REDIMAT
Journal of Research in Mathematics Education

Hipatia Press
www.hipatiapress.com

Instructions for authors, subscriptions and further details:

http://redimat.hipatiapress.com

Predictors of Achievement When Virtual Manipulatives are Used for Mathematics Instruction

Patricia S. Moyer-Packenham[1], Joseph Baker[2], Arla Westenskow[1], Katie L. Anderson-Pence[3], Jessica F. Shumway[1], Kerry E. Jordan[1]

1) Utah State University, United States of America.
2) Stanford University, United States of America.
3) University of Colorado-Colorado Springs, United States of America.

Date of publication: June 24th, 2014
Edition period: June 2014-October 2014

To cite this article: Moyer-Packenham, P.S., Baker, J., Westenskow, A., Anderson-Pence, K.L., Shumway, J.F., and Jordan, K.E. (2014). Predictors of Achievement When Virtual Manipulatives are Used for Mathematics Instruction. *REDIMAT*, *Vol 3*(2), 121-150. doi: 10.4471/redimat.2014.46

To link this article: http://dx.doi.org/10.4471/redimat.2014.46

REDIMAT, Vol. 3 No. 2 June 2014 pp. 121-150

Predictors of Achievement When Virtual Manipulatives are Used for Mathematics Instruction

Patricia S. Moyer-Packenham
Utah State University

Joseph Baker
Stanford University

Arla Westenskow
Utah State University

Katie L. Anderson-Pence
University of Colorado – Colorado Springs

Jessica F. Shumway
Utah State University

Kerry E. Jordan
Utah State University

(*Received: 14 October 2013; Accepted: 9 June 2014; Published: 24 June 2014)*

Abstract

The purpose of this study was to determine variables that predict performance when virtual manipulatives are used for mathematics instruction. This study used a quasi-experimental design. This design was used to determine variables that predict student performance on tests of fraction knowledge for third- and fourth-grade students in two treatment groups: classroom instruction using texts and physical manipulatives (CI), and computer lab instruction using virtual fraction applets (VM). The Pre-test, Post-test 1, and Post-test 2 measured learning and retention of fraction concepts. Observation ethograms documented representation use. The results revealed that fewer demographic predictors of student performance (e.g., socio-economic status, English language learner status, and gender) exist during fraction instruction when virtual manipulatives were used. When instructors used virtual manipulatives, there was an equalizing effect on achievement in third and fourth grade classrooms, in that fewer demographic factors were influential for VM groups compared to CI groups.

Keywords: Virtual manipulatives, fraction learning, grade 3 and 4

2014 Hipatia Press
ISSN: 2014-3621
DOI: 10.4471/redimat.2014.46

REDIMAT, Vol. 3 No. 2 Junio 2014 pp. 121-150

Predictores de los Logros Cuando se Utilizan Materiales Virtuales para la Enseñanza de las Matemáticas

Patricia S. Moyer-Packenham
Utah State University

Joseph Baker
Stanford University

Arla Westenskow
Utah State University

Katie L. Anderson-Pence
University of Colorado – Colorado Springs

Jessica F. Shumway
Utah State University

Kerry E. Jordan
Utah State University

(*Recibido: 14 Octubre 2013; Aceptado: 9 Junio 2014; Publicado: 24 Junio 2014)*

Resumen

El propósito de este artículo fue cómo determinar las variables que predicen el aprendizaje cuando se utilizan materiales virtuales para enseñar matemáticas. En este estudio utilizamos una metodología cuasi-experimental. El diseño se utilizó para determinar variables que predicen el aprendizaje de los estudiantes en tests sobre conocimiento de fracciones con estudiantes de tercero y cuarto en dos grupos: uno donde se utilizaron textos y materiales físicos (CI) y otro donde se usaron applets virtuales (VM). El pre-test, post-test 1 y post-test 2 midieron el aprendizaje y retención de los conceptos de fracciones. Se usaron etnogramas para documentar el uso de las representaciones. Los resultados rebelan que los predictores demográficos (i.e. estatus socio-económico, conocimiento del inglés, género) tienen poca incidencia cuando se usan materiales virtuales. El uso de dichos materiales produce un efecto nivelador entre ambos grupos.

Palabras clave: Materiales virtuales, aprendizaje de fracciones, grados 3 y 4

2014 Hipatia Press
ISSN: 2014-3621
DOI: 10.4471/redimat.2014.46

A virtual manipulative (VM) is defined as "an interactive, web-based visual representation of a dynamic object that presents opportunities for constructing mathematical knowledge," (Moyer, Bolyard & Spikell, 2002, p. 373). In the past two decades, numerous studies have shown teachers using virtual manipulatives in a variety of ways (Beck & Huse, 2007; Bolyard & Moyer, 2003; Moyer & Bolyard, 2002; Moyer, Bolyard & Spikell, 2001; Moyer, Niezgoda, & Stanley, 2005; Moyer-Packenham, 2005; Moyer-Packenham, Salkind, Bolyard & Suh, 2013; Suh & Moyer, 2007; Suh, Moyer, & Heo, 2005) and research on virtual manipulatives has produced mixed results (Baturo, Cooper, & Thompson, 2003; Clements, Battista & Sarama, 2001; Deliyianni, Michael, & Pitta-Pantazi, 2006; Haistings, 2009; Highfield & Mulligan, 2007; Izdorczak, 2003; Moyer-Packenham & Suh, 2012; Steen, Brooks & Lyon, 2006; Takahashi, 2002). A recent meta-analysis comparing the effects of virtual manipulatives on student achievement using 32 research reports and 82 effect size scores reported that virtual manipulatives produced an overall moderate effect (0.34) on student achievement when compared with other instructional treatments (Moyer-Packenham & Westenskow, 2013). In contrast, one of the largest studies comparing virtual manipulatives with physical manipulatives and text-based materials in third and fourth grade demonstrated no significant differences in achievement between the treatments (Moyer-Packenham, Baker, Westenskow, Anderson, Shumway, Rodzon, & Jordan, The Virtual Manipulatives Research Group at Utah State University, 2013). The results of previous studies demonstrate that the relationship between virtual manipulatives and student achievement still remains an important area for further study.

The research reported in this paper is part of a larger research project in which researchers examined the use of virtual manipulatives for mathematics instruction in third- and fourth-grade classrooms. Multiple papers were produced to examine the large amount of data produced in the larger study (Anderson-Pence, Moyer-Packenham, Westenskow, Shumway, & Jordan, in press; Moyer-Packenham et al., 2013; Westenskow, Moyer-Packenham, Anderson-Pence, Shumway, & Jordan, 2014). The first of the two larger complementary papers reports on a comparison between physical manipulatives and virtual manipulatives and focuses on student achievement results in the study (Moyer-Packenham et al., 2013). This present paper is the second of the two complementary papers, and here we focus on reporting results about those variables that were predictors of student achievement in the larger study. It is not the intention of the authors to repeat much of what has already been discussed in the first paper about the achievement results, but to build on those results and examine variables that predict achievement results when virtual manipulatives are used for mathematics instruction.

