With regard to non-academic outcomes, of the original sample size of 44 students, 18 Achievers made up the lower-dosage group while 26 Achievers made up the higher-dosage one. The higher-dosage group showed greater mean effect sizes on three of the four attitudinal survey statements evaluated (Table 9). More specifically, the higher-dosage group showed a modest improvement (ES = 0.20) in favorable attitudes toward learning (item #2) while the lower-dosage group showed no gains on the same measure (ES = 0.00). The most significant gain (ES = 0.62) was made by the higher-dosage group in response to item #3, in which Achievers expressed a greater sense of confidence in having an individual in their lives to help with homework when needed. The lower-dosage group also showed a positive albeit significantly smaller mean effect size on the same measure (ES = 0.21). Similarly, Achievers in the higher-dosage group indicated an increased sense of accomplishment in school when it comes to grades (ES = 0.47 for item #4), while the lower-dosage group again showed a positive albeit significantly smaller mean effect size (ES = 0.19) for changes in responses to the same item between the entrance and exit surveys. Finally, however, when it came to self-perceived ability in mathematics, students in the lower-dosage group showed greater gains (ES = 0.36) in comparison to the larger-dosage group (ES = 0.25).

Table 9. Results of Attitudinal Shifts by Dosage

Non-Academic Outcome Assessment		**Tutoring Dosage**	
Item	**Survey Statement**	**Lower (≤ 13 hours) (N = 18)**	**Higher (> 13 hours) (N = 26)**
2	*I like learning.*	0.00	0.20
3	*When I'm having trouble with homework, I have someone who can help me.*	0.21	0.62
5	*I am good at math.*	0.36	0.25
7	*Overall, I get good grades.*	0.19	0.47

Tutee Grade

The results of the present study were also analyzed to identify trends, if any, for Achievers in different grades in school. Tables 10, 11, and 12 show the results for both sample groups (academic and non-academic outcomes) according to tutee grade. Since the sample size of 56 students was relatively small to split into the four grades (5th, 6th, 7th, and 8th) that Achievers represent, the sample was instead divided into two subgroups: (1) students in 5th or 6th grade, of which there were 29; and (2) students in 7th or 8th grade, of which there were 27.

Table 10. Academic Performance Assessment Results by Tutee Grade for Written Computation in Addition and Subtraction

Academic Performance: Addition/Subtraction	**Tutee Grade**	
Measure	**5th-6th (N = 29)**	**7th-8th (N = 27)**
Average Effect Size (Raw Score)	0.08	0.19
Average Effect Size (Percentile)	-0.05	0.09
Average Time between Tests (in months)	3.25	3.28
Average Change in Age Equivalency (in months)	3.31	5.70

Table 10 summarizes the results according to tutee grade for the "Written Computation in Addition and Subtraction" portion of the *KeyMath-3 Diagnostic Assessment*. Overall, students in 7th-8th grade group showed greater gains. Specifically, Achievers in the 7th-8th grade group showed a greater improvement in raw scores (ES = 0.19) between the pre-assessment and post-assessment in comparison to Achievers in the 5th-6th grade group (ES = 0.08). The difference between the two groups was less pronounced with respect to percentile change, where the 7th-8th grade group (ES = 0.09) showed only a small gain, and the 5th-6th grade group showed a slightly negative mean effect size (ES = -0.05). With the change in age equivalency, the 5th-6th grade group showed no significant gains: a growth of 3.31 months in 3.25 months, on average. Achievers in the 7th-8th grade group, however, displayed a greater improvement: a gain of 5.70 months in the 3.28 months, on average, between the assessment dates.

