

A Systematic Replication and Extension of Using Incremental Rehearsal to Improve Multiplication Skills: An Investigation of Generalization

Robin S. Coddling · Jillian Archer · James Connell

Published online: 6 February 2010
© Springer Science+Business Media, LLC 2010

Abstract The purpose of this study was to replicate and extend a previous study by Burns (*Education and Treatment of Children* 28: 237–249, 2005) examining the effectiveness of incremental rehearsal on computation performance. A multiple-probe design across multiplication problem sets was employed for one participant to examine digits correct per minute and percentage of digits correct on problems targeted during treatment and response maintenance sessions. Treatment occurred twice weekly across 12 weeks. Retention of target facts was measured prior to each treatment session. Treatment impact on generalized skills including single skill mastery probes, multiplying fractions, and word problems was examined. Results suggested that the participant achieved accurate and fluent performance on the problem sets. This performance was maintained over time and generalized to SSM, fraction, and word problem probes.

Keywords Mathematics · Interventions · Generalization

Introduction

According to a report from the National Mathematics Advisory Panel (2008), U.S. students cannot solve single digit addition, subtraction, multiplication, or division problems as quickly or efficiently as students from other countries. The National Mathematics Advisory Panel suggested that differences in computational fluency

R. S. Coddling (✉) · J. Archer
Department of Counseling & School Psychology, Graduate College of Education,
University of Massachusetts, 100 Morrissey Boulevard, Boston, MA 02125, USA
e-mail: robin.coddling@umb.edu

J. Connell
School Psychology Program, Psychological Studies in Education, Temple University, Philadelphia,
PA, USA

were related to the quantity and quality of practice within the classroom. That is, teachers do not provide enough opportunities to practice basic facts (Daly et al. 2007) and few textbook curricula include sufficient activities to facilitate fluent computation performance (National Mathematics Advisory Panel 2008). Insufficient practice may limit students' ability to access and understand higher level math concepts (Gersten and Chard 1999) and the absence of computation fluency is common among students identified as having a specific mathematical learning disability (Gersten et al. 2005). Given the frequent absence of practice in school, a first-line intervention for students struggling in mathematics may be to provide interventions that provide multiple opportunities to practice basic facts.

Binder (1996) suggested as much as 70% of instructional time be devoted to practice activities. Practice has been described as brief opportunities consisting of modeling, feedback, and reinforcement (Daly et al. 2007) that use appropriately challenging materials (Burns et al. 2006). Two forms of practice have been defined in the literature (Haring and Eaton 1978; Johnson and Layng 1996): (a) drill, which is the practice of isolated items; and (b) composite practice, which requires the use of learned component responses in combination with previously learned responses. Some researchers have suggested that drill is the most critical form of practice (Cohen et al. 1992) while others (Haring and Eaton 1978) have suggested that drill and composite practice each serve separate roles in fluency building. Haring and Eaton postulated that the former may facilitate proficiency and the latter may lead to retention, maintenance, and generalization. To date, limited evidence is available to describe the relationship between prerequisite computation skills, retention, and generalization.

Unfortunately, a paucity of research exists on mathematics computation interventions (Coddling et al. 2009). Of the 12 interventions identified in a recent review of computation treatment research (Coddling et al. 2009), only half had been evaluated in three or more studies. Incremental rehearsal is a drill intervention that incorporates many of the aforementioned critical elements of practice as each item is practiced in isolation, students have frequent opportunities to practice items, and items are presented within an appropriate range of challenge by pairing known with unknown items in the practice sequence (Burns 2005). Although this intervention has produced promising findings (Coddling et al. 2009), only one study to date has examined the use of incremental rehearsal with mathematics (Burns 2005). Using a multiple-baseline across students design, Burns administered incremental rehearsal to three students with learning disabilities in mathematics using multiplication as the target skill. Incremental rehearsal led to increases in digits correct per minute (DCPM) for all students and the percentage of non-overlapping data was 100. Although promising, no maintenance data were collected for this study and generalization of these skills (other than the M-CBM probes) to other related skills was not measured.

The extent to which an intervention produces results that generalize across time, stimuli, and responses is the ultimate test of any treatment (Daly et al. 2006) and may only then correspond with continued use in real-world settings (Evidence-based Intervention Work Group 2005). Despite being a critical test for treatment use (American Psychological Association 2002), Coddling et al. (2009) found that only

38% of all computation intervention studies evaluated maintenance of performance following treatment and 11% examined generalized treatment effects. Although some researchers have suggested that generalization may not be spontaneous and often requires programming (Stokes and Baer 1977), other researchers postulate that generalization is associated with the level of skill proficiency achieved (Haring and Eaton 1978; VanDerHeyden and Burns 2008). According to Haring and Eaton (1978), learning progresses through a series of stages beginning with acquisition of the skill (i.e., accurate responding), leading to skill fluency (i.e., rapid and accurate responding), and progressing to generalization. However, research has yet to confