Ross, J.A. & Bruce, C.D. (2009). Student achievement effects of technology-supported remediation of understanding of fractions. *International Journal of Mathematical Education in Science and Technology*, 40(6), 713-727.

Student Achievement Effects of Technology-Supported Remediation of Understanding of Fractions

Abstract

Students have difficulty learning fractions and problems in understanding fractions persist into adulthood, with moderate to severe consequences for everyday and occupational decision-making. Remediation of student misconceptions is hampered by deficiencies in teachers' knowledge of the discipline and pedagogical content knowledge. We theorized that a technology resource could provide the sequencing and scaffolding that teachers might have difficulty providing. Five sets of learning objects, called CLIPS, were developed to provide remediation on fraction concepts.

In this article we describe one stage in a research program to develop, implement and evaluate CLIPS. Two studies were conducted. In the first, 14 grade 7-10 classrooms were randomly assigned, within schools, to early and late treatment conditions. A pre-post, delayed treatment design found that that CLIPS had no effect on achievement for the Early Treatment group due to unforeseen implementation problems. These hardware and software issues were mitigated in the late treatment in which CLIPS contributed to student achievement (Cohen's $d=.30$).

Study 2 was a pre-post, single group replication involving 18 grade 7 classrooms. The independent variable was the number of CLIPS completed. Completion of all five CLIPS contributed to higher student achievement: Cohen's $d=.53$, compared to students who completed none ($d=.00$) or 1-4 CLIPS ($d=.02$).

The two studies indicate that a research-based set of learning objects is effective when the full program is implemented. Incomplete sequences deprive students of instruction in one or more constructs linked to other key ideas in the conceptual map and reduce the amount of practice required to remediate student misconceptions.

1. Introduction

Fractions have long been described by educational researchers as a challenging area of the curriculum for students in mathematics (Behr, Lesh, Post & Silver, 1983; Carraher & Schliemann, 1991; Hiebert 1988, to name just a few). The failure to master simple fractions tasks can have serious consequences, for example:

A newly graduated registered nurse...administered one-half grain of morphine when, in fact, one-eighth grain was ordered, reasoning that since 4 plus 4 equals 8, $\frac{1}{4}$ plus $\frac{1}{4}$ equals $\frac{1}{8}$ (instead of $\frac{1}{2}$). Although the patient survived, the dose was enough to depress her respiration to a life threatening level. This was not an isolated incident...(Grillo, Latif, & Stolte, 2001, p. 168).

In this article we will argue that many students emerge from elementary school (i.e., kindergarten to grade 8 or ages 4-14) with insufficient knowledge of fractions to complete tasks routinely encountered by adults. We will describe a set of learning objects designed to provide remediation in fractions and report the results of field tests involving three samples of students in grades 7-10 (ages 13-16). The research reported is part of a multi-year project to develop, implement and assess learning objects for core objectives in mathematics.

2. Literature Review

2.1 Evidence of low achievement on fractions

The challenge of learning fractions has been documented in case studies of middle school children (Armstrong & Larsen, 1995; Kamii & Clark, 1995). Fractions are difficult to learn because they require deep conceptual knowledge of part-whole (how much of an object or set is represented by the fraction symbol), measurement (fractions are made up of numbers that can be ordered on a number line), and ratios (Hecht, Close, & Santisi, 2003). A representative national US sample found that only one-third of 13-year old students were able to correctly place a simple fraction on a number line, a learning objective for 11-year olds (Kamii & Clark, 1995).

Manipulating fractions is particularly difficult when embedded in word problems. Boaler (1993) found that 12-13 year old students in a school with a traditional math curriculum had difficulty recognizing the math when simple tasks requiring the comparison of fractions were presented in a context-laden form. Students in a school committed to teaching for deep understanding were more likely to find context enabling, although the overall performance in both schools was poor.

2.2 Importance of fractions in post-school experience

School success in mathematics contributes to higher earnings. A national longitudinal study found that after controlling for the effects of number of years in school, adults with stronger mathematics skills (including fractions) had significantly higher wages than adults with lower mathematics achievement (Murnane, Willet, & Levy, 1995). The effects increased from the 1970s to the 1980s in response to increased employer demand for basic cognitive skills. The importance of fractions in blue collar occupations was demonstrated in an observational study of jobs in a car plant (Smith, 1999). Among the most important occupational demands for knowledge of fractions is for nurses and pharmacists dispensing medications, especially in pediatrics where fractional doses are the norm (Cartwright, 1996), as noted in the quotation that opened this article. In a contrary view, Hoyles, Noss, and Pozzi (2001) provided ethnographic

evidence that pediatric nurses invent algorithms specific to particular medications that preclude the need for mathematical understanding.

