UNIVERSITY OF CALIFORNIA
RIVERSIDE

Examining the Effect of Gesture-Use on Teaching Mathematical Equivalence in Elementary School Classrooms

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Arts

in

Education

by

Kelsey R. Rocha

June 2020

Thesis Committee:
Dr. Kinnari Atit, Chairperson
Dr. Marsha Ing
Dr. Rebekah Richert

The Thesis of Kelsey R. Rocha is approved:

____

____

____

Committee Chairperson

University of California, Riverside

ABSTRACT OF THE THESIS

Examining the Effect of Gesture-Use on Teaching Mathematical Equivalence in Elementary School Classrooms

by

Kelsey R. Rocha

Master of Arts, Graduate Program in Education
University of California, Riverside, June 2020
Dr. Kinnari Atit, Chairperson

Teachers regularly use gestures—meaningful movements of the hands (Alibali, 2005)—when conveying STEM content to their students. Researchers have found that gestures can be particularly useful as a scaffold for student learning of mathematical equivalence (i.e. Goldin-Meadow, Kim, & Singer, 1999; Novack et al., 2014; Singer & Goldin-Meadow, 2005). Mathematical equivalence is a fundamental math concept taught early in elementary school that many students struggle with fully understanding (McNeil, 2008). Though there is evidence that supports gesture-use when teaching mathematical equivalence in a laboratory setting, there is a lack of evidence supporting its use in a classroom setting. In this study we examined if teacher's gesture-use would be beneficial to students' mathematical equivalence understanding in second grade classrooms. Further, we investigated whether gesture-use would affect learning outcomes differently within English Language Learning (ELL) students. Participants (*n*=119) were from seven classes and approximately half of the sample was of ELL status. Results did not support the use of gestures being more effective than the "business as usual" control; however, there was not a significant difference between the learning outcomes of ELL students and their non-ELL counterparts. This study highlights the necessity for further research on the use of gestures (and their impact) in the classroom setting.

Table of Contents

List of Figures and Tables

Figures

Tables

According to the 2017 National Assessment of Educational Progress (NAEP), only 40% of fourth-grade students in the United States are proficient or above in math (nationsreportcard.gov). While learning math is a challenge for many students, one topic that has been shown to be especially difficult for elementary students to master, specifically from second to fifth grade, is mathematical equivalence (McNeil, 2008). Mathematical equivalence (ME) is the fundamental idea that quantities on either side of an equal sign are equal and interchangeable (Kieran, 1981). Experts think that a formal understanding of ME may be a core concept that drives the progression of understanding arithmetic operations and algebra (Charles & Carmel, 2005; Common Core Standards, NGA Center & CCSSO, 2010; McNeil et al., 2019; Alibali et al., 2007).

Prior research suggests that learning the principle concept of ME is a process which takes place over the duration of many elementary school years and it is crucial for further algebraic learning (Rittle-Johnson, 2011). Rittle-Johnson and colleagues' (2011) construct map for ME knowledge contains four levels which indicate increasingly complex understanding. These levels span from the least sophisticated level of knowledge (rigid operational knowledge) to the most sophisticated (comparative relational knowledge). At the operational level students are able to master the most basic structure of equations $a+b=c$ and operationally define the equal sign. Then, at the flexible operational level, students are able to solve atypically structured equations (such as $c=a+b$) that retain the operational view of the equal sign. In the third level, the basic relational level, students are able to solve more difficult problems which have operations on both sides of the equal sign (i.e. $a+b=c+d$). Finally, in the comparative relational level, students are able to successfully compare the expressions on both sides of the equal sign (i.e. ax+b=c) and recognize the relational definition of the equal sign. In this study, the researchers found that by fifth grade, most students seem to be firmly at the third level (basic relational) and begin gaining sophistication in their understanding, leading them into the comparative relational level.