Previous Research on Virtual Manipulatives

Virtual manipulatives are typically considered "cognitive technology tools" (Pea, 1985), and as such, are considered a "medium that helps transcend the limitations of the mind, such as memory, in activities of thinking, learning, and problem solving" (p. 168). As cognitive technology tools, virtual manipulatives may change how students approach mathematical tasks. For example, a virtual manipulative representing the multiplication of fractions (see Fractions – Rectangle Multiplication on http://nlvm.usu.edu) allows students to immediately observe the consequences of changing the numerator or denominator of either factor (See Figure 1.) By providing simultaneous representations in pictorial and symbolic forms, virtual manipulatives provide a different type of mathematics experience that interacting with physical manipulatives or with text-based materials.

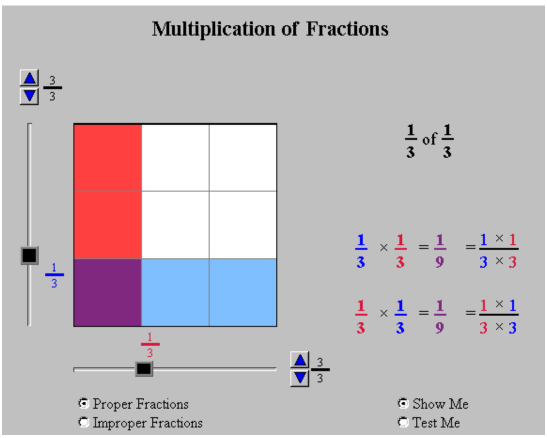


Figure 1. Example of fractions –rectangle multiplication at http://nlvm.usu.edu

Among the studies conducted on student achievement and virtual manipulatives, there are 12 reported results in favor of virtual manipulatives. These studies include: 65 Pre-K to first graders in the domain of partitioning (Manches, O'Malley, & Benford, 2010), 68 Pre-K to second graders in four mathematics domains (number, geometry, measurement, and patterns) (Clements & Sarama, 2007); 31 first graders (Steen, Brooks, & Lyon, 2006), 48 third graders (Clements & Battista, 1989), 560 eighth graders (Pleet, 1991), and 194 tenth graders in the domain of geometry (Hauptman, 2010); 32 third graders in the domain of measurement (Daghestani, Al-Nuaim & Al-Mshat, 2004), 91 fourth graders in the domain of fractions (Ball, 1988), 89 sixth and eighth graders in the domain of integers (Smith, 1995), 47 (Cavanaugh, Billan, & Bosnick, 2008) and 34 (Guevara, 2009) ninth through twelfth graders in the domain of algebra, and 48 university pre-service teachers in the domain of fractions (Lin, 2010). Together, these 12 studies represent all of the studies conducted to date comparing virtual manipulatives with other instructional methods in which there were significant results reported in favor of the virtual manipulatives.

As these studies show, the research on virtual manipulatives has included children at different grade levels, different mathematical domains, different numbers of participants, and different instructional methods. There could be many different variables that contribute to the results obtained in these studies. Most of these studies focus on the achievement results when different treatments were used with virtual manipulatives. However they do not focus on what factors may have predicted those achievement results. Therefore the research reported in this paper adds a new dimension to the research on virtual manipulatives by investigating what variables might predict the outcomes obtained when virtual manipulatives are used for mathematics instruction.

Research Questions

The study compared two instructional treatments – Treatment 1: classroom instruction using physical manipulatives and text-based materials (CI), and Treatment 2: computer lab instruction using virtual manipulatives (VM). We examined a number of different variables to determine whether or not the variable predicted student achievement in a fraction unit with third- and fourth-grade students. Our research questions in the present study looked beyond student achievement and focused on predictors of student performance.

1. Which demographic variables predict third- and fourth-grade student achievement, learning, and retention during fraction instruction in two treatments (CI and VM) as indicated by scores on a Pre-test, Post-test 1, and Post-test 2 (a delayed post-test)? Within the context of this broad research question, we addressed the following sub-question: Does student *gender*, *race*, *objective ability*, *subjective ability*, *Socio-Economic Status (SES)*, and *English Language Learner (ELL) status* predict fraction achievement, learning, and retention in either CI or VM classrooms?

2. Do mathematics representations, used during instruction or appearing on test items, modulate student achievement in CI and VM classrooms? Within the context of this broad research question, we addressed the following sub-questions: a) Do students in CI and VM classrooms score differently on symbolic, pictorial, and combined question types? b) Do students in CI and VM classrooms score differently on assessments based on their use of different types of representations (e.g., pictorial, symbolic, manipulatives) in different participation settings (e.g., individual vs. group)?

Methods

Participants

Third graders. A total of 156 third-grade students from eight public school classrooms located in four different elementary schools in two school districts participated in the study. Third-grade students were 46% male/54% female, predominantly Caucasian (75%; 14% Hispanic, 4% Mixed Race, 3% African American, 3% Asian, 1% Pacific Islander,), with 42% of students living in low Socio-Economic Status (SES) households. SES households are defined by the school district as those that receive free and reduced lunch services. A pre-test identified the group as 14% low-, 48% average-, and 38% high ability. Only 4% of the third graders were identified as English Language Learners (ELL). ELL students are defined as those receiving services from the school district to support English language development.

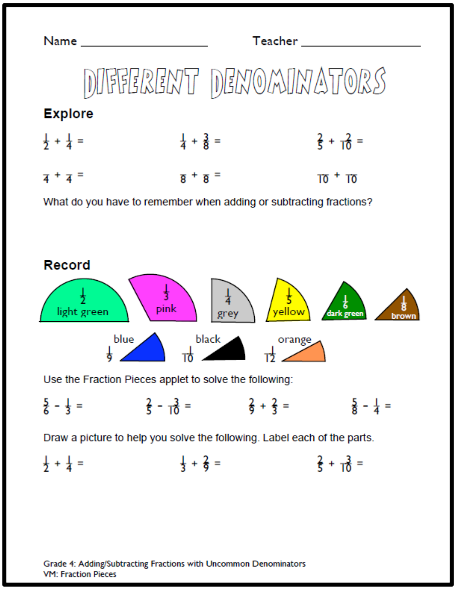
Fourth graders. A total of 194 fourth-grade students from nine public school classrooms located in six different elementary schools in two school districts participated in the study. Fourth-grade students were 48% male/52% female, predominantly Caucasian (78%; 14% Hispanic, 5% Mixed Race, 1% Pacific Islander, 1% Asian), with 54% of students living in low SES households. A pre-test identified the group as 21% low-, 49% average-, and 30% high ability. There were 8% of fourth graders identified as ELL.

Instructional Treatment Groups

Researchers used within-class random assignment to assign students to one of two treatment groups: Treatment 1: classroom instruction using physical manipulatives and text-based materials (CI), and Treatment 2: computer lab instruction using virtual manipulatives (VM).

Classroom instruction (CI) treatment groups. The CI treatment groups participated in classroom instruction using physical manipulatives and text-based materials. Seven third-grade public school classroom teachers taught the eight third-grade CI classes. Seven fourth-grade public school teachers taught the nine fourth-grade CI classes. Third-grade teacher experience ranged from 5 to 23 years (*M*=17.3); Fourth-grade teacher experience ranged from 3 to 32 years (*M*=14.6). All teachers (except two) taught third or fourth grade for three years or more. The CI teachers used Pearson SuccessNet curriculum materials (Scott Foresman/Addison Wesley Mathematics 2005 textbook) to teach the fraction unit. SMART Board™ technology was used during mathematics instruction in 50% of third-grade and 89% of fourth-grade classrooms. Teachers used manipulatives, worksheets, and teacher-created resources during instruction. The teachers and classrooms in the CI group did not use virtual manipulatives during the study.