In addition to occupational requirements, knowledge of fractions is required in everyday decision making, such as self-monitoring medication doses. Less numerate individuals have poorer medical outcomes when their treatment requires following complicated medication instructions (Estrada, Martin- Hryniewicz, Peek, Collins, & Byrd, 2004).

Low performance on fractions tasks continues into adulthood. The National Adult Literacy Study found that almost twice as many adult Americans (22%) scored below basic on numeracy tasks that included fractions than on other literacy dimensions (Reyna & Brainerd, 2007), a finding also reported in other studies (Lipkus, Samsa, & Rimer, 2001). Skilled professionals are not exempt: low scores on simple fractions tasks have been reported for samples of nurses (Kapborg, 1995; Pozehl, 1996) and pharmacy students (Grillo et al., 2001).

The continuation of low performance on fractions tasks into adulthood can be attributed to curriculum priorities in schools. Students who fail to learn fractions in elementary school have little opportunity to acquire these skills in high school because the mathematics curriculum is so crowded. This is particularly the case in Ontario, the province in which our research was conducted, because five years of high school were compacted in to four beginning in 1999.

2.3 Why create a technology resource?

Teachers' ability to identify the conceptual origins of student difficulties, predict misconceptions, and relate current to future curriculum topics is enabled by their disciplinary and pedagogical content knowledge (Lloyd & Wilson, 1998; Spillane, 2000). Generalist teachers in elementary schools and high school teachers assigned mathematics classes outside their specialization, might not have had the opportunity to develop the conceptual foundations

required to promote deep understanding of fractions. Previous studies found evidence of generalist teacher confusion around core fractions concepts (Ball, 1990; Lehrer & Franke, 1992; Marnich, 2002). We theorized that a technology resource could provide additional support to teachers in the form of sequencing and scaffolding. Scaffolding in this context consists of learning materials that provide temporary support to enable a student to achieve a desired performance. With sequenced practice, the amount of scaffolding required for success is gradually reduced and the student is able to perform independently (see Vygotsky, 1978). In addition, having a resource that students could complete relatively independently of teacher guidance might increase teacher willingness to provide remediation to students who need it.

3. CLIPS

CLIPS (Critical Learning Instructional Paths Supports) are multi-media learning objects focused on fractions. The term CLIPS was coined by Hill and Crévola as “devices for bringing expert knowledge to bear on the detailed daily decisions that every teacher must make in teaching a coherent domain of the curriculum” (Fullan, Hill & Crévola, 2006, p. 56). The Ontario Ministry of Education, applied the term CLIPS to a series of software programs for students needing additional support for learning fractions. To view the fractions CLIPS go to:
<http://oame.on.ca/CLIPS/index.html>.

Students begin by viewing a video which shows why students should care about fractions. Students are presented with a menu of five sets of activities: (i) representing simple fractions; (ii) forming and naming equivalent fractions; (iii) comparing simple fractions; (iv) forming equivalent fractions by splitting or merging parts; and (v) representing improper fractions as mixed numbers. Within each set of activities there are introductory instructions, interactive tasks, consolidation quizzes and extension activities. The CLIPS are designed as a

research-based sequence of lessons of 15-20 minutes per day for five days but each can stand alone and teachers may assign less than five on the basis of their assessment of student needs.

For example, CLIPS A has an introduction activity on representing simple fractions. There is a voice over with area models presented on the screen. In the second activity students are asked to describe a fraction by entering the numerator and denominator and showing what the fraction looks like in an area model. There are three additional mini-sets of activities. The student is given a quiz on representing simple fractions. Students drag their answers to a box and receive immediate feedback. If incorrect, they are given an explanation. The final component of CLIPS A is a “show what you know” screen which suggests five different activities (e.g., a fractions card game) students could do as consolidation. The same structure is repeated for each of the five CLIPS.

4. Purpose of the study

The research program was an extended-term mixed-method design (Chatterji, 2004) involving four stages: (1) We conducted a systematic needs assessment (Ross, Ford, & Bruce, 2007) which combined student performance data, student beliefs about their mathematical competencies, and teacher perceptions about the difficulty and importance of learning particular fractions objectives. The needs assessment provided a rank ordered list of the most urgent student learning needs. (2) This list was transformed into an integrated learning agenda that prescribed an instructional sequence designed to overcome student deficits. A design team drew upon research on teaching fractions (particularly Gould, Outhred, & Mitchelmore, 2006; Moss & Case, 1999; Streefland, 1993) to develop interactive activities addressing each learning need. (3) The CLIPS were pilot tested in two classrooms to assess their functioning and to generate detailed recommendations for their revision (Bruce and Ross, in press). (4) The CLIPS were revised on

the basis of qualitative data collected in the pilot test and field tested (the research reported here).