Table 11. Academic Performance Assessment Results by Tutee Grade for Written Computation in Multiplication and Division

Academic Performance: Multiplication/Division	**Tutee Grade**	
Measure	**5th-6th (N = 29)**	**7th-8th (N = 27)**
Average Effect Size (Raw Score)	0.42	0.21
Average Effect Size (Percentile)	0.17	0.10
Average Time between Tests (in months)	3.25	3.28
Average Change in Age Equivalency (in months)	11.86	5.30

Table 11 summarizes the results according to tutee grade for the "Written Computation in Multiplication and Division" portion of the *KeyMath-3 Diagnostic Assessment*. Overall, in contrast with the results for addition/subtraction above, students in the 5th-6th grade group showed greater gains than those in the 7th-8th grade group. Specifically, while Achievers in both groups showed some gains in raw scores, those in the 5th-6th grade group showed a greater improvement (ES = 0.42) between the pre-assessment and post-assessment than that of the 7th-8th grade group (ES = 0.21). The 5th-6th grade group showed smaller gains with respect to percentile change (ES = 0.17), but the 7th-8th grade group's improvement was, again, more modest (ES = 0.10). With respect to change in age equivalency, the results indicated that in an average of 3.25 months between the administration of the two assessments, students in the 5th-6th grade group achieved a considerable 11.86-month improvement in performance. Achievers in the 7th-8th grade group showed a smaller but still significant 5.30-month improvement in the 3.28 months, on average, that transpired between the pre and post test dates.

Lastly, with respect to the assessment of non-academic outcomes, of the original sample size of 44 students who met the inclusion criteria for the attitudinal survey data, the 5th-6th grade included 22 Achievers, as did the 7th-8th grade group. Overall, the results were split: the 7th-8th grade group showed greater mean effect sizes on two of the four attitudinal survey items evaluated for changes between the entrance and exit surveys, while the 5th-6th grade group showed greater gains on the other two attitudinal items (Table 12). More specifically, the 5th-6th grade group showed a significant improvement (ES = 0.30) in its agreement with the statement of favorable attitude toward learning (item #2), while the 7th-8th grade group showed a negative mean effect size in responding to the same item (ES = -0.16). Again, the greatest gain (ES = 0.54) was made by the 7th-8th grade group in response to item #3, in which Achievers agreed to a greater degree since beginning their participation in the intervention that they had access to an individual who could help with homework when needed. The 5th-6th grade group also showed a positive albeit smaller mean effect size on the same measure (ES = 0.35).

Similarly, the results show that Achievers in the 7th-8th grade group indicated a significant increase in their sense of success regarding grade performance in school (ES = 0.51 for item #4), while the 5th-6th grade group showed a positive albeit significantly smaller mean effect size (ES = 0.23) in responses to the same item. Finally, however, with regard to Achievers' self-perceived abilities in mathematics, students in the 5th-6th grade group showed significantly greater gains (ES = 0.45) in comparison to the 7th-8th grade group (ES = 0.14).

Table 12. Results of Attitudinal Shifts by Tutee Grade

Non-Academic Outcome Assessment		**Tutee Grade**	
Item	**Survey Statement**	**5th-6th (N = 22)**	**7th-8th (N = 22)**
2	*I like learning.*	0.30	-0.16
3	*When I'm having trouble with homework, I have someone who can help me.*	0.35	0.54
5	*I am good at math.*	0.45	0.14
7	*Overall, I get good grades.*	0.23	0.51

Chapter 5 Discussion

Regrettably, the need to exclude all but one of the students in the control group rendered the evaluation of the three hypotheses as originally described in the present study unfeasible. However, the shift to evaluating the data using a more single-case research approach by generating mean effect sizes for changes in each student's performance on the assessments and responses to attitudinal surveys allows the results to nevertheless be evaluated with consideration to the intention and essence of the original hypotheses. In particular, the first hypothesis expects that results will indicate a significant improvement in performance on the academic assessments given to Achievers; the second, which focuses on dosage, expects that students who received a greater of hours of tutoring will show greater gains than those who received fewer hours of tutoring; and the third anticipates an improvement in attitudinal outcomes as evidenced by the tutees' survey responses.

Beyond the calculations of effect sizes for raw scores, the addition of mean effect size calculation based on percentiles as well as age equivalency data, both of which are based on a normalized general population of test takers outside the experimental group, strengthen the trend analysis. Unfortunately, however, no such additional measures on students outside the experimental group was available for the non-academics outcome component of the study.