Virtual manipulatives (VM) treatment groups. The VM treatment groups participated in computer lab instruction where teachers introduced concepts to students who were working at their own individual computers using virtual manipulatives. Four individuals from the local university taught the third- and fourth-grade VM groups, including three doctoral-level university graduate students and one university faculty member. The public school teaching experience of the university-based teachers ranged from 7 to 30 years of experience (M=13.7), which was similar to the experience of the CI teachers. Three of the VM instructors had prior public school experience teaching third and fourth grade. The VM teachers also used the Pearson SuccessNet curriculum program (Scott Foresman/Addison Wesley Mathematics 2005 textbook) and VM instructor-developed task sheets specifically designed to teach fraction concepts using the virtual manipulatives. These task sheets were designed to mirror the mathematical content being taught to the CI group, with tasks specific to problem exploration using the virtual manipulatives. These lesson materials were evaluated by an expert group of experienced teachers to determine the mathematical content match between CI and VM lessons. The lesson materials were piloted, reviewed and revised in preparation for the research project. (See Figure 2 for a sample task sheet.)


Name ____________ Teacher ____________

DIFFERENT DENOMINATORS

Explore

$\frac{1}{2} + \frac{1}{4} =$ $\frac{1}{4} + \frac{3}{8} =$ $\frac{2}{5} + \frac{2}{10} =$

What do you have to remember when adding or subtracting fractions?

Record

Use the Fraction Pieces applet to solve the following:

$\frac{5}{6} - \frac{1}{3} =$ $\frac{2}{5} - \frac{3}{10} =$ $\frac{2}{9} + \frac{2}{3} =$ $\frac{5}{8} - \frac{1}{4} =$

Draw a picture to help you solve the following. Label each of the parts.

$\frac{1}{2} + \frac{1}{4} =$ $\frac{1}{3} + \frac{2}{9} =$ $\frac{2}{5} + \frac{3}{10} =$

Grade 4: Adding/Subtracting Fractions with Uncommon Denominators
VM: Fraction Pieces

Figure 2. Sample virtual manipulatives task sheet created for the study

Procedures for Third –and Fourth–Grade Treatment Groups

At the beginning of the study, each classroom teacher reported demographic information including: gender, race, English Language

Learner (ELL) status, Socio-Economic Status (SES), and two measures of mathematical ability: subjective and objective. Teachers subjectively rated students' mathematical ability as high, medium, or low based on their observations of students' prior performance in mathematics. Objective ability was established by comparing each student's pre-test score to the class pre-test average and standard deviation. Standardized scores on the pre-test, that were one standard deviation or more above the mean, were classified as *high*, while scores one standard deviation or more below the mean were classified as *low*. All other scores were classified as *average*. At the conclusion of the fraction unit, teachers identified any students absent for more than 40% of the fraction unit and we removed these students from data analyses.

For a period of two to three weeks during the study, third- and fourth-grade students participated in the study of fractions in the CI or the VM groups. The CI students learned fraction content in a regular public school classroom setting, sitting at their desks, and using manipulatives and text-based materials. The CI teachers introduced the lesson concepts and children worked on the mathematics concepts individually or in small groups. The VM students also learned fraction content in a regular public school setting, in a computer classroom where each child was seated at their own computer. The VM teachers introduced the lesson concepts and how to use specific virtual manipulatives and children worked independently using the virtual manipulatives to complete a variety of mathematical tasks using the VM task sheets.

To ensure instructional fidelity across treatment groups, each paired CI teacher and VM teacher met prior to beginning instruction to specify the number of days allotted for the fraction unit and to correlate their lessons with the state's mathematics guidelines. The goal was to ensure that students would receive instruction on similar content regardless of their treatment groups. Each CI and VM teacher met daily to discuss plans for the following day to further ensure conformity between the two groups. During some meetings, teachers determined that students were struggling with a particular concept and together, the instructors decided to re-teach a concept. The purpose of this frequent, daily check-in was to ensure that students were learning the same mathematical content on each day of the study.

The length of the fraction units in the eight third-grade classrooms ranged from 9 to 12 days (average of 10.1 days) and in the nine fourth-grade classrooms ranged from 10 to 17 days (average of 11.4 days). Classroom teachers administered all pre- and post-tests to students in both treatment groups. All instruction occurred during regularly scheduled mathematics classes in the classrooms (CI groups) and computer labs (VM groups) of the participating schools. Students in the VM treatment spent almost every day of the fraction unit in the computer lab using their own individual computers for approximately 50 minutes each day.

Third-grade lessons addressed the following concepts: identify the denominator of a fraction as the number of equal parts of the unit whole and the numerator of a fraction as the number of equal parts being considered; define regions and sets of objects as a whole and divide the whole into equal parts using a variety of objects, models and illustrations; name and write a fraction to represent a portion of a unit whole for halves, thirds, fourths, sixths, and eighths; place fractions on the number line and compare and order fractions using models, pictures, the number line and symbols; find equivalent fractions using concrete and pictorial representations. Fourth-grade lessons addressed the following concepts: divide regions, lengths, and sets of objects into equal parts using a variety of models and illustrations; name and write a fraction to represent a portion of a unit whole length or set for halves, thirds, fourths, fifths, sixths, eighths, and tenths; generate equivalent fractions and simplify fractions using models, pictures, and symbols; order simple fractions; use models to add and subtract simple fractions where one single digit denominator is one, two, or three times the other; add and subtract simple fractions where one single digit denominator is one, two, or three times the other.

Instruments

Two primary instruments were used to collect data during the study: mathematics content tests and observation ethograms.

Pre- and post-tests. Third-grade tests contained 13 items total with two different types of test items: *pictorial* items (i.e., pictorial models with a written question stem) and *combined* items (i.e., numerals and operations combined with pictorial models with a written question stem). Fourth-grade tests contained 19 items total with three different types of test items:

pictorial items, *combined* items, and *symbolic* items (i.e., numerals and operations only). Test questions were selected to: (1) align with the third- and fourth-grade objectives for fractions in the state's mathematics curriculum standards; (2) represent a wide range of difficulty levels to differentiate among students; and (3) vary the types of representations used (i.e., pictorial, symbolic, and combined). During development, 27 multiple-choice questions and 3 open-ended questions were compiled into one form and checked for content validity by five mathematics educators with elementary school experience. Items were piloted with over 500 students from 23 classrooms in 14 elementary schools in 6 school districts prior to the study. A complete description of the development and validation of these instruments is discussed in the complementary paper (Moyer-Packenham et al., 2013). We administered the pre-test immediately prior to the fraction unit in each classroom, Post-test 1 on the day after the fraction unit concluded, and Post-test 2 seven weeks after the conclusion of the fraction unit.