This article reports the results of Stage 4. The purpose was to measure the effects of CLIPS on student achievement in conditions approximating normal use. The research question was: Did CLIPS contribute to students' understanding of fractions?

5. Study 1

5.1 Research Design

In Study 1 grade 7-10 classrooms were randomly assigned, within schools, to early and late treatment conditions, as shown in Table 1. Teachers selected students they believed would benefit from CLIPS (see below). These students were tested on three occasions. Occasions 1 and 2 compared the Early CLIPS group to a control group; i.e., students who had been selected for CLIPS but had not received the treatment. Occasions 2 and 3 compared the late CLIPS group to the early group that was not using CLIPS at the time. All teachers who volunteered for the project had equal opportunity to implement CLIPS, thereby avoiding demoralization of the control group and denial of treatment to students who could benefit from it. Since teachers, not researchers, selected which students received CLIPS, the design was quasi-experimental, requiring a demonstration that the groups were equivalent on entry.

Table 1 about here

5.2 Study 1 Participants

In Study 1, 14 grade 7-10 teachers from one public school district in Ontario, Canada volunteered to participate. The district served a student population in which 98% were Canadian born, only 1% spoke a language other than English at home, 24% were identified as having special needs, and average family income in the district was near the mean for the province of Ontario. Classrooms were randomly assigned, in schools, to the early and late treatment groups.

Teachers used the results of the first test occasion and their knowledge of students to identify 91 of their 364 students (25%) for CLIPS training.

5.3 Study 1 Data Sources

Student Achievement was measured on three occasions. Test 1 consisted of six fractions items drawn from the PRIME placement tests for Number and Operations (PRIME, 2005). Tests 2 and 3 each consisted of 16 items generated for the study. The items included procedural tasks, e.g., “Write two fractions that are equivalent to $\frac{5}{9}$ ” and conceptual e.g., “ $\frac{2}{10}$ is less than $\frac{2}{5}$. How do you know?” All items were scored 0-2 by trained markers. Inter-rater reliability, based on samples of 200, 210, and 210 items for the three test occasions, was high. Perfect agreement was 93%, 96%, and 94% and chance adjusted agreement was Kappa=.86, .92, and .88 respectively. Student achievement was operationalized at each test occasion as the mean item score.

At Test occasion 1 we administered a battery of affect measures to determine the equivalence of the early and late treatment groups. *Math self-efficacy* consisted of eight Likert items measuring expectations about future mathematics performance (from Ross, Hogaboam-Gray, & Rolheiser, 2002; e.g., “as you work through a math problem how sure are you that you can...explain the solution”). There were six response options, anchored by “not sure” and “really sure”.

Functional beliefs about mathematics learning consisted of five statements about participating in mathematical discussions from Jansen (2006) and Schoenfeld (1985). The items were Likert scales; e.g., “If you are there throwing out your ideas, you could find a new way of doing a math problem.” with six response options, anchored by “strongly agree” and “strongly disagree”.

Dysfunctional beliefs about mathematics learning consisted of eight items from Schommer-Aitkins, Duell and Hunter (2005), measuring belief in quick/fixed learning (i.e., that learning occurs quickly or not at all and that intelligence is fixed rather than incremental); e.g., “If I cannot understand something quickly, it usually means I will never understand it.” There were six response options, anchored by “strongly agree” and “strongly disagree”.

Fear of failure consisted of six items (e.g., “I worry a lot about making errors on my math work”) from Turner, Meyer, Midgley, and Patrick (2003). There were six response options, anchored by “not at all true” and “very true”.

Effort was measured with eight items from Ross et al., 2002 (e.g., “how hard do you study for your math tests?”). There were six response options, anchored by “not hard at all” and “as hard as I can”. Students also reported their gender and grade (7-10).

After administering CLIPS in their classrooms, teachers completed an implementation survey consisting of 12 open and fixed response items. The survey asked teachers how they assigned students to CLIPS, the amount of time spent on CLIPS, perceptions of the adequacy of time allotted, where students worked on CLIPS (e.g., the classroom, the library, at home), perceptions of student like/dislike of the program, perceptions of student success, difficulties in accessing CLIPS software and whether the teacher was able to resolve technical problems, strategies used by the teacher to interact with students while working with CLIPS and to debrief them after completion, whether teachers felt the right students had received CLIPS, whether they would use CLIPS again, and fractions topics taught during CLIPS implementation. Teachers also indicated how many (0-5) CLIPS each student completed.

5.4 Teacher Training

CLIPS were designed to be used with relatively little teacher training. Teachers who volunteered for the project received 60 minutes of instruction that focused on how to access the program (through disks or the web), hardware requirements, the structure of the five CLIPS, and the projected benefits for students.