Some researchers claim that individual differences in ME understanding can have a big impact on students' subsequent learning. In one study, McNeil and colleagues (2019) found the following: (1) that levels of ME understanding were low, overall, for children ages 7-11 and (2) ME understanding in the second grade significantly predicts math achievement in third grade, even after controlling for second-grade math achievement, IQ, gender and SES. These researchers hypothesize that existing knowledge structures are the key aspect that drives this relationship because of the fundamental role of ME knowledge for advancing math knowledge, overall (McNeil et al., 2019).

Overall, supporting students' acquisition of ME understanding is important for their later success in math. While second grade is where this topic is traditionally introduced as recommended by the Common Core State Standards, a secure understanding of the concept does not seem to develop until at least the fifth grade (Rittle-Johnson, 2011). Some researchers suggest that ME is a difficult concept for students to learn because of students' misinterpretation of the equal sign (McNeil, 2008; Knuth et al., 2006). Further, incorrect knowledge around the equal sign is thought to stem from the pattern that many textbooks and teachers use in teaching arithmetic where the computation always is done after the equal sign, which indicates to students an operational purpose (McNeil et al., 2006).

Gesture's Role in Facilitating ME Learning

As suggested by the previous research on ME learning, examining new ways to instruct students on this subject is necessary for changing the current trajectory that leaves a majority of students not fully comprehending this fundamental topic until late in the fifth grade (Rittle-Johnson, 2011). One possible solution for bolstering students' understanding of ME may literally lie in the hands of our educators. Researchers have found that teachers' gestures play an important role in shaping students' math learning in general (Goldin-Meadow et al., 1999). Many prior studies have found that gestures--purposeful movements of the hands during cognitive activity, such as speaking or problem solving (Alibali, 2005)—can be used to effectively facilitate students' understanding of ME (Goldin-Meadow, Cook & Singer, 2009; Novack et al., 2014; Singer & Goldin-Meadow, 2005).

Goldin-Meadow, Kim and Singer (1999) studied teachers' gesture use when teaching individual students mathematical equivalence. They found that when teachers used gestures to convey problem solving strategies, children were able to comprehend the task-relevant information and incorporate the information from the gestures into their own speech (Goldin-Meadow, Kim & Singer, 1999). In other words, the students' ability to ascertain the procedural information necessary to correctly solve ME problems was enhanced when teachers used gestures.

Singer and Goldin-Meadow (2005) looked more into exactly which kinds of gestures, accompanied by what kinds of speech resulted in the highest gains in ME learning in third and fourth grade students. In order to find what amount of strategies taught (either one or two) and combination of verbal and gestural instruction (either matching or not matching) would provide the greatest gains in ME learning, the researchers randomly assigned the participants into six conditions. Findings suggested that students in the condition where gestures were accompanied by verbal instructions that did not match learned the most (Singer & Goldin-Meadow, 2005).

But what about when students produce the gestures, too? Goldin-Meadow, Cook & Mitchell (2009) investigated if third and fourth grade students would learn ME better when asked to replicate the gestures used when they were taught. The researchers wanted to understand if ME understanding could be contingent upon producing (and, further, correctly producing) gestures. Findings of this study suggested that students who correctly replicated the gestures learned the most; however, students who replicated partially correct gestures still learned more than students who did not gesture at all. This finding was mediated by whether or not the gesture was successfully integrated into the students' speech (Goldin-Meadow, Cook & Mitchell, 2009).

All of the previously mentioned studies and more have established that both teachers' and students' gestures can support student ME understanding (Goldin-Meadow, Kim & Singer, 2009; Novak et al., 2014; Goldin-Meadow, Cook & Mitchell, 2009), but the details about the cognitive mechanisms underlying learning from teachers' gestures versus a students' own gestures are not well understood. Prior research seems to indicate that students' gestures-use supports learning by freeing up some of their working memory (WM) resources, aligning with ideas presented by Baddely's (2000) model of WM. Baddeley's (2000) model of WM, posits that information is processed in WM through two stores--the phonological loop (where language is processed) and the visuospatial sketchpad (where visual semantics are processed)--and then combined with long term memories in the episodic buffer for encoding into long term memory. When students use gestures in addition to verbalizing relevant information to support their learning, this activates both stores of their WM (the phonological loop and the visuospatial sketchpad)--according to Baddeley's model--which makes it a cognitive offloading mechanism.