Observation ethograms. Observation ethograms documented instruction in each of the classrooms throughout the study. The purpose for these observations was to determine if there were variables that predicted achievement that could be observed during instruction. Observers visited and observed 70% of the lessons using a modified ethogram protocol. Ethograms are instruments most often used by animal behavior researchers to efficiently and accurately describe the frequency and duration of behaviors made by a species observed in the field without any evaluation of the observed behaviors (MacNulty, Mech, & Smith, 2007). The use of this instrument results in a cohesive inventory of behavioral patterns that describe what a particular species spends its time doing in a particular environment. Using an ethogram modified for naturalistic classroom observations of humans, observers recorded the types of representations used by teachers and students at 5-minute intervals during observations. Observers specifically recorded information on mathematical content presented, terminology used, mathematical procedures presented, use of pictorial models, use of symbolic models, use of physical models, use of virtual manipulatives, and students' access to manipulatives (i.e., passive group viewing or active individual manipulation). The four VM teachers documented their use of representations in each lesson using instructor logs. The logs were subsequently coded and converted to an ethogram protocol.

The ethograms provided a quantitative measure of students' exposure and access to various fraction concepts, terminology, and representations in each of the treatment groups (CI and VM).

Results

As described previously in this article, the information described in this paper is part of a larger study in which student achievement results were examined. In the complementary paper to this one, there were no statistically significant differences between the VM and CI groups in terms of student achievement on the pre-, post-, and delayed post-tests (Moyer-Packenham et al., 2013). To provide a context for the predictor variables presented in this present paper, we have included (in Appendix A and B) the achievement outcomes for each class that were reported in the Moyer-Packenham et al. (2013) paper. In the results that follow, we discuss the variables that were predictors of student achievement in our study. Essentially these predictors are hidden behind the non-significant statistical results when achievement results are presented.

RQ#1: Post-test 1 and 2, Learning and Retention Analyses

Our first research question focused on demographic variables that predict student achievement, learning, and retention. Data on student demographics, ability ratings, and descriptions of the instructional environment were used as variables in multiple linear regression analyses to determine how well each variable predicted student test scores, and which variables combined to provide the best fitting predictive model. Using an alpha of .05 for each analysis, we conducted 24 linear regression analyses at Grade 3 and 24 at Grade 4 to describe the relationship between the demographic variables and students' test scores (pre-test, Post-test 1, and Post-test 2), students' learning (defined as the difference between Post-test 1 and Pre-test scores), and students' retention (defined as the difference between Post-test 2 and Post-test 1 scores). Table 1 displays a summary of the regression results for Grades 3 and 4 based on the mediating variables.

RQ#1: third-grade analysis. The scores for the VM and CI groups were split to investigate the effects of our predicting variables for third grade. For the CI group subjective ability, gender, SES, and ELL status, were significant predictors of Post-test 1 scores, $F(4, 82)= 9.703$, $MSE= 3124.657$, $p< .000$, $R^2= .332$, while subjective ability was the only significant predictor of Post-test 1 scores for the VM group, $F(1,72)= 6.186$, $MSE= 1768.818$, $p= .015$, $R^2= .080$. The best-fitting model for the third-grade CI group on Post-test 2 included subjective ability and ELL status, $F(2, 82)= 30.143$, $MSE= 7574.587$, $p< .000$, $R^2= .43$, while the model for the VM group consisted only of subjective ability, $F(1, 72)= 30.972$, $MSE= 8683.746$, $p< .000$, $R^2= .304$. Next, we calculated "learning" scores for each student, subtracting Pre-test scores from Post-test 1 scores. For the third-grade CI group, the model included gender, $F(1, 82)= 4.253$, $MSE= 1773.269$, $p= .042$, $R^2= .050$, while the model for the VM group included subjective ability, $F(1, 72)= 6.362$, $MSE= 2589.249$, $p= .014$, $R^2= .082$.

Table 1
Linear regression analysis of predictive variables: Test results for grade 3 and 4

	Treatment group	Subjective Ability	Objective Ability	Gender	SES	ELL	Race
Post-test 1	CI	Δ, □		Δ	Δ, □	Δ	
	VM	Δ					□
Post-test 2	CI	Δ, □			□	Δ	
	VM	Δ					
Learning	CI			Δ			
	VM	Δ					□
Retention	CI		Δ	Δ			
	VM	Δ			□		□

Note. Δ indicates that the variable was a significant predictor of students' test scores for Grade 3; □ indicates that the variable was a significant predictor of students' test scores for Grade 4.

These results suggest that students' gender significantly affected performance in the CI group, while subjective ability affected performance in the VM group. Next we calculated a "retention" score to describe the amount of fraction content retained between Post-test 1 and Post-test 2. A positive retention score indicates that information was gained between post-tests. For the third-grade CI group, the best-fitting model included objective ability and gender, $F(2, 82)= 5.305$, $MSE= 2115.710$, $p= .007$, $R^2= .117$, while the model for the VM group included only subjective ability, $F(1, 72)= 7.098$, $MSE= 2614.206$, $p= .010$, $R^2= .091$. This indicates that objective ability and gender significantly predicts long-term retention of fraction concepts in the CI group. These data show a trend in which a broader range of demographic variables influence students' performance in the CI group compared to the VM group.

RQ#1 fourth-grade analyses. Next we conducted the same analyses in Grade 4 to describe the relationship between demographic variables and students' test scores, learning, and retention. The scores for the VM and CI groups were split to investigate the effects of our predicting variables for fourth grade. For the CI group the best fitting model for Post-test 1 consisted of subjective ability and SES, $F(2, 102)= 9.027$, $MSE= 3210.881$, $p< .000$, $R^2= .153$), while the model for the VM group included only race, $F(1, 90)= 7.865$, $MSE= 3934.904$, $p= .006$, $R^2= .081$. The preferred model for the CI group alone for predicting Post-test 2 scores included subjective ability and SES, $F(2, 102)= 14.124$, $MSE= 3444.874$, $p< .000$, $R^2= .22$; there were no significant predictors of Post-test 2 for the VM group. For the CI group, there were no significant predictors for learning in Grade 4, while the best fitting model for the VM group included race, $F(1, 72)= 8.244$, $MSE= 3005.143$, $p= .005$, $R^2= .085$. For the CI group, there were no significant predictors for retention in Grade 4, while the best fitting model for the VM group included SES and race, $F(2, 72)= 9.576$, $MSE= 2293.532$, $p< .000$, $R^2= .179$. These results indicate that race and SES significantly predict fourth-grade students' retention scores in the VM but not the CI groups.

RQ#2: Test Item Question Type Analyses

The first part of Research Question #2 focused on relationships between demographic variables and test item question types (e.g., pictorial items, symbolic items, or combined items). To investigate the effect of our predicting variables on the question types that appeared on the third- and fourth-grade tests, we conducted linear regression analyses on one variable while controlling for the others. Essentially this examination helped us to determine if certain variables predict how well a student will do on different types of test items (pictorial items, symbolic items, or combined items). See Table 2.

Table 2

Linear regression analysis of predictive variables: Question types in grade 3 and 4

	Treatment group	Subjective ability	Objective ability	Gender	SES	ELL	Race
Post 1:	CI	Δ, □			Δ	Δ	
Pictorial	VM	Δ			□		
Post 1:	CI				□		
Symbolic	VM						□
Post 1:	CI	Δ, □		Δ		Δ	
Combined	VM	□				Δ	□
Post 2:	CI	Δ			□	Δ	
Pictorial	VM	Δ					
Post 2:	CI	□			□		
Symbolic	VM						
Post 2:	CI	Δ, □		□	□	Δ	Δ
Combined	VM	Δ	Δ				

Note. Δ indicates that the variable was a significant predictor of students' test scores for Grade 3; □ indicates that the variable was a significant predictor of students' test scores for Grade 4.