5.5 Study 1 Analysis Procedures

After establishing the reliability of the measures used in the study, we used t-tests to determine the equivalence of the early and late treatment groups. To assess the effects of CLIPS on the early treatment group, we conducted analysis of covariance (using GLM in SPSS 16.0) in which the dependent variable was posttest achievement O_2 ; the covariate was pretest achievement O_1 ; and the independent variable was experimental condition (the early group that received CLIPS versus the late group that did not). To assess effects of CLIPS on the late treatment group, we conducted analysis of covariance in which the dependent variable was posttest achievement O_3 ; the covariate was pretest achievement O_2 ; and the independent variable was experimental condition (the late group that received CLIPS versus the early group that did not).

5.6 Study 1 Results

There were small amounts of achievement data missing completely at random [Little's MCAR test: $\chi^2=10.41$, $df=6$, $p=.108$]: one case for achievement at O_1 and four cases at each of O_2 and O_3 . We used Expectation Maximization to impute missing values. We used the same procedure to replace a small number of missing values for the affect variables. All variables were normally distributed: skewness and kurtosis were <1.0 on all measures. Table 2 shows that the variables were reliable: internal consistency was $\alpha >.70$ for all variables.

Table 2 about here

The t-tests comparisons displayed in Table 2 indicate there were no statistically significant differences on O₁ achievement or on any of the attitude or demographic measures. In both groups there were more boys than girls: 57% of the students in the Early Treatment were male and slightly more, 66%, in the Late Treatment were male. We interpreted these results to mean that the groups were equivalent prior to implementation.

We used data from students who did not complete any CLIPS ($N=273$) to determine whether the achievement tests were of equivalent difficulty. O₁ and O₂ were equivalent [$t(272)=-0.22, p=.824$]; the means (and standard deviations) were 1.52 (.44) and 1.52 (.39). O₂ and O₃ were different [$t(272)=3.62, p<.001$]; the mean item score (and standard deviation) for O₃ was significantly lower 1.45 (.34) than the mean for O₂.

We conducted a univariate analysis of covariance using GLM in SPSS. In the first ancova, the dependent variable was O₂ achievement; the covariate was O₁ achievement; the independent variable was study condition (early or late, i.e., exposure or no exposure to CLIPS). CLIPS had no statistically significant effect on student achievement for students in the Early Treatment group [$F(1,88)=.122, p=.728$]. Although the model explained 47% of the achievement variance, virtually all the variance was attributable to the pretest score [$F(1,83)=7.873, p<.001$]. Students who performed poorly on the achievement pretest continued to perform poorly on the posttest, regardless of whether they had completed CLIPS.

Responses to the teacher survey revealed obstacles to CLIPS implementation. Technology concerns focused on both hardware and software. Hardware issues included: sound (“lost audio”), inability to launch the program (“CD ROMs on computers didn’t work”) and sporadic problems running the program (“screen froze”). The diversity of hardware issues suggested that the system requirements of CLIPS pushed the limits of the hardware available in

the schools. Software issues included: the web version of the program (as opposed to disk delivery) was so slow students became discouraged, students were “unable to open the disk”, “program did not run”, and “one disk didn’t work”. The lack of equipment in their homerooms and school policies that assigned computers to non-classroom locations meant that the majority of students in the Early Treatment group accessed CLIPS outside their math classroom. In these non-classroom locations students were either unsupervised (e.g., home use) or monitored for behaviour only (i.e., if students appeared to be working on the program, regardless of how, the librarian or supervisor of the computer room did not intervene). For further discussion of these implementation issues, see Bruce and Ross (in press).

CLIPS had a positive impact on student achievement in the Late Treatment group. In the second analysis of covariance, the dependent variable was achievement O_3 ; the covariate was achievement O_2 ; the independent variable was experimental condition. In this phase of the study, the Late Treatment received CLIPS and the Early Treatment did not. Students who received CLIPS between O_2 and O_3 had higher achievement than students who did not receive CLIPS during this time period [$F(1,88)=4.71, p=.033$]. CLIPS accounted for 5% of the achievement variance, Cohen’s $d=.30$.

Figure 1 shows the achievement test means for the Early and Late Treatment groups at each test occasion. Also displayed in the figure are the means for students who were not assigned CLIPS. Their means were substantially higher than students identified by teachers as needing support in learning fractions (the non-CLIPS group obtained 73% of the maximum score across all three tests) and the non-CLIPS group declined from O_2 and O_3 while the means from the two CLIPS groups improved. The most likely explanation for the decline of the scores of the non-CLIPS group was measurement error.