Academic Performance

With respect to the first hypothesis, Table 5 shows that only the mean effect size in raw score for multiplication/division (ES = 0.30) appears significant albeit modest. The other results are minimal: effect sizes of 0.13 and 0.02 in addition/subtraction for raw score and percentile, respectively; even in multiplication/division, the mean effect size for percentile rank (ES = 0.14) was small. At the same time, results indicate that students in the experimental group progressed by by 4.46 months in their addition and subtraction abilities compared to the KeyMath-3 Diagnostic Assessment's normalized group of test takers, who showed about 3.26 months of growth during that time. In multiplication and division, the age equivalency results were even more pronounced: Achievers progressed by by 8.70 months compared to the normalized group of their test-taking peers, who, again, showed about a gain of 3.26 months during the same amount of time. The reasons for the somewhat inconclusive nature of the data are unclear. It is known that Educational Justice chose to administer the post-assessment during the same time that many of the Achievers were experiencing consecutive days of state-mandated testing in the local school district. Thus, test fatigue may have played a role in the Achievers performance on the post-assessment. Also, because EJA does not have a prescribed curriculum, it is possible that a greater number of Achievers worked on multiplication and division skills (or higher math that required review of those skills) during tutoring sessions with their Activists, or perhaps even in the sessions in the weeks leading up to the assessments. At the same time, many Activists may have shifted towards devoting more time to reading and language proficiency in the sessions before the post-assessment. Regardless, it is apparent that Achievers did see gains from their participation in EJA. However, based on the results of the present study, these gains are, generally speaking, not as significant as those found by the meta-analyses cited in the background (see Table 2), which appears to suggest that Hypothesis A would have been rejected if the control group had remained intact.

Tutoring Dosage

With the exception of the Jun et al. (2010) study, the meta-analyses discussed in the background suggest that mean effect sizes are equivalent if not slightly greater for groups receiving lower dosages of tutoring than those receiving higher ones (Table 2).

In considering the second hypothesis, the results (Tables 7 and 8) reveal that, for the experimental group in the present study, this trend is generally consistent with the meta- analyses in multiplication/division but not in addition/subtraction. The higher-dosage group showed greater gains than the lower-dosage group in effect sizes for raw score (0.20 compared to 0.04) and for percentile rank (0.03 compared to -0.01), as well as in age equivalency over a similar timeframe (4.90 months of gain in 3.33 months compared to 3.92 months of gain in 3.18 months). While it clearly cannot be concluded in the absence of a control group, Hypothesis B seems to be strengthened by these results for written computation in addition/subtraction. Hypothesis B appears to be contradicted by the results for multiplication/division, in which greater gains were calculated for raw score and percentile effect sizes as well as for age equivalency. The explanation for this conflicting trend is unclear, but the same kinds of possible causes may apply as those outlined above, including test fatigue among tutees or the content covered, particularly most recently, in the Achievers' tutoring sessions with their Activists.

Tutee Grade

The meta-analyses discussed in the background lack agreement regarding a trend, if one even exists, in the effect of peer-tutoring interventions on older versus younger students (Table 2). While no hypothesis was posed in the present study regarding a trend for effect sizes of students in higher grades in school (or older groups of tutees) in comparison to those in lower grades (or younger in age), it is interesting to note the results according to tutee grade (Tables 10 and 11). A similar pattern emerged as that for tutoring dosage: in the addition/subtraction assessment, older students (7th and 8th grade Achievers) showed greater gains than younger students (5th and 6th grade Achievers); in multiplication and division, however, it was reversed: the younger group made the larger gains. Interestingly, the 7th-8th grade group maintained relatively comparable gains with respect to all three metrics in both addition/subtraction and multiplication/division: raw score mean effect size in (0.19 and 0.21, respectively), percentile rank mean effect size (0.09 and 0.10, respectively), and age equivalency (5.30 and 5.70 months of gain, respectively, over the course of an average of 3.28 months). The 5th-6th grade group, however, showed low, insignificant, or even negative gains in addition/subtraction (raw score and percentile effect sizes of 0.08 and -0.05, respectively, and only 3.31 months of gain in age equivalency over 3.25 months), while in multiplication/division, the results were considerable more favorable (raw score and percentile effect sizes of 0.42 and 0.17, respectively, and only 11.86 months of gain in age equivalency over the same 3.25-month period). The best explanation for such a disparity appears to be, again, the freedom that Activists are allotted in determining the content and path of tutoring sessions with their Achievers. If the 5th-6th grade students required more remediation in multiplication and division than in addition/subtraction -- perhaps due to what was being in school at that time -- it is likely that their Activists directed more of their tutoring time to addressing that content area. The 7th-8th graders, on the other hand, may have merely focused, for example, on the Achievers' math assignments rather than on remediation of particular areas of arithmetic, and, as a result, displayed less polarized gains.