Third-grade analyses. For Post-test 1 pictorial questions alone, the best-fitting model for the CI group consisted of subjective ability, SES, and ELL status, $F(3, 82)= 9.897$, $MSE= 4994.049$, $p< .000$, $R^2= .273$, while the best-fitting model for the VM group included only subjective ability, $F(1,72)=4.572$, $MSE= 1522.539$, $p= .036$, $R^2= .060$. For the combined question type (e.g., numerals and operations combined with pictorial models with a written question stem), the preferred model for the CI group on Post-test 1 consisted of subjective ability, gender, and ELL status, $F(3, 82) = 8.369$, $MSE = 3023.058$, $p < .000$, $R^2 = .241$, while the preferred model for the VM group included ELL only, $F(1, 72) = 5.483$, $MSE = 3265.207$, $p = .022$, $R^2 = .072$. For the third-grade CI group, the preferred model for pictorial question types on Post-test 2 included subjective ability and ELL status, $F(2,82) = 23.817$, $MSE = 7824.735$, $p < .000$, $R^2 = .373$, while the model for the VM group included only subjective ability, $F(1, 72) = 29.907$, $MSE = 11787.375$, $p < .000$, $R^2 = .296$. For the CI group, the best-fitting model for the combined question types included subjective ability, ELL status, and race, $F(3, 82) = 13.589$, $MSE = 9693.585$, $p < .000$, $R^2 = .340$, while the model for the VM group included subjective and objective ability, $F(2, 72) = 7.923$, $MSE = 6048.99$, $p = .001$, $R^2 = .185$. Consistent throughout our analyses, results show fewer significant predicting variables in the VM group (compared to the CI group), suggesting that fewer demographic factors mediate students' performance when fraction concepts are taught using virtual manipulatives.

Fourth-grade analyses. The best fitting model for pictorial question types for the CI group on Post-test 1 included only subjective ability, $F(1, 102) = 12.166$, $MSE = 13490.445$, $p = .001$, $R^2 = .108$, while the preferred model for the VM group consisted of only SES, $F(1, 90) = 10.037$, $MSE = 11761.01$, $p = .002$, $R^2 = .101$. For the symbolic question types on Post-test 1, the preferred model for the CI group included SES, $F(1, 102) = 6.545$, $MSE = 4653.088$, $p = .012$, $R^2 = .061$, while the preferred model for the VM group included race, $F(1, 90) = 5.024$, $MSE = 4401.562$, $p = .027$, $R^2 = .053$. Finally, for the combined question types on Post-test 1, the preferred model for the CI group consisted of only subjective ability, $F(1, 102) = 10.141$, $MSE = 4042.156$, $p = .002$, $R^2 = .091$), while the model for the VM group included subjective ability and race, $F(2, 90) = 7.469$, $MSE = 3437.615$, $p = .001$, $R^2 = .145$. For the fourth-grade CI group, the preferred model for pictorial question types on Post-test 2 consisted of only SES, $F(1, 102) = 7.284$, $MSE = 3960.422$, $p = .008$, $R^2 = .067$. There were no significant predictors identified for the VM group. For the symbolic question types, the preferred model for the CI group included subjective ability and SES, $F(2, 102) = 9.722$, $MSE = 6878.904$, $p < .000$, $R^2 = .163$. There were no significant predictors identified for the VM group. Finally, for the combined question types, the preferred model for the CI group included subjective ability, gender, and SES, $F(3, 102) = 6.725$, $MSE = 2464.694$, $p < .000$, $R^2 = .169$. There were no significant predictors identified for the VM group. Fewer significant predicting variables in the VM group compared to the CI group for each test item type suggest that fewer demographic factors mediate students' performance when fraction concepts are taught using virtual manipulatives.

RQ#2: Use of Representations Analyses

The second part of Research Question #2 focused on the relationship between time students spent using each representation (pictorial, symbolic, and physical/virtual manipulatives), type of learning participation (individually or in groups), and students' test scores (Post-test 1, Post-test 2, learning, and retention). This examination helped us to understanding how using different representations predicted students' performance on tests and how working individually or in groups predicted students' performance on tests. This produced six instructional combinations for analysis: 1) pictorial used by an individual student (PI), 2) symbolic used by an individual student (SI), 3) manipulatives used by an individual student (MI), 4) pictorial used by students in a group (PG), 5) symbolic used by students in a group (SG), and 6) manipulatives used by students in a group (MG). These abbreviated notations in parentheses are used in the section that follows. This enabled comparison of student time spent using each instructional representation in groups or working individually, and their impact on test scores across treatment groups. We analyzed the ethogram observation data by treatment group (CI and VM) in relation to the mathematics test results. These results are summarized in Table 3.

Third-grade analyses. A linear regression analysis identified two significant predictors of Post-test 1 scores in the CI group, PI and SI, $F(2,8) = 14.58$, $p = .005$, $R^2 = .911$. Two significant predictors of Post-test 2 scores were identified in the CI group, PI and MG, $F(2,8) = 17.320$, $p = .003$, $R^2 = .923$. One significant predictor of retention scores in the CI group was identified: SI, $F(1,8) = 8.199$, $p = .024$, $R^2 = .734$. No significant predictors of learning scores were identified for the CI group. Furthermore, no significant predictors were found for Post-test 1, Post-test 2, learning, or retention scores for the VM group. This suggests that the representations used with individual students or with groups of students mediate students' performance in the CI groups, but not in the VM groups.

Table 3

Impact of representations on tests performance by treatment group in grade 3 and 4

	Treatment group	Pictorial individual (PI)	Symbolic Individual (SI)	Manips Individual (MI)	Pictorial group (PG)	Symbolic group (SG)	Manips groups (MG)
Post-test 1	CI	Δ	Δ		□		
	VM						
Post-test 2	CI	Δ	□			□	Δ
	VM						
Learning	CI						
	VM						
Retention	CI		Δ, □				
	VM						

Note. Δ indicates that the variable was a significant predictor of students' test scores for Grade 3; □ indicates that the variable was a significant predictor of students' test scores for Grade 4.

Fourth-grade analyses. The fourth-grade analysis identified one significant predictor of Post-test 1 scores in the CI group: PG, $F(1,1)= 6.414$, $p= .045$, $R^2= .719$. Two significant predictors of Post-test 2 scores in the CI group were identified: SI and SG, $F(2,7)= 10.95$, $p= .015$, $R^2= .902$. One significant predictor of retention scores in the CI group was identified: SI, $F(1,7)= 14.94$, $p= .008$, $R^2= .845$. No significant predictors of learning scores for the CI group were identified. Furthermore, no significant predictors of Post-test 1, Post-test 2, learning, or retention score for the VM group were identified. Similar to Grade 3, no significant predicting variables in the VM group (compared to the CI group) suggest that the representations used with individual students or with groups of students mediate students' performance in the CI groups, but not in the VM groups.