Figure 1 about here

We re-ran the analysis to explore whether the number of CLIPS completed influenced achievement. We compared students who completed no CLIPS ($N=273$) to those who completed 1-4 ($N=65$) and all five CLIPS ($N=26$). Table 3 shows the results of the ancova in which the dependent variable was O_3 achievement; the covariate was O_2 achievement; and the independent variables were group (Early or Late), number of CLIPS completed (none, some or all), and the interaction between CLIPS completed and group. We found an interaction: the number of CLIPS completed made a difference for the Late Treatment group (who completed CLIPS during the O_2 to O_3 period) but not for the Early Treatment (who were not working on CLIPS during the O_2 to O_3 period). The means indicated that students who completed all five CLIPS improved more than students who completed some or none: Cohen's $d=.32$ for the all CLIPS versus none comparison and $d=.16$ for the all versus some comparison.

Table 3 about here

5.7 Study 1 Discussion

Study 1 found that CLIPS had a statistically significant effect on student achievement in the Late but not the Early Treatment group. The most likely explanation is that unforeseen technology-related challenges could not be overcome by the first group of teachers using CLIPS. By the time of the second round, implementation problems had been identified and teachers could draw on their colleagues for strategies to reduce technical problems and the need for greater support for students working on CLIPS. For example, CLIPS was loaded on the district server so that students avoided problems of the web-based version and had to work on the program in school.

We also found that the effects of CLIPS were stronger if students completed all five. Even though CLIPS were designed to be an individualized program, teachers tended to assign

each CLIPS to a particular day, e.g., A=Monday, B=Tuesday, etc. If students were absent on a particular day, teachers typically had them skip that CLIPS if it had been assigned to them and continue with the next CLIPS scheduled on the day of their return. We found many instances of completed CLIPS sequences that were not likely planned by teachers, e.g., A-B-E or B-D.

In Study 2 we investigated these issues further. We were particularly interested in determining whether the number of CLIPS completed influenced student performance.

6. Study 2

6.1 Study 2 Methodology

Study 2 was a pre-post single group design in which the dependent variable was posttest achievement, the covariate was pretest achievement, and the independent variable was number of CLIPS completed (0, some, or all five). Participants were 307 grade 7 students in 18 classrooms from the Catholic school district that was co-terminus with the public school district in Study 1. The two districts served overlapping regions with virtually identical student population characteristics: 98% were Canadian born, less than 1% spoke a language other than English at home, 26% were identified as special needs, and average family income in the district was near the mean for the province of Ontario.

The sources of data were the same achievement tests administered in Study 1 and the number of CLIPS completed by students as reported by their teachers. Teacher training included the information given to Study 1 teachers, except that Study 2 teachers were given specific guidance, drawn from Study 1, in how to optimize implementation. There were three sets of conditions (Bruce and Ross, in press): students (e.g., encourage students to think of their assignment to CLIPS as a reward or learning opportunity rather than a punishment), technical

(e.g., ensure that all students have headphones), and teaching (e.g., provide a warm up that matches the procedural and conceptual focus of CLIPS).

6.2 Study 2 Results

There were no missing data. Both achievement scales were reliable ($\alpha=.71$ and $.78$ respectively) and normally distributed (skewness $=-.918$ and -1.208 ; kurtosis $=-.116$ and 1.369). We grouped the number of CLIPS assigned into the three categories of use developed for Study 1: none, some (i.e., 1-4), and all. Teachers assigned at least one CLIPS to 42% of their students. The pretest and posttest means for each student group are shown in Table 4.

Table 4 about here

Table 4 shows there was virtually no change from pre to post for the groups that received no CLIPS or less than five CLIPS, in contrast with the means for the group that completed all five. We conducted an analysis of covariance in which the outcome variable was the student achievement posttest, the covariate was the student achievement pretest score, and the independent variable was the number of CLIPS assigned. Table 5 shows that after controlling for pretest effects, there was a statistically significant effect for CLIPS: Students who were assigned all the CLIPS outperformed students who were assigned none or less than the complete package. The within-subjects effect sizes were $d=0.0$ for students receiving no CLIPS, $d=.02$ for students receiving less than all five CLIPS and $d=.53$ for students completing five CLIPS.

Table 5 about here

6.3 Study 2 Discussion

Study 2 replicated the findings of Study 1: completion of all five CLIPS contributed to student achievement but completion of less than all five had no effect. The effect sizes were larger than in Study 1, increasing from $ES=.30$ in the late treatment of Study 1 to $ES=.53$ in Study 2. Our

interpretation is that the technology-related difficulties students encountered in Study 1 and the lack of supervision/support they received, made it doubtful that students completed CLIPS as intended. For example, some students may have entertained themselves by manipulating the flashier aspects of the program without engaging in serious learning. The teachers in the Study 2 sample had the benefit of this experience and were able to reduce the problems.