The cognitive processes associated with students learning from watching the teachers' gestures have yet to be ascertained. Gestures convey information in a visuospatial nature (McNeill, 1992) which is different from the verbal information that is conveyed in speech (Kendon, 1980; McNeill, 1992; Morrel-Samuels & Krauss, 1992). Potentially, gestures could be useful to learners when teachers use them because they are using more than one modality to reinforce information. Some research has been conducted to evaluate whether or not teaching concepts with multiple learning modalities (i.e. visual, auditory, kinesthetic) is beneficial. Previous work confirms that a multimodal approach to teaching can be beneficial for teaching a variety of topics including fractions (Chahine, 2013) and complex software installation (Gellevij et al., 2002). This framework could explain the benefits of using speech and gesture. To my knowledge, a direct comparison of unimodal instruction (speech only) versus multimodal (speech and gesture) has not been examined in ME learning, specifically. Additionally, it is unknown whether the teachers' gestures accompanying speech free up WM resources for the student (i.e., the listener) in the same way as when the students are actively partaking in the gesturing.

Gesture-Use and English Language Learners (ELLs)

An additional detail that is unclear is the extent to which understanding the language used (phonological loop inputs) may impact the usefulness of the gesture. Many interventions for ELLs focus on their literacy skills, even interventions that address their math learning mainly focus on helping them learn to solve math word problems (examples: Orosco et al., 2013; Martinello, 2008; Driver & Powell, 2017). Due to this disparity, it is no surprise that students from low income and minority backgrounds are at an even greater risk for poor math achievement (Siegler et al., 2012). However, looking at gesture use in the ELL population is not an entirely new concept. Research by Wermelinger and colleagues (2020) suggests that preschool aged bilingual children produce more iconic gestures (gestures which indicate form or movement of some referent; McNeill, 1992) than their monolingual counterparts. The researchers found, first, that there was no difference in understanding iconic gestures between monolingual and bilingual students. They also found that bilingual students were able to better produce iconic gestures than their monolingual counterparts. These researchers theorize this effect is due to bilingual students possessing increased communicative sensitivity--or being more in tune with a conversation partner's needs and intentions (Wermelinger et al., 2020).

This still does not indicate the effect of the teacher's gestures on their learning in math. Perhaps if gestures aid learning by spreading WM resources over both memory stores (cognitive offloading), even when done by the teacher, then this intervention could prove to be just as beneficial for ELL students as non ELL students. The current study will not only examine the effects of gesture-use on all students, but will also look particularly at students from an ELL background.

Research Questions/Aims of this Study

Being that this is the first study to examine gesture-use in teaching ME in second-grade classrooms that have many ELLs, there are many unknowns. The methods that we have decided to use were established to accomplish three main goals. The research team hoped to (1) determine--with reasonable certainty--if our intervention was placed at an appropriate time in the existing curriculum, as all previous studies used older participants, (2) validate our measure within that age group--again, due to a lack of previous work with children of the age group we worked with and (3) test the length of our intervention with actual teachers in classrooms--a novel location for the intervention and experimentation.

Methods

Participants

Participants in this study were seven second grade teachers (all seven teachers were White, Non-Hispanic females) from two different elementary schools on the West Coast of the United States. Teachers were recommended for recruitment by curriculum specialists in the partnering school district. Teachers were recruited via email, and informed consent was obtained prior to their participation in the study.

After informed consent was obtained, 4 of the teachers were assigned to be in the experimental group and the remaining 3 were assigned to be in the control group. Each group had both veteran and early career teachers. Teachers received a $50 gift card for participating in this study.