Non-Academic Outcomes

Returning to the meta-analyses discussed in the background, the research suggests that, when it comes to non-academic outcomes, positive gains can generally be expected among tutees in response to peer tutoring interventions (Tables 3a and 3b) and that the corresponding effect sizes are often significant (gains of about 25% of a standard deviation or greater).

In considering the third hypothesis posed by the present study, the attitudinal results (Table 6) reveal that, for Achievers in the experimental group, positive gains were shown for all four of the survey items. With the exception of the item intended to assess a general shift in tutees' attitude towards learning (item #2), which showed a relatively low mean effect size (0.14), the other three items showed more significant gains for items 3, 5, and 7 (mean effect sizes of 0.43, 0.28, and 0.36, respectively). In spite of the lack of a control group, to further explore the assertion put forth by Hypothesis C that Achievers would display significant positive attitudinal shifts toward school and learning, the attitudinal survey data results were also analyzed according to tutoring dosage and tutee grade in a similar manner as was conducted for the academic performance data (Tables 9 and 12).

For item #2, the higher-dosage group showed higher gains (ES = 0.20) than the lower-dosage group, which showed no gain at all (ES = 0.00). This may suggest that a threshold of tutoring hours must be reached before tutees begin to experience a shift to a more favorable attitude toward learning. For the same item, the 5th-6th grade group showed higher gains (ES = 0.30) than the 7th-8th grade group, which showed a negative gain (ES = -0.16). These results may be reflective of social pressures, often experienced to a greater degree among older students in the later middle school years than their younger peers, that stigmatize positive sentiments toward learning (or at least the expression thereof). Regardless of the underlying causes, these two elements may at least partially explain the relatively low effect size (0.14) for item #2 by the experimental group as a whole and further suggest that younger (5th-6th grade) tutees who receive a high dosage of tutoring may experience the greatest gains in attitude toward learning in general.

Achievers' attitudes toward their self-perceived abilities in mathematics (item #5) also yielded interesting results. Tutees in the lower-dosage group and those in the 5th-6th grade group showed greater attitudinal gains (mean effect sizes of 0.36 and 0.45, respectively) on this survey item than tutees who received more tutoring or were a few years older (mean effect sizes of 0.25 and 0.14, respectively). It is possible that only at higher doses of tutoring hours do tutees begin to confront more challenging mathematical problems (perhaps with reduced assistance from their tutors) or to realize the extent of their learning gaps in mathematics, and, as a result, their self-concept when it comes to their math abilities does not rise nearly as quickly as it does among newer and younger tutees.

Clearly, Hypothesis C cannot be concluded in the absence of a control group; however, its general expectation that the EJA peer tutoring intervention will yield significant gains in attitudinal shifts among tutees toward learning, academic support, mathematical ability, and overall school performance appears to be supported by the results for non-academic outcomes investigated in the present study.

Chapter 6 Conclusion

In general agreement with the body of research on peer tutoring interventions, the results of the present study suggest that tutees benefit -- both in academic performance and in non-academic outcomes -- from participating in the EJA peer tutoring program. To better evaluate peer tutoring trends in the EJA model as a means of (1) informing Educational Justice on the design strengths and areas for growth in its intervention efforts and effectiveness; and (2) further contributing to the general body of educational research on peer tutoring interventions, it may be beneficial to consider the following suggestions.

Limitations of the Data and Implications for Practice

In an effort to be most accommodating to the low-income families they serve, the staff at Educational Justice provided a window of dates during which the pre-assessment and post-assessment may be taken. The span of of these windows (over a month long) was relatively large, however, and created greater variability in time between pre-test and post-test dates within the experimental group. Specifically, students may have experienced as little as 62 days between pre-test and post-test to as much as 138 days between the assessments. Given that the results of the pre-assessment are provided to the Activists to guide instruction in their tutoring sessions, it is reasonable to expect that those who had less time to benefit from the test results might yield less substantial gians than those who had more time with the pre-assessment results to serve as instructional guidance before the post-assessment.