7. General Discussion

The two studies demonstrate that learning objects completed by students with relatively little teacher direction can improve student performance on fractions tasks that are difficult to learn. The effect sizes were modest, when compared to the results of efficacy studies of treatments implemented in optimal conditions. Our studies, although they used volunteers (required by our ethical review protocol) approximated effectiveness studies because we did not screen the teachers, teacher training on the use of the CLIPS was minimal, and no additional hardware was provided to the schools. For example, the CLIPS effect sizes for Study 1 late treatment ($ES=.30$) and Study 2, $ES=.53$, compare favourably to the median $ES=.50$ in a study of 300 meta-analyses of educational and psychological interventions (Lipsey & Wilson, 2001) and to the $ES=.15$ for all Comprehensive School Reform programs in the meta-analysis of Borman, Hewes, Overman and Brown (2003) and $ES=.11$ for all Title I programs for schools serving disadvantaged students (Borman & D'Agostino, 1996).

The key finding of the study was that students needed to complete all five CLIPS. We suggest three reasons: (1) Student absences had negative effects, in part because teachers unnecessarily aligned CLIPS with days of the week. Students who missed a CLIPS to which they were assigned did not make it up. (2) The storyboards for the CLIPS followed an instructional sequence based on research on fractions learning (Gould, Outhred, & Mitchelmore, 2006; Moss

& Case, 1999; Streefland, 1993). Incomplete sequences deprived students of instruction in one or more constructs linked to other key ideas in the conceptual map. (3) The entire sequence involved modest amounts of instructional time (5 * 20 minutes). Incomplete sequences provided insufficient practice to remediate student misconceptions.

The two studies highlight the importance of evaluating the effects of technology interventions through a program of research, rather than one-shot summative evaluations. Researchers are too quick to generalize from a single evaluation (Burkhardt & Schoenfeld, 2003). Even though we pilot tested an earlier version of CLIPS, neither we nor our practitioner partners anticipated the severity of the hardware and software challenges teachers would face when the program was rolled out to a larger group of classes. If we had focused exclusively on the results from the Early Treatment group in Study 1, we would have concluded that the program had no student achievement effects. It was only when the implementation issues were reduced that student benefits began to emerge.

8. Conclusion and Limitations

We were not able to implement a true experiment as we had intended in Study 1. Genuine experiments are the gold standard in educational research and provide greater protection from internal validity threats than quasi-experiments (Wortman, 1994), although establishing causality is challenging even within experimental designs (McMillan, 2007). In Study 1 unforeseen difficulties in implementing study conditions resulted in serious attrition, which required that we treat the design as a pre-post quasi-experiment with matched controls. Quasi-experiments provide useful information provided there is evidence of the equivalence of groups on entry to the study and an appropriate predictor of achievement (the most powerful is prior achievement) is included as a covariate (Bloom, Richburg-Hayes, & Black, 2007). We provided evidence of

pretest group equivalence and included prior achievement as a covariate. Slavin and Lake (2008) in a best-evidence synthesis of elementary math programs reported that only 36 of the 87 they included as high quality studies used random assignment.

The task of developing effective learning sequences both online and offline is worth continuing effort. Fractions are difficult for students to learn, there is relatively little attention to them in the high school years, and flaws in performing simple fractions tasks can have moderate to severe consequences in adult life. The findings of our research to date suggest that CLIPS are promising: students who completed the sequence of five CLIPS improved their understanding, using modest amount of instructional time and only modest teacher intervention.

References

- Armstrong, B. E., & Larson, C. N. (1995). Students' use of part-whole and direct comparison strategies for comparing partitioned rectangles. *Journal for Research in Mathematics Education*, 26(1), 2-19.
- Ball, D. L. (1990). Prospective elementary and secondary teachers' understanding of division. *Journal for Research in Mathematics Education*, 21(2), 132-144.
- Behr, M. J., Lesh, R., Post, Th. R., & Silver, E. A. (1983). Rational-Number concepts. In Lesh Richard, & M. Landau (Eds.), *Acquisition of mathematical concepts and processes*. (pp. 91-126). New York: Academic Press.
- Bloom, H. S., Richburg-Hayes, L., & Black, A. R. (2007). Using covariates to improve precision for studies that randomize schools to evaluate educational intentions. *Educational Evaluation and Policy Analysis*, 29(1), 30-59.
- Boaler, J. (1993). Encouraging the transfer of 'school' mathematics to the 'real world' through the integration of process and content, context and culture. *Educational Studies in Mathematics*, 25, 314-373.
- Borman, G., & D'Agostino, J. (1996). Title I and student achievement: A meta-analysis of federal evaluation results. *Educational Evaluation and Policy Analysis*, 18, 309-326.
- Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73(2), 125-230.
- Bruce, C. D., & Ross, J. A. (in press). Conditions for effective use of interactive on-line learning objects: The case of a fractions computer-based learning sequence. *Electronic Journal of Mathematics and Technology*