Gesture-Training

Prior to teaching mathematical equivalence in their classrooms, experimental teachers attended a brief twenty minute training session on the role of gestures in supporting students' mathematics learning. Two training sessions were conducted -- one for each school. At this training, the research team presented previous research showing how the equalizer gesture supports students' understanding of mathematical equivalence (Goldin-Meadow, Kim & Singer, 1999; Goldin-Meadow, Cook, Mitchell, 2009; Singer & Goldin-Meadow, 2005). An image of the equalizer gestures is shown in Figure 1. Training sessions were conducted by the first and second authors of this study. The training provided teachers the opportunity to see the gesture displayed, practice how they would use it during instruction and ask any questions about using the gesture. Experimental teachers were instructed to incorporate the equalizer gesture into their instruction on mathematical equivalence.

Teacher Demographics and Background

At the training sessions, teachers were asked to fill out a brief demographic questionnaire. The questionnaire asked the following: (1) the teachers' gender, (2) amount of time teaching, (3) amount of time teaching elementary school and (4) degrees held. The questionnaire also included two questions about comfort and enjoyment in teaching math, specifically, with 5 point likert style responses.

Fidelity of Implementation

Observation-protocol trained members of the research team observed teachers in the classroom during one of their two lessons on ME and took detailed notes on gesture-use (or lack thereof in the case of teachers in the control group). Observations took place over the span of 30 minutes per each classroom. The form researchers used to complete fidelity of implementation has been provided in Appendix A.

Mathematical Equivalence Pretest, Posttest, and Practice Sheets

Researchers created a six-item ME Pre-Test and a similar six-item ME Post-Test. We referred to a previous study (Goldin-Meadow, Cook & Mitchell, 2009) to inform our creation of these tasks. Both of the tests were created to be of equal level of difficulty. Each test had 3 items in the a+b+c=_+c format and 3 items in the d+e+f=d+_ format in randomized order. Teachers were also given two worksheets of the same composition (six-item, same level of difficulty with items that were different) to use for student practice during teaching their ME lessons. Sample mathematical equivalence items are provided in Appendix B.

Procedure

Teachers in the experimental group were asked to use the equalizer gesture during instruction on mathematical equivalence. Teachers in the control group were instructed to teach mathematical equivalence using "business as usual" methods. Study activities took place over the span of four consecutive school days within which all teachers were instructed to teach about mathematical equivalence for two consecutive class "lessons" (about 30 minutes each). Teachers administered the pretest on day 1, which was followed by two days of mathematical equivalence instruction during which students completed the two sets of practice problems. Teachers administered the posttest on the fourth and final day of experimentation.

Results

All analyses were done with IBM SPSS Statistics version 24.

Student Data

After discarding data from students whose parents did not consent to their child participating in the study, data from 151 students was examined for data analysis. Of those 151 students, 58 were designated as having English Language Learner (ELL) status. Students who had a perfect score on the ME Pre-Test were not included in the analysis, (*n*=146) because we wanted our analysis to only look at students who had the chance to learn from the intervention. Additionally, work from students who did not complete both a ME Pre-Test, ME Post-Test and at least one practice worksheet were also taken out of the sample, leaving us data from 119 students to analyze. The number of students whose data was included in the analysis from each participating classroom, and the number of ELL students from each participating classroom is shown in Table 1.

Examining ELL and Non-ELL Performance

As indicated, we suspected that there may have been some difference in performance on the measures due to ELL status (Table 2 displays the distribution of both

Pre-Test and Post-Test scores for both ELL and non-ELL groups). In order to examine if this was true, we ran independent samples t-tests for the Pre-Test, the Post-Test and the overall gain in scores from pre to post. Results, as seen in Table 3, established that there were no significant differences between ELL students and non ELL students in Pre Test Scores, Post Test Scores or overall gains in scores.

Regression Analysis Examining if Gesture-Use Predicts ME Post Test

Regression analyses were conducted in order to determine if the use of the equivalence gesture during instruction predicts ME Post Test scores. Two models were examined in order to assess various potential additional covariates, after the initial analysis to determine the predictive nature of condition. Model 1 only looked at Pre-Test scores as a covariate in the model. We did this, as opposed to just using gain scores as the outcome measure, because this method of analysis, according to Griffin et al. (1999) is maximally informative. Results of this model, as shown in Table 4, indicated that condition was not a significant predictor for Post-Test scores.