Discussion

The purpose of this study was to examine a number of variables that predict achievement results in two instructional treatments: Treatment 1: classroom instruction using physical manipulatives and text-based materials (CI), and Treatment 2: computer lab instruction using virtual manipulatives (VM).

Beyond Student Achievement

The main overarching finding from this study is that there are less extraneous variables (particularly demographic) that predict students' fraction achievement, learning, and retention in third- and fourth-grade classrooms when using virtual manipulatives versus text-based materials and physical manipulatives. For example, in Grade 3 CI classrooms, subjective ability, gender, SES, and ELL status predicted students' Post-test 1 scores, and subjective ability and ELL status predicted students' Post-test 2 scores. Similarly in Grade 4, subjective ability and SES predicted CI students' Post-test 1 scores, and subjective ability and SES predicted CI students' Post-test 2 scores. Gender predicted overall learning in Grade 3 CI classrooms, and objective ability and gender predicted retention in Grade 3 CI classrooms, while there were no significant predictors of learning or retention in Grade 4 for the CI group. In contrast, there were many fewer demographic and related predictors of student performance, learning and retention in the VM classrooms. In Grade 3, only subjective ability predicted VM students' Post-test 1, Post-test 2, learning, and retention scores. In Grade 4, race was the only predictor of students' Post-test 1, learning and retention scores.

Studies of virtual manipulatives compared with other instructional treatments often produce mixed results, or small effects in favor of the virtual manipulatives treatments. As this study showed, there is a story that goes beyond simply looking at the test results. This is a story about how other variables interact while students are learning mathematics. Our results shine a spotlight on potential hidden factors that influence student learning in mathematics classrooms. Variables that influence mathematics learning may also have the potential to equalized learning when students use virtual manipulatives.

Test Item Question Type Analyses

For both the CI and VM groups, subjective ability, which was identified by the classroom teacher based on prior student performance, was a frequent predictor of student performance for each of the test item question types. Other research has shown that students often perform to the level of a teacher's expectations, whether those expectations are high or low for their students. The most frequent predictor of student performance on each question item type in CI groups (after subjective ability) was students' SES, followed by ELL status and gender. SES predicted students' performance on Post-tests 1 and 2 in both Grades 3 and 4 in the CI groups. SES, ELL status and gender were rarely predictors in the VM group. The most frequent predictor of student performance in VM groups (after subjective ability) was race, but this predictor only appeared in Grade 4, and appeared less frequently than SES and ELL as predictors in the CI group.

Results from the test question type analyses suggest that there are fewer predicting variables that mediate overall fraction achievement, learning, and retention in third- and fourth-grade classrooms when virtual manipulatives are used to teach fraction concepts, rather than text-based materials and physical manipulatives. These findings generally hold across test item type, whether the question is pictorial, symbolic, or combined. Such findings suggest that, regardless of test question type, classroom learning opportunities using virtual manipulatives—rather than textbooks and physical manipulatives—minimize the impact of extraneous demographic variables on learning fractions in third and fourth grade.

It is important to note, however, that subjective ability, as expected by the classroom teacher based on prior student performance, still exerted a pervasive influence on student learning in both VM and CI classrooms. Previous studies have demonstrated that this variable strongly impacts learning in many contexts (e.g., Rosenthal & Jacobson, 1968). While using virtual manipulatives in the classroom may help eliminate the influence of some demographic variables—such as SES, ELL status, and gender—on student mathematics learning, further research is necessary to determine ways in which to reduce other social influences such as subjective teacher expectations. Recent research has begun to address this question in contexts other than virtual manipulatives, suggesting that teachers can change negative behaviors associated with subjective expectations for students, which in turn may positively affect classroom dynamics and student learning (e.g., Hamre et al., 2012).

Students' Use of Representations

In terms of representation use, the percentage of time students in third- and fourth-grade CI groups spent using pictorial and symbolic representations predicted students' scores on post-tests. Additionally, the time that third- and fourth-grade CI groups spent working individually with symbolic representations predicted student retention of concepts. In contrast, the percentage of time third- and fourth-grade VM students spent using different representations was not a significant predictor of performance.

These results suggest that time spent using certain types of instructional representations—whether pictorial, symbolic, or physical/virtual manipulatives—and the time devoted to individual vs. group work when using these representations mediates students' learning in the CI groups but not in the VM groups. Thus, similar to results discussed above concerning the influence of demographic variables on learning in the VM classroom, fewer variables relating to time spent using specific representations or individual/group contexts predict overall fraction achievement, learning, and retention in third- and fourth-grade classrooms when using virtual manipulatives versus text-based materials and physical manipulatives. Such results are important, as previous studies have shown relationships between early mathematics learning and types of representations used for such learning with other technology (e.g., Jordan & Baker, 2011; Jordan, Suanda, & Brannon, 2008; Moyer-Packenham et al., 2014). Taken together, results suggest that student learning with the virtual manipulatives—rather than through textbooks and physical manipulatives—helps minimize the impact of extraneous demographic and social variables on learning of mathematics.

Conclusion

In sum, this study determined which demographic variables predict student performance in third and fourth grade fraction learning by implementing a large-scale, random-assignment, delayed effects study comparing virtual manipulative use with other forms of instruction in the elementary classroom. The results revealed that, when this quasi-experimental design was used, SES and ELL status more strongly predicted achievement and learning in the CI as compared with the VM classrooms.

Results in the current study suggest that VM and CI classrooms provide differential opportunities to learn. So what do these results mean for children? These results mean that a child from a low-socio economic status or who is an English Language Learner in the CI classrooms already had a predetermined outcome on the tests; their ELL and SES status already predicted how well they would do on the tests in the CI classrooms. However, in contrast, because there were far fewer predictor variables in the VM classrooms, a child's demographics did not determine what their learning outcomes would be. Every child in the VM classrooms began with an equal chance. This is a crucial finding. STEM educators must seek to minimize the influence of demographic and other extraneous variables on student learning and achievement so that all students have access to STEM content and STEM-based careers. Such findings highlight the importance of further exploring the impact on teaching and learning of multiple potentially predictive variables when mathematics students use virtual manipulatives in a computer lab setting to learn concepts other than fractions, and in grades other than third and fourth.

Acknowledgement

Funding for this project was provided by a SPARC grant from the Vice-President for Research Office at Utah State University.

References

Anderson-Pence, K. L., Moyer-Packenham, P. S., Westenskow, A., Shumway, J. F., & Jordan, K. (in press). Visualizing mathematics: Relationships between visual static models and students' written solutions. *International Journal for Mathematics Teaching and Learning.*

Ball, S. (1988). Computers, concrete materials and teaching fractions. *School Science and Mathematics, 88*, 470-475. doi: 10.1111/j.1949-8594.1988.tb11839.x

Baturo, A. R., Cooper, T. J., & Thomas, K. (2003). Effective teaching with virtual materials: Years six and seven case studies. *Proceedings of the 27th Annual Conferences of International Group for the Psychology of Mathematics Education, 4*, 299-306.

Beck, S. A., & Huse, V. E. (2007). A "virtual spin" on the teaching of probability. *Teaching Children Mathematics, 13*(9), 482-486.

Bolyard, J. J., & Moyer, P. S. (2003). Investigations in algebra with virtual manipulatives. *ON-Math, Online Journal of School Mathematics, 2*(2), 1-10.