- Burkhardt, H., & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher*, 32(9), 3-14.
- Carraher, D. W., & Schliemann, A. D. (1991). Children's understanding of fractions as expressions of relative magnitude. In F. Furinghetti (Ed.), *Proceedings of the fifteenth Psychology of Mathematics Education conference* (Vol. 1pp. 184-191). Assisi Italy: Psychology of Mathematics Education.
- Cartwright, M. (1996). Numeracy needs of the beginning registered nurse. *Nurse Education Today*, 16, 137-143.
- Chatterji, M. (2004). Evidence on "What Works": An argument for extended-term mixed-method (ETMM) evaluation designs. *Educational Researcher*, 33(9), 3-13.
- Estrada, C. A., Martin-Hryniewicz, M., Peek, B. T., Collins, C., & Byrd, J. C. (2004). Literacy and numeracy skills and anticoagulation control. *American Journal of the Medical Sciences*, 328, 88-93.
- Fullan, M. G., Hill, P., & Crvola, C. (2006). *Breakthrough*. Thousand Oaks, CA: Corwin.
- Gould, P., Outhred, L., & Mitchelmore, M. (2006). *One-third is three quarters of one-half*. Paper presented at the annual meeting of the Mathematics Education Research Group of Australia, Australia.
- Grillo, J. A., Latif, D. A., & Stolte, S. K. (2001). The relationship between preadmission indicators and basic math skills at a new school of pharmacy. *The Annals of Pharmacotherapy*, 35(2), 167-172.
- Hecht, S. A., Close, L., & Santisi, M. (2003). Sources of individual differences in fraction skills. *Journal of Experimental Child Psychology*, 86, 277-302.

- Hiebert, J. (1988). A theory of developing competence with written mathematical symbols. *Educational Studies in Mathematics*, 19, 333-355.
- Hoyles, C., Noss, R., & Pozzi, S. (2001). Proportional reasoning in nursing practice. *Journal for Research in Mathematics Education*, 32(1), 4-28.
- Jansen, A. (2006). Seven graders' motivations for participating in two discussion-oriented mathematics classrooms. *Elementary School Journal*, 106(5), 409-4028.
- Kamii, C., & Clark, F. B. (1995). Equivalent fractions: Their difficulty and educational implications. *Journal of Mathematical Behavior*, 14(4), 365-378.
- Kapborg, I. D. (1995). An evaluation of Swedish nurse students' calculating ability in relation to their earlier educational background. *Nurse Education Today*, 15, 69-74.
- Lehrer, R., & Franke, M. L. (1992). Applying personal construct psychology to the study of teachers' knowledge of fractions. *Journal for Research in Mathematics Education*, 23(3), 223-241.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21(1), 37-44.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Lloyd, G. M., & Wilson, M. (1998). Supporting innovation: The impact of a teacher's conceptions of functions on his implementation of a reform curriculum. *Journal for Research in Mathematics Education*, 29(3), 248-274.
- Marnich, D. B. (2002). *How the knowledge base of pre-service teachers affects the diagnosis and remediation of student errors when adding fractions with unlike denominators*. Unpublished doctoral dissertation, Preston University, Cheyenne, Wyoming .
- McMillan, J. H. (2007). *Randomized field trials and internal validity: Not so fast my friend*.

- Practical Assessment Research & Evaluation* 12(15). Retrieved February 28, 2008, from <http://pareonline.net/getvn.asp?v=12&n=15>
- Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education*, 30(2), 122-148.
- Murnane, R. J., Willet, J. B., & Levy, F. (1995). The growing importance of cognitive skills in wage determination. *Review of Economics and Statistics*, 77(2), 251-266.
- Pozehl, B. J. (1996). Mathematical calculation ability and mathematical anxiety of baccalaureate nursing students. *Journal of Nursing Education*, 35(1), 37-39.
- PRIME (Professional Resources and Instruction for Mathematics Educators). (2005). *Number and Operations Diagnostic Tests*. Toronto: Thomson Nelson.
- Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences*, 17(2), 147-159.
- Ross, J. A., Hogaboam-Gray, A., & Rolheiser, C. (2002). Student self-evaluation in grade 5-6 mathematics: Effects on problem solving achievement. *Educational Assessment*, 8(1), 43-58.
- Ross, J. A., Ford, J., & Bruce, C. D. (2007). Needs assessment for the development of learning objects. *Alberta Journal of Educational Research*, 53(4), 430-433.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando: Academic Press.
- Schommer-Aitkins, M., Duell, O. K., & Hutter, R. (2005). Epistemological beliefs, mathematical problem-solving beliefs, and academic performance of middle school students. *Elementary School Journal*, 105(3), 289-304.