This effect was compounded by the fact that the results of the pre-assessment could not be provided to Activists expeditiously. Because the *KeyMath-3 Diagnostic Assessment* required hand-scoring (not to mention repeatedly, in an effort to ensure accuracy) and results needed to be organized into easy-to-follow reports, Educational Justice staff members required as much as six weeks to deliver results to Activists and Achiever families. This may have left as few as two or three tutoring sessions, each an hour long, for the Activist to attempt to address the academic needs of the Achiever based on results in two sections of a mathematics assessment (addition/subtraction and multiplication/division) and two sections of reading assessment (general vocabulary and paragraph reading), all while continuing to assist the Achiever with homework and other school assignments.

In addition, as mentioned previously, the testing window scheduled by Educational Justice for the post-assessment overlapped with state-mandated, week-long standardized testing dates at the local school district, which may have contributed to test fatigue among some Achievers and, therefore, negatively impacted their results. In fact, some Achievers reported that they had completed the post-assessment at the Educational Justice offices in the late afternoon or evening on a day when they had endured a full day of testing at school.

Furthermore, the testing requirement was added to the EJA model mid-year, which may have created some confusion and even resistance among parents/guardians of Achievers who likely did not anticipate such a requirement when enrolling in the program. In addition, the only testing site provided for Achievers to complete the assessments was the Educational Justice central office located on the east side of Louisville; many Achiever families, however, reside on the west or south end of the city, as much as 15-20 miles from the site. Given the considerable hurdle that arranging transportation often poses, particularly to underserved populations, the added burden likely prevented many Achievers from testing. Indeed, it is reasonable to assume that all of these factors contributed to the nearly 40% attrition in the number of students testing between the two assessments.

It is known that Educational Justice's testing budget is limited, and the assessment test forms were donated by a local university. Nonetheless, the age of the exam should also be called into question, the *KeyMath-3 Diagnostic Assessment* being a ten-year-old exam. Not only did the assessment require hand-scoring -- which was not only time-consuming but introduces the likelihood of scoring errors), many of its questions for the written computation in both addition/subtraction and multiplication/division allowed for little leniency in scoring and did not accommodate for partial credit. One item number, for example, asked students to subtract two fractions and to provide the response "as a mixed number in lowest terms." It turned out that a few students arrived at the correct answer but either did not convert it to a mixed number or did not reduce it (or both), and, as such, received zero credit for the response, in accordance with the assessments scoring guidelines.

Finally, with regard to non-academic outcomes, Educational Justice currently only administers attitudinal surveys, from which students' attitudes and self-concept assessments may be derived. However, no means of assessing other non-academic measures (such as behavioral or social-emotional outcomes) are employed, yielding a limited picture of Achievers' response to the intervention beyond academic performance.

To address these concerns, the following recommendations are offered:

(1) To the extent that funding allows, newer computer-adaptive assessments should be used, particularly those that provide results according the quantile and lexile framework. By obviating hand-scoring, such assessments are able to provide immediate results to Educational Justice staff and Activists as well as to Achievers and their families, in addition to user-friendly quantile-based and lexile-based reports that link to free resources targeted to address areas for growth revealed by the testing.

(2) The assessments should be administered more frequently. Even adding one additional test administration may prove beneficial, such as by having a pre-test at the start of each school year, a post-test during the middle of the year (e.g., before winter break), and a post-post-test near the conclusion of the academic year.

(3) The assessments and surveys should become a standard requirement for Achievers to participate in the program, thereby providing parents/guardians prior notice to prepare for these assessments that occur throughout the year.

(4) The assessments and surveys should be administered using mobile devices (such as laptop computers or tablets) at any of the approved locations within the Educational Justice network of tutoring sites. Doing so would reduce the transportation burden associated with testing on Achiever families, while still allowing an Educational Justice staff member to supervise the test administration.