In Model 2, Pre-Test Scores were included as a covariate and ELL status was included as a predictor. Similar to Model 1, we have Pre-Test scores in the analysis as a covariate. Additionally, we looked at ELL status as a predictor because we wanted to see if the gesture-use intervention would produce the same types of outcomes for students of ELL status as for students who are not--considering that nearly half of our population was in the ELL category. Results of this model, as shown in Table 4, indicated that ELL status did not significantly predict ME Post Test Score.

Further Analysis Examining the Relationship Between Teacher Background Data and Student Outcomes

The teachers in the sample had varying levels of teaching experience (from 1 year to 30 years) and comfort/enjoyment in teaching math. Pearson's correlations, as seen in Table 3, were calculated. According to Cohen's (1988) standards, teachers' experience and their enjoyment for teaching math were both negatively correlated with students' gain scores (ME Post Test Scores minus ME Pre-Test Scores) to a significant extent. Conversely, teachers' comfort in teaching math was not significantly correlated to students' gain scores.

Regression Analyses Examining if Teacher Enjoyment and Experience Predict Students' Overall Gain Scores

Correlational data indicated that there is a relationship between gain scores and teacher enjoyment of teaching math and teacher experience. In order to understand this relationship further, we conducted regression analyses to see if these teacher characteristics could predict students post test scores. Model 3 examines if teacher enjoyment of math and teacher experience are predictive of Post-Test scores. Similar to the previous analyses, Pre-Test scores were included as a covariate. Model 4 examines the same relationship as Model 3 but looks also at ELL status as a covariate. Results of these analyses, which can be seen in Table 6, indicate that both teacher enjoyment of teaching math and teacher experience are significantly predictive of post test scores; however, in Model 4 when ELL status was added into the model, teacher enjoyment became insignificant and ELL status was not a significant predictor.

To conclude, the data suggests that gesture-use does not significantly predict ME outcomes. Also, the analysis revealed that ELL status was not a predictor of ME Post Test scores, which is a positive indication that the students in our sample (including those in the gesture-use condition) were able to produce statistically equivalent gains to their non-ELL counterparts. Finally, our analyses indicated that varying teacher background characteristics may have had an effect on student outcomes.

Discussion

The results of this study did not provide sufficient evidence that using the equalizer gesture to teach ME would improve learning outcomes more than with traditional instruction in a classroom setting. Contrary to our hypothesis, students who received instruction on ME that included the gestures did not score significantly better on the ME Post-Test than the students who received instruction on ME in a business as usual format. However, we found that the performance of the ELL students versus the non-ELL students was not significantly different, overall, which was more aligned with our hypotheses. Upon examining the relationship between teacher characteristics and their students' post-test performance, teacher enjoyment in teaching math and teacher experience in teaching math were both negative predictors.

Many previous studies (Goldin-Meadow, Kim & Singer, 1999; Goldin-Meadow, Cook & Mitchell, 2009; Singer and Goldin-Meadow, 2005) set the precedent that teaching with gestures aid in student acquisition of ME problem solving. These studies provided the fundamental support for the use of gestures by teachers to better articulate information about how to solve ME problems in individualized instruction. Our findings did not corroborate the conclusions of these studies; however, our study was the first to attempt to determine whether or not gestures were more useful in teaching ME than a control *and* if it would be useful in classroom instruction, which could account for our differing results.