Cavanaugh, C., Billan, K. J., & Bosnick, J. (2008). Effectiveness of interactive online algebra learning tools. *Journal of Educational Computing Research, 38(*1), 67-95.

Clements, D. H., & Battista, M. T. (1989). Learning of geometric concepts in a logo environment. *Journal for Research in Mathematics Education, 20*(5), 450-467.

Clements, D. H., Battista, M. T., & Sarama, J. (2001). Logo and Geometry [Monograph]. *Journal for Research in Mathematics Education, 10,* 1-177.

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the *Building Blocks* project. *Journal for Research in Mathematics Education, 38*(2), 136-163.

Daghestani, L., Al-Nuaim, H., & Al-Mshat, A. (2004). Case study: Investigating the use of VLE in teaching mathematical concepts to third grade students. *Proceedings of the 2nd Saudi Science Conference*, 61-69.

Deliyianni, E., Michael, E., & Pitta-Pantazi, D. (2006). The effect of different teaching tools in overcoming the impact of the intuitive rules. In J. Novotná, H. Moraová, M. Krátká & N. Stehliková (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 409-416). Prague: PME.

Guevara, F. D. (2009). *Assistive technology as a cognitive developmental tool for students with learning disabilities using 2d and 3d computer* *objects. (*Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1465252)

Haistings, J. L. (2009). *Using virtual manipulatives with and without symbolic representation to teach first grade multi-digit addition* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3366234)

Hamre, B. K., Pianta, R. C., Burchinal, M., Downer, J. T., Howes, C., LaParo, K., & Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge, and observed practice. *American Educational Research Journal, 49*(1), 88-123. doi:10.3102/0002831211434596

Hauptman, H. (2010). Enhancement of spatial thinking with Virtual Spaces 1.0. *Computers and Education, 54*(1), 123-135. doi: 10.1016/j.compedu.2009.07.013

Highfield, K., & Mulligan, J. (2007). The role of dynamic interactive technological tools in preschoolers' mathematical patterning. In J. Watson & K. Beswick (Eds.), *Proceedings of the 30th annual conference of the Mathematics Education Research Group of Australasia* (pp. 372-381): MERGA Inc.

Izydorczak, A. E. (2003). *A study of virtual manipulatives for elementary mathematics.* Unpublished doctoral dissertation, State University of New York at Buffalo.

Jordan, K. E., & Baker, J. (2011). Multisensory information boosts numerical matching abilities in young children. *Developmental Science, 14*(2), 205-213. doi: 10.1111/j.1467-7687.2010.00966.x

Jordan, K. E., Suanda, S., & Brannon, E. M. (2008). Intersensory redundancy accelerates preverbal numerical competence. *Cognition, 108*(1), 210-221. doi: 10.1016/j.cognition.2007.12.001

Lin, C. (2010). Web-based instruction on pre-service teachers' knowledge of fraction operations. *School Science and Mathematics, 110*(2), 59-71. doi: 10.1111/j.1949-8594.2009.00010.x

MacNulty, D. R., Mech, D. L., & Smith, D. W. (2007). A proposed ethogram of large-carnivore predatory behavior, exemplified by the wolf. *Journal of Mammalogy, 88*(3), 595-605.

Manches, A., O'Malley, C., & Benford, S. (2010). The role of physical representations in solving number problems: A comparison of young children's use of physical and virtual materials. *Computers & Education, 54*(3), 622-640. doi: 10.1016/j.compedu.2009.09.023

Moyer, P. S., & Bolyard, J. J. (2002). Exploring representation in the middle grades: Investigations in geometry with virtual manipulatives. *The Australian Mathematics Teacher, 58*(1), 19-25.

Moyer, P. S., Bolyard, J., & Spikell, M. A. (2001, August). Virtual manipulatives in the K-12 classroom. In A. Rogerson (Ed.). *Proceedings of the International Conference New Ideas in Mathematics Education* (pp. 184-187), Queensland, Australia. Autograph.

Moyer, P. S., Bolyard, J. J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching Children Mathematics, 8*(6), 372-377.

Moyer, P. S., Niezgoda, D., & Stanley, J. (2005). Young children’s use of virtual manipulatives and other forms of mathematical representations. In W. J. Masalski & P.C. Elliott (Eds.), *Technology-supported mathematics learning environments: Sixty-seventh yearbook* (pp. 17-34). Reston, VA: NCTM.

Moyer-Packenham, P. S. (2005). Using virtual manipulatives to investigate patterns and generate rules in algebra. *Teaching Children Mathematics, 11*(8), 437-444.

Moyer-Packenham, P., Baker, J., Westenskow, A., Anderson, K., Shumway, J., Rodzon, K., & Jordan, K., The Virtual Manipulatives Research Group at Utah State University. (2013). A study comparing virtual manipulatives with other instructional treatments in third- and fourth-grade classrooms. *Journal of Education, 193*(2), 25-39.

Moyer-Packenham, P. S., Salkind, G. W., Bolyard, J., & Suh, J. M. (2013). Effective choices and practices: Knowledgeable and experienced teachers’ uses of manipulatives to teach mathematics. *Online Journal of Education Research, 2*(2), 18-33.

Moyer-Packenham, P. S., Shumway, J. F., Bullock, E., Tucker, S. I., Anderson-Pence, K., Westenskow, A., Boyer-Thurgood, J., Maahs-Fladung, C., Symanzik, J., Mahamane, S., MacDonald, B., & Jordan, K. (2014, April). *Young children’s learning performance and efficiency when using virtual manipulative mathematics iPad apps.* Paper presented at the annual National Council of Teachers of Mathematics Research Conference (NCTM), New Orleans, Louisiana.

Moyer-Packenham, P. S., & Suh, J. M. (2012). Learning mathematics with technology: The influence of virtual manipulatives on different achievement groups. *Journal of Computers in Mathematics and Science Teaching, 31*(1), 39-59.

Moyer-Packenham, P. S., & Westenskow, A. (2013). Effects of virtual manipulatives on student achievement and mathematics learning. *International Journal of Virtual and Personal Learning Environments, 4*(3), 35-50. doi: 10.4018/jvple.2013070103

Pea, R. D. (1985). Beyond amplification: Using the computer to reorganize mental functioning. *Educational Psychologist, 20*(4), 167-182.

Pleet, L. J. (1991). *The effects of computer graphics and mira on acquisition of transformation geometry concepts and development of mental rotation skills in grade eight* (Unpublished doctoral dissertation). Oregon State University, Corvallis, Oregon.

Reimer, K., & Moyer, P. S. (2005). Third graders learn about fractions using virtual manipulatives: A classroom study. *Journal of Computers in Mathematics and Science Teaching, 24*(1), 5-25.

Rosenthal, R., & Jacobson, L. F. (1968). Teacher expectations for the disadvantaged. *Scientific American, 218*, 19-23.

Smith, J. P. (1995). *The effects of a computer microworld on middle school students' use and understanding of integers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9544692)

Steen, K., Brooks, D., & Lyon, T. (2006). The impact of virtual manipulatives on first grade geometry instruction and learning. *Journal of Computers in Mathematics and Science Teaching, 25*(4), 373-391.

Suh, J., & Moyer, P. S. (2007). Developing students' representational fluency using virtual and physical algebra balances. *Journal of Computers in Mathematics and Science Teaching, 26*(2), 155-173.