(5) An additional assessment for non-academic outcomes (or the expansion of the current attitudinal surveys to include such an assessment) should be implemented to assess social-emotional and/or behavioral measures.

(6) A system for tracking of content covered by Activist-Achiever pairs in their tutoring session along with the amount of time spent on each subject should be implemented. Such a system may allow Educational Justice staff and researchers of the EJA model to more easily dismiss or support the various speculated explanations for the calculated results in academic performance and non-academic outcomes among Achievers.

Limitations in Methodology and Future Research

The present study suffered from a number of methodological flaws that should be addressed in any future study of the EJA model. First and foremost, the loss of the control group rendered the evaluation of the original hypotheses unfeasible. The results, nevertheless, provided some insights along the lines of inquiry targeted by the hypotheses, but a re-evaluation of the EJA model with a proper control group would be much preferred. The approach of using students on the EJA waitlist is ideal, as they meet all of the requirements to become Achievers and are virtually indistinguishable from other Achievers in the EJA program with the exception that they have yet to be paired with an Activist. However, as occured with the present study, Achievers on the waitlist may become paired with an Activist before post-assessments are administered, thereby excluding their results from the control group data. At the same time, preventing Achievers in the control group from being paired and receiving the academic support they need would pose ethical concerns, and, in fact, Educational Justice operating guidelines forbid doing so. It is, therefore, suggested that a much larger initial control group be used that would allow for considerable attrition of students (presumably as they became paired) without reduction in the sample size to insignificant levels by the time post-assessments are administered.

Beyond implementing a control group design, future studies of the EJA model would ideal span greater amounts of time. More specifically, studies of Achievers participating throughout at least one academic year in the EJA intervention -- particularly if implemented according to the assessment recommendations outlined in the implications for practice above and accommodating three or more testing administrations (e.g., following the pre, post, and post-post schedule suggested above) per student throughout the school year -- would provide two central benefits: (1) better conditions under which to evaluate the effect of dosage; and (2) the opportunity to conduct a more rigorous single-case research analysis by enabling multiple assessments that establish a baseline of results for each student before treatment (e.g., with Achievers who begin the program on the waitlist and complete the pre-test and post-test but become paired with an Activist before the post-post-test).

With regard to dosage, analysis over greater spans of time would allow for the evaluation of the type of trend suggested by Hypothesis B. Namely, instead of creating two dosage groups above and below the median number of tutoring hours received by tutees in the experimental group (as was done in the present study and in many of the research studies cited in the background section), a much more granular assessment of correlation, if any, between the number hours (on the x-axis) and effect size (on the y-axis) could be conducted. Furthermore, since Achievers and Activists may be paired for up to four years, a more lengthy study of the EJA model would not only enable investigation of results for ultra-high doses of one-on-one peer tutoring hours but may also shed additional light on some of the conflicting results in the literature (including in the present study) when it comes to identifying a clear trend for tutoring dosage.

As for further exploration of the mediating factor of tutee grade, to better assess any trends, a greater initial sample size and implementation of the assessment recommendations above would prevent the need to group Achievers into multi-grade groups (5th-6th and 7th-8th), as was done in the present study. Instead, the experimental group could be divided according to the four grades of students in the program (5th, 6th, 7th, and 8th) without concern for reducing the sample sizes in each subgroup to insignificant levels.

Lastly, some of the literature reviewed for this study suggests that same-sex dyads display considerably greater gains than mixed-sex pairs. Since Achievers in the EJA model choose their own Activists, tutor-tutee pairs of every binary sex permutation (female tutor-female tutee, female tutor-male tutee, male tutor-female tutee, and male tutor-male tuee) abound in the program. Thus, it would be feasible and suggested that further studies on the EJA model explore hypotheses along these lines of inquiry.

References

Bowman-Perrott, L., Burke, M. D., Zhang, N., & Zaini, S. (2014). Direct and collateral effects of peer tutoring on social and behavioral outcomes: A meta-analysis of single-case research. *School Psychology Review, 43*(3), 260.