There are a few methodological reasons that could account for the lack of evidence supporting our initial hypothesis in the study; however, there are some positive implications that can be taken away from the results. Beyond just looking at the intervention and its effects on student learning, overall, there is an interesting finding when we look at student learning between students who are ELLs and those who are not. Additionally, there are some interesting findings about how teacher characteristics impacted the students' learning outcomes. Of the limitations in the methods of the study, one problematic component is the length and make up of our ME testing measures. One of the aims of this study was to work on validating our measure, because they were researcher-created, and to determine if the length of the measures was sufficient. Both the pre-test and the post-test were six items, which did not leave very much room for variation in scores. In order for our measure to be developmentally accessible to second grade students, it cannot be too long; however, it seems as though six items is not enough to achieve a standard/normal distribution of scores, which we were expecting to see. Additionally, all of the items were designed to be of equal difficulty, which may not have been the best strategy to determine a score that indicates level of ME understanding.

Another limitation imposed by our methods was that we had less statistical power to detect differences in our fairly small sample than we would have had if we used a different procedure. Moving forward the procedure will be changed in order to examine changes in score across three time points--pre test one, pre test two and post test--in which teachers are asked to teach business as usual after pre test one, then use the gesture intervention after pre test two. This will give us more statistical power because it is a between-subjects design.

A last limitation that we faced with the methods that is worth mentioning is that the business as usual control group did not prove to be an equal control across classrooms, as the three teachers in that condition all chose to use different teaching methods (that did not include the equalizer gesture). Moving forward, more systematic controls (and more methods to assess what cannot be controlled for) will be put in place to ensure a more accurate understanding of whether or not using gestures while teaching ME in the classroom is beneficial to student learning.

One thing to note about the results that was favorable to the hypotheses that we set out was that the data did not indicate that there was not any significant difference on ME post test scores between ELLs and non-ELLs (within the whole sample, experimental and control conditions combined). This seems to imply that using the gesture intervention was the same as traditional instruction for ELL students, which aligns with the hypothesis that the intervention would not benefit non-ELL students more than ELL students.

Finally, when looking at the background characteristics of the teachers to see if the differences between teachers experience, enjoyment of teaching mathematics and comfort in teaching math was related to gains in scores we noticed that there was significant negative correlations between (1) gains and teacher enjoyment in teaching math and (2) gains and teacher experience. Upon analyzing whether or not these factors were predictors of ME post test scores, they indeed were. To further clarify what these findings imply, we see that students of teachers who report more enjoyment of teaching math and more experience in teaching elementary math have smaller gains in ME learning. One reason that we think this relationship could occur might be that teachers who enjoy teaching math and who have more experience in teaching math are less open to change/teaching a concept in more or less depth than they normally would because they are a part of the study. However, both teacher experience and teacher enjoyment were assessed through self report on the Teacher Demographics and Background Questionnaire where one item was asked about each construct, which does not constitute this finding to be more than preliminary and something that we'd like to focus on more in the future to fully understand this relationship.

References

Alibali, M. W. (2005) Gesture in spatial cognition: expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation 5*(4), 307-331.

Alibali, M. W., Knuth, E. J., Hattikudur, S., McNeil, N. M., & Stephens, A. C. (2007). A longitudinal examination of middle school students' understanding of the equal sign and equivalent equations. *Mathematical Thinking and learning*, *9*(3), 221-247.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, *4*(11), 417-423.

Chahine, I. C. (2013). The impact of using multiple modalities on students' acquisition of fractional knowledge: An international study in embodied mathematics across semiotic cultures. *The Journal of Mathematical Behavior*, *32*(3), 434-449.

Charles, R. I., & Carmel, C. A. (2005). Big ideas and understandings as the foundation for elementary and middle school mathematics. *Journal of Mathematics Education*, *7*(3), 9-24.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Driver, M. K., & Powell, S. R. (2017). Culturally and linguistically responsive schema intervention: Improving word problem solving for English language learners with mathematics difficulty. *Learning Disability Quarterly*, *40*(1), 41-53.

Gellevij, M., Van Der Meij, H., De Jong, T., & Pieters, J. (2002). Multimodal versus unimodal instruction in a complex learning context. *The Journal of Experimental Education*, *70*(3), 215-239.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teachers hands tell the students mind about math. *Journal of Educational Psychology, 91*(4), 720-730.