Suh, J., Moyer, P. S., & Heo, H. J. (2005). Examining technology uses in the classroom: Developing fraction sense using virtual manipulative concept tutorials. *The Journal of Interactive Online Learning, 3*(4), 1-22.

Takahashi, A. (2002). *Affordances of computer-based and physical geoboards in problem solving activities in the middle grades*

(Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3070452)

Westenskow, A., Moyer-Packenham, P. S., Anderson-Pence, K. L., Shumway, J. F., & Jordan, K. (2014). Cute Drawings? The disconnect between students' pictorial representations and their mathematics responses to fraction questions. *International Journal for Research in Mathematics Education, 1*(1), 81-105.

Patricia S. Moyer-Packenham is professor of mathematics education at Utah State University, United States of America.

Joseph Baker is a post-doctoral researcher in psychology at Stanford University, United States of America.

Arla Westenskow is Director of the Tutoring Intervention and Mathematics Enrichment (TIME) Clinic at Utah State University, United States of America.

Katie L. Anderson-Pence is assistant professor of mathematics education at the University of Colorado – Colorado Springs, United States of America.

Jessica F. Shumway is a doctoral candidate in mathematics education at Utah State University, United States of America.

Kerry Jordan is associate professor of psychology at Utah State University, United States of America.

Contact Address: Direct correspondence concerning this article, should be addressed to the author. Postal address: Utah State University, 2605 Old Main Hill, Logan UT 84322 – 2605, United States of America. **Email**: patricia.moyer-packenham@usu.edu

Appendix A. Mean (SD) Third Grade Students Performance: Teacher x Treatment Group x Test (Moyer-Packenham et al., 2013)

Teacher	Treatment group	Pre-test %	Post-test 1 %	Post-test 2 %	Learning Avg.	Retention Avg.
Mrs. Alpha	VM	31.81	72.72	43.94	40.91	-28.79
		(12.53)	(19.07)	(14.56)	(24.89)	(23.29)
	PM	50.90	85.45	50.90	34.55	-34.55
		(20.93)	(8.13)	(26.19)	(19.71)	(24.39)
Mrs. Bravo	VM	69.70	85.86	72.73	16.16	-13.13
		(13.63)	(8.01)	(7.87)	(13.46)	(10.27)
	PM	71.71	79.80	61.62	8.10	-18.20
		(21.53)	(19.69)	(20.21)	(11.53)	(12.02)
Mr. Charlie	VM	60.61	75.00	54.55	14.40	-20.45
		(18.32)	(11.05)	(15.98)	(14.22)	(16.95)
	PM	48.95	74.13	57.34	25.17	-16.78
		(22.12)	(19.58)	(21.44)	(22.31)	(16.94)
Mrs. Delta*	VM	56.06	59.10	58.57	3.03	-.51
		(19.53)	(20.26)	(25.37)	(25.52)	(23.01)
	PM	58.90	67.99	56.52	9.10	-11.46
		(21.21)	(17.52)	(19.56)	(18.99)	(21.27)
Mrs. Echo	VM	67.27	72.72	69.10	5.45	-3.64
		(19.73)	(18.68)	(18.77)	(15.56)	(13.68)
	PM	67.27	75.45	71.81	14.54	-3.63
		(24.26)	(25.37)	(18.40)	(12.27)	(23.93)
Mrs. Foxtrot	VM	63.64	69.32	72.72	5.68	3.41
		(17.52)	(18.78)	(13.74)	(10.79)	(8.32)
	PM	66.94	78.51	72.72	11.57	-5.78
		(16.40)	(13.65)	(11.49)	(18.67)	(11.69)

Teacher	Treatment group	Pre-test %	Post-test 1 %	Post-test 2 %	Learning Avg.	Retention Avg.
Mrs. Golf	VM	74.54	70.00	76.40	-4.54	6.36
		(13.41)	(10.54)	(13.68)	(10.71)	(14.87)
	PM	58.33	46.97	54.54	-11.36	7.57
		(21.39)	(22.21)	(24.51)	(15.07)	(19.31)
Total	VM	60.51	72.10	63.99	11.58	-8.10
		(16.38)	(15.20)	(15.71)	(16.45)	(15.77)
	PM	59.52	72.61	60.78	13.09	-11.83
		(21.12)	(18.02)	(20.26)	(16.93)	(18.51)

Note. The asterisks indicate teachers who taught more than one class.

Appendix B

Appendix B. Fourth-Grade Students Performance: Teacher x Treatment Group x Test (Moyer-Packenham et al., 2013)

Teacher	Treatment group	Pre-test %	Post-test 1 %	Post-test 2 %	Learning Avg.	Retention Avg.
Mr. Hotel*	VM	50.98 (18.63)	61.06 (22.51)	42.85 (19.25)	10.08 (16.12)	-18.21 (17.34)
	PM	52.45 (17.93)	60.78 (19.13)	40.68 (16.63)	8.33 (17.25)	-20.09 (16.89)
Mrs. India*	VM	49.85 (20.00)	67.50 (20.93)	54.34 (14.98)	17.64 (20.46)	-13.16 (20.45)
	PM	49.41 (21.62)	69.11 (19.91)	46.47 (16.84)	19.40 (22.92)	-22.64 (16.88)
Mrs. Juliet	VM	47.05 (20.88)	67.37 (18.51)	45.45 (15.07)	20.32 (20.60)	-21.92 (13.94)
	PM	47.89 (17.39)	60.51 (23.28)	38.65 (16.65)	12.60 (15.27)	-21.84 (11.86)
Mrs. Kilo	VM	47.05 (20.65)	57.98 (19.92)	45.37 (22.19)	10.92 (12.44)	-12.60 (20.77)
	PM	43.53 (24.04)	58.82 (25.41)	40.00 (17.71)	15.29 (15.23)	-18.82 (21.79)
Mrs. Lima	VM	35.94 (16.49)	67.32 (30.86)	57.51 (22.74)	31.37 (24.07)	-9.80 (12.82)
	PM	39.57 (14.43)	69.51 (19.65)	55.08 (14.94)	29.94 (14.03)	-14.43 (11.87)
Mrs. Mike	VM	61.17 (18.84)	80.58 (25.27)	59.41 (24.08)	19.41 (17.54)	-21.17 (9.86)
	PM	50.98 (19.32)	75.49 (20.96)	64.21 (15.35)	24.50 (20.19)	-11.27 (13.35)

Teacher	Treatment group	Pre-test %	Post-test 1 %	Post-test 2 %	Learning Avg.	Retention Avg.
Mrs. November	VM	42.64 (16.47)	59.80 (20.80)	45.09 (20.42)	17.15 (23.32)	-14.70 (14.94)
	PM	41.17 (14.18)	67.64 (19.18)	51.96 (13.23)	53.47 (21.35)	-15.68 (17.25)
Total	VM	47.81 (18.58)	65.94 (23.37)	50.01 (20.04)	18.13 (19.22)	-15.94 (16.23)
	PM	*46.43* *(18.42)*	*65.98* *(20.39)*	*48.15* *(15.68)*	*19.55* *(18.03)*	*-17.83* *(15.20)*

Note. The asterisks indicate teachers who taught more than one class.