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review, 42*(1), 39. Retrieved from https://search.proquest.com/docview/1349929936

Bowman-Perrott, L., deMarín, S., Mahadevan, L., & Etchells, M. (2016). Assessing the academic, social, and language production outcomes of english language learners engaged in peer tutoring: A systematic review. *Education & Treatment of Children, 39*(3), 359-388. doi:10.1353/etc.2016.0016

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*(2), 237-248. doi:10.3102/00028312019002237

Cole, M. W. (2013). Rompiendo el silencio: Meta-analysis of the effectiveness of peer-mediated learning at improving language outcomes for ELLs. *Bilingual Research Journal, 36*(2), 146-166. doi:10.1080/15235882.2013.814609

Cole, M. W. (2014). Speaking to read: Meta-analysis of peer-mediated learning for english language learners. *Journal of Literacy Research, 46*(3), 358-382. doi:10.1177/1086296X14552179

Collins, T. A., Hawkins, R. O., & Flowers, E. M. (2017). Peer-mediated interventions: A practical guide to utilizing students as change agents. *Contemporary School Psychology,* doi:10.1007/s40688-017-0120-7

Connolly, A. J. (2007). KeyMath-3 diagnostic assessment: Manual forms A and B. Minneapolis, MN: Pearson.

Cook, S. B., Scruggs, T. E., Mastropieri, M. A., & Casto, G. C. (1985). Handicapped students as tutors. *Journal of Special Education, 19*(4) Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=4726630&site=ehost-live&scope=site

Debbie R. Robinson, Janet Ward Schofield, & Katrina L. Steers-Wentzell. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review, 17*(4), 327-362. doi:10.1007/s10648-005-8137-2

Debra M. Kamps, Charles Greenwood, Carmen Arreaga-Mayer, Mary Baldwin Veerkamp, Cheryl Utley, Yolanda Tapia, . . . Harriett Bannister. (2008). The efficacy of ClassWide peer tutoring in middle schools. *Education and Treatment of Children, 31*(2), 119-152. doi:10.1353/etc.0.0017

Dietrichson, J., Bøg, M., Filges, T., & Klint Jørgensen, A. (2017). Academic interventions for elementary and middle school students with low socioeconomic status. *Review of Educational Research, 87*(2), 243-282. doi:10.3102/0034654316687036

Elbaum, B., Vaughn, S., Hughes, M. T., & Watson Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology, 92*(4), 605-619. doi:10.1037/0022-0663.92.4.605

Fantuzzo, J. W., Polite, K., & Grayson, N. (1990). An evaluation of reciprocal peer tutoring across elementary school settings. *Journal of School Psychology, 28*(4), 309-323. doi:10.1016/0022-4405(90)90021-X

Galbraith, J., & Winterbottom, M. (2011). Peer‐tutoring: What's in it for the tutor? *Educational Studies, 37*(3), 321-332. doi:10.1080/03055698.2010.506330

Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006a). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology, 98*(4), 732-749. doi:10.1037/0022-0663.98.4.732

Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006b). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology, 98*(4), 732-749. doi:10.1037/0022-0663.98.4.732

Greenwood, C. R., Terry, B., Arreaga-Mayer, C., & Finney, R. (1992). The classwide peer tutoring program: Implementation factors moderating students' achievement. *Journal of Applied Behavior Analysis, 25*(1), 101-116. doi:10.1901/jaba.1992.25-101

Gregory F. Harper, Barbara Mallette, Larry Maheady, Anita E. Bentley, & Jill Moore. (1995). Retention and treatment failure in classwide peer tutoring: Implications for further research. *Journal of Behavioral Education, 5*(4), 399-414. doi:10.1007/BF02114540

International journal of educational research. *International Journal of Educational Research,* Jun, S. W., Cumming, A., & Ramirez, G. (2010). Tutoring adolescents in literacy: A meta-analysis. *McGill Journal of Education, 45*(2), 219-238. doi:10.7202/045605ar

Kunsch, C. A., Jitendra, A. K., & Sood, S. (2007). The effects of peer-mediated instruction in mathematics for students with learning problems: A research synthesis. *Learning Disabilities Research & Practice, 22*(1), 1-12. doi:10.1111/j.1540-5826.2007.00226.x