Goldin-Meadow, S. Wagner Cook, S., Mitchell, Z. A. (2009) Gesturing gives children new ideas about math. Psychological Science *20*(3), 267-272

Griffin, D., Murray, S., & Gonzalez, R. (1999). Difference score correlations in relationship research: A conceptual primer. *Personal Relationships*, *6*(4), 505-518.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational studies in Mathematics*, *12*(3), 317-326.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the. *The relationship of verbal and nonverbal communication*, (25), 207.

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for research in Mathematics Education*, 297-312.

Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, *78*(2), 333-368.

McNeil, N. M., Grandau, L., Knuth, E. J., Alibali, M. W., Stephens, A. C., Hattikudur, S., & Krill, D. E. (2006). Middle-school students' understanding of the equal sign: The books they read can't help. *Cognition and instruction*, *24*(3), 367-385.

McNeil, N. M. (2008). Limitations to teaching children 2+2=4: typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development, 79*(5), 1524-1537.

McNeil, N. M., Hornburg, C. B., Devlin, B. L., Carrazza, C., & McKeever, M. O. (2019). Consequences of individual differences in children's formal understanding of mathematical equivalence. *Child development*, *90*(3), 940-956.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 615.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). Common Core State Standards. Washington, DC: Authors.

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science*, *25*(4), 903-910.

Orosco, M. J., Swanson, H. L., O'Connor, R., & Lussier, C. (2013). The effects of dynamic strategic math on English language learners' word problem solving. *The Journal of Special Education*, *47*(2), 96-107.

Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology*, *103*(1), 85.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological science*, *23*(7), 691-697.

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science*, *16*(2), 85-89.

Wermelinger, S., Gampe, A., Helbling, N., & Daum, M. M. (2020). Do you understand what I want to tell you? Early sensitivity in bilinguals' iconic gesture perception and production. *Developmental Science*, e12943.

Figure 1.

Equalizer Gesture

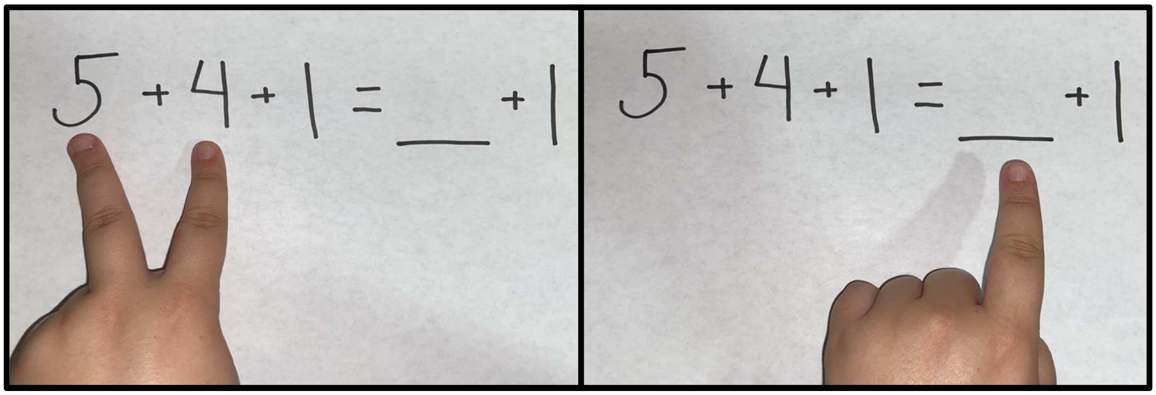

Table 1.

Sample Distribution Amongst Classrooms

Classroom Number	Condition	*N*	ELL	Pre Test Score *M*	Post Test Score *M*
1	Control	13	7	0.23 (0.83)	5.46 (0.78)
2	Experimental	19	8	0.58 (1.12)	3.89 (2.36)
3	Control	18	7	0.33 (1.19)	4.89 (1.81)
4	Experimental	17	10	0.29 (0.69)	5.41 (0.87)
5	Control	19	11	1.68 (1.34)	4.05 (2.12)
6	Experimental	18	10	0.00 (0.00)	2.39 (2.66)
7	Experimental	15	5	0.47 (1.25)	4.73 (2.12)
Total		119	58	0.54 (1.13)	4.34 (2.17)

Note: Standard deviations are listed in parentheses.