- Slavin, R. E., & Lake, C. (2008). Effective programs for elementary school mathematics: A best evidence synthesis. *Review of Educational Research*, 78(3), 427-515.
- Smith, J. P. (1999). Tracking the mathematics of automobile production: Are schools failing to prepare students for work. *American Educational Research Journal*, 36(4), 835-878.
- Spillane, J. P. (2000). A fifth-grade teacher's reconstruction of mathematics and literacy teaching: Exploring interactions among identity, learning, and subject matter. *Elementary School Journal*, 100(4), 307-330.
- Streefland, L. (1993). Fractions: A realistic approach. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 289-325). Hillsdale, NJ: Earlbaum.
- Turner, J. C., Meyer, D. K., Midgley, C., & Patrick, H. (2003). Teachers' discourse and sixth graders' reported affect and achievement behaviors in two high-mastery/high-performance mathematics classrooms. *Elementary School Journal*, 103(4), 357-382.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wortman, P. (1994). Judging research quality. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 98-109). New York: Russell Sage Foundation.

Table 1

Research Design for Study 1

Early Treatment (N=44)	O ₁	X _{CLIPS}	O ₂		O ₃
Late Treatment (N=47)	O ₁		O ₂	X _{CLIPS}	O ₃

O₁=Test occasion 1 consisting of six Fraction items and a battery of affect measures

O₂=Test occasion 2 consisting of 16 Fraction items

O₃=Test occasion 3 consisting of 16 Fraction items

Table 2

Alphas, Means, Standard Deviations, and t-test Comparisons for Early (N=44) and Late (N=47)

Treatment Groups in Study 1

Variable	Alpha	# items	Group	Mean	SD	Comparison of Means
Math self efficacy			early	3.45	0.84	
	0.88	8	late	3.71	0.89	$t(89)=-1.41, p=.164$
Beliefs about math			early	3.58	0.81	
	0.81	5	late	3.60	0.88	$t(89)=-0.11, p=.916$
Quick/fixed learning			early	2.60	0.90	
	0.94	8	late	2.57	1.03	$t(89)=0.15, p=.884$
Fear of failure			early	3.09	1.18	
	0.81	6	late	3.00	1.20	$t(89)=0.35, p=.731$
Effort			early	3.94	1.14	
	0.89	8	late	3.87	0.96	$t(89)=-0.28, p=.784$
Grade*			early	2.00	1.08	
		1	late	1.79	0.62	$t(89)=1.41, p=.257$
O ₁ achievement			early	0.97	0.42	
	0.75	6	late	0.82	0.49	$t(89)=1.57, p=.120$
O ₂ achievement			early	0.96	0.38	
	0.84	16	late	0.88	0.50	$F(1,88)=.122, p=.728$
O ₃ achievement			early	1.03	0.35	
	0.82	16	late	1.10	0.42	$F(1,88)=4.71, p=.033$

*Grade: grade 7=1, grade 8=2, grade 9=3, grade 10=4

Table 3 Effects of Group Assignment and Number of CLIPS on O₃ Achievement in Study 1

Source	SS	df	MS	F	Sig.	Partial Eta ²
Corrected Model*	34.87	6	5.81	105.97	<.001	0.640
Intercept	9.23	1	9.23	168.21	<.001	0.320
None-some-all CLIPS * group	0.55	2	0.27	5.01	.007	0.027
None-some-all CLIPS	0.11	2	0.05	0.96	.384	0.005
Group	0.12	1	0.12	2.27	.133	0.006
O ₂ mean	23.11	1	23.11	421.28	<.001	0.541
Error	19.58	357	0.05			
Total	720.84	364				
Corrected Total	54.45	363				

* R Squared = .640 (Adjusted R Squared = .634)

Table 4 Pre- and Posttest Means of Students Assigned CLIPS in Study 2

	N	N		
	CLIPS	Students	Mean	SD
pretest	none	178	1.48	0.51
pretest	some	90	1.46	0.45
pretest	all	39	1.27	0.56
posttest	none	178	1.48	0.39
posttest	some	90	1.47	0.35
posttest	all	39	1.52	0.35

Table 5 Effect of Number of CLIPS and Pretest Scores on Posttest Student Achievement in Study 2

Source	SS	df	MSS	F	Sig.	Partial Eta ²
Corrected Model*	17.36	3	5.785	81.294	0.000	0.446
Intercept	22.41	1	22.409	314.900	0.000	0.510
pretest	17.28	1	17.284	242.888	0.000	0.445
# CLIPS	0.68	2	0.340	4.782	0.009	0.031
Error	21.56	303	0.071			
Total	714.82	307				
Corrected Total	38.92	306				

*R Squared = .446 (Adjusted R Squared = .440)