Larry Maheady, Gregory F. Harper, & M. Katherine Sacca. (1988). Peer-mediated instruction: A promising approach to meeting the diverse needs of LD adolescents. *Learning Disability Quarterly, 11*(2), 108-113. doi:10.2307/1510988

Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement.*Journal of Educational Psychology, 107*(2), 558-579. doi:10.1037/a0037698

Lisa J. Bowman-Perrott, Charles R. Greenwood, & Yolanda Tapia. (2007). The efficacy of CWPT used in secondary alternative school classrooms with small teacher/pupil ratios and students with emotional and behavioral disorders. *Education and Treatment of Children, 30*(3), 65-87. doi:10.1353/etc.2007.0014

Miller, D., Topping, K., & Thurston, A. (2010). Peer tutoring in reading: The effects of role and organization on two dimensions of self-esteem. *The British Journal of Educational Psychology, 80*(Pt 3), 417-433. doi:10.1348/000709909X481652

Oddo, M., Barnett, D. W., Hawkins, R. O., & Musti-Rao, S. (2010). Reciprocal peer tutoring and repeated reading: Increasing practicality using student groups. *Psychology in the Schools, 47*(8), 842-858. doi:10.1002/pits.20508

Okilwa, N. S. A., & Shelby, L. (2010). The effects of peer tutoring on academic performance of students with disabilities in grades 6 through 12: A synthesis of the literature. *Remedial and Special Education, 31*(6), 450-463. doi:10.1177/0741932509355991

Reardon, S. F. (2013). The widening income achievement gap. Educational Leadership, 70(8), 10-16. Retrieved February 20, 2018 from: http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=87529513&site=ehost-live&scope=site

Rees, E. L., Quinn, P. J., Davies, B., & Fotheringham, V. (2016). How does peer teaching compare to faculty teaching? A systematic review and meta-analysis (.). *Medical Teacher, 38*(8), 829-837. doi:10.3109/0142159X.2015.1112888

Renee O. Hawkins, Shobana Musti-Rao, Cynthia Hughes, Laura Berry, & Shannon McGuire. (2009). Applying a randomized interdependent group contingency component to classwide peer tutoring for multiplication fact fluency. *Journal of Behavioral Education, 18*(4), 300-318. doi:10.1007/s10864-009-9093-6

Rod D. Roscoe, & Michelene T. H. Chi. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research, 77*(4), 534-574. doi:10.3102/0034654307309920

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students. *Journal of Educational Psychology, 95*(2), 240-257. doi:10.1037/0022-0663.95.2.240

Ryan, J. B., Reid, R., & Epstein, M. H. (2004). Peer-mediated intervention studies on academic achievement for students with EBD. *Remedial and Special Education, 25*(6), 330-341. doi:10.1177/07419325040250060101

Sheldon, D. A. (2001). Peer and cross-age tutoring in music. *Music Educators Journal, 87*(6), 33-38. doi:10.2307/3399690

Southern Education Foundation. (2015). *A New Majority.* Retrieved February 20, 2018 from SouthernEducation.org: http://www.southerneducation.org/getattachment/4ac62e27-5260-47a5-9d02-14896ec3a531/A-New-Majority-2015-Update-Low-Income-Students-Now.aspx

Stephen Houghton, & Alan Bain. (1993). Peer tutoring with ESL and below-average readers. *Journal of Behavioral Education, 3*(2), 125-142. doi:10.1007/BF00947032

Ya-yu Lo, & Gwendolyn Cartledge. (2004). Total class peer tutoring and interdependent group oriented contingency: Improving the academic and task related behaviors of fourth-grade urban students. *Education and Treatment of Children, 27*(3), 235-262. Retrieved from http://www.jstor.org/stable/42900545

Zeneli, M., & Tymms, P. (2015). A review of peer tutoring interventions and social interdependence characteristics. *International Journal for Cross-Disciplinary Subjects in Education, 5*(Special 2), 2504-2510. doi:10.20533/ijcdse.2042.6364.2015.0341

Zeneli, M., Tymms, P., & Bolden, D. (2016). The impact of interdependent cross-age peer tutoring on social and mathematics self-concepts. *International Journal of Psychology and Educational Studies, 3*(2), 1-13. doi:10.17220/ijpes.2016.02.001