Table 2.

Distribution of ELL students' scores versus non-ELL students' scores

Group	*N*	Pre Test *M*	Post Test *M*
ELL	58	0.57 (1.23)	4.33 (2.17)
Non-ELL	61	0.50 (1.01)	4.34 (2.19)

Note: Standard deviations are listed in parentheses.

Table 3.

Results of t-tests Examining Differences between ELL and non-ELL Students

	ELL		Non-ELL		
Measure	*M*	*SD*	*M*	*SD*	*t*
Pre Test	.57	1.23	.50	1.01	.356
Post Test	4.33	2.17	4.34	2.19	-0.042
Overall Gain	3.75	2.36	3.84	2.27	-0.214

Note. All p's > .05.

Table 4.

Linear Regression Models Examining whether Condition and ELL Status Predict Post Test Scores

	Model 1	Model 2
Intercept	4.547*** (0.34)	4.536*** (0.40)
Condition	-0.559 (0.41)	-0.559 (0.41)
Pre Test	-0.211 (0.18)	0.212 (0.18)
ELL Status		0.021 (0.40)
R^2	0.03	0.03

Note. Numbers in parentheses are standard errors. Outcome measure used in both models was Post Test Score. **p*<0.05, ***p*<.01, ****p*<.001

Table 5.

Correlational Data Among Teacher Background and Gain Scores

Variable	1	2	3	4
1. Gain Score				
2. Teacher Experience	-0.293**			
3. Teacher Comfort	0.107	0.605***		
4. Teacher Enjoyment	-0.228*	0.539**	.511**	

Note: *p<0.05, **p<.01, ***p<.001

Table 6

Linear Regression Models Examining whether Teacher Enjoyment and Experience Predict Student Post Test Scores

	Model 3	Model 4
Intercept	0.582 (2.30)	0.58 (2.31)
Teacher Enjoyment	1.04 (0.52)*	1.04 (0.52)
Teacher Experience	-0.08 (0.02)**	-0.08 (0.02)**
Pre Test	.469 (0.19)*	0.40 (0.18)*
ELL Status		-0.01 (0.38)
R^2	0.11	0.11

Note: Unstandardized coefficients reported. Numbers in parentheses are standard errors. Outcome measure used was Post Test Score.*$p<0.05$, **$p<.01$, ***$p<.001$

Appendix A -- Teacher Observation Sheet

Participant: ___________________________

Date of Observation: _______________________

1. Does the teacher use gestures when explaining mathematical equivalence?

2. Does the teacher use the **combining** gesture when explaining mathematical equivalence?

3. Does he/she use the combining gesture correctly?

4. If they do not use the combining gesture, what kind of gesture does he/she use when teaching the topic?

5. Does the teacher ask the students to use gestures when completing class work on mathematical equivalence?

6. Does the teacher instruct students on using the *combining* gesture when completing class work? (meaning does the teacher directly address using the combining gesture)

7. How long did the teacher spend teaching the concept?

8. How long did the teacher designate for students to complete classwork pertinent to the concept

9. Other notes:

Statement	1	2	3	4	5 or more
How many times does the teacher use gestures when explaining mathematical equivalence?					
How many times does the teacher use the **combining** gesture when explaining material?					
How many times does the teacher encourage the students to use the combining gesture when solving the problems?					

Appendix B

Sample Mathematical Equivalence Problems

$2 + 1 + 3 = \_\_ + 3$

$2 + 2 + 4 = \_\_ + 4$

$5 + 4 + 1 = 5 + \_\_$

$1 + 4 + 5 = \_\_ + 5$

$2 + 5 + 3 = 2 + \_\_$

$6 + 4 + 2 = 6 + \_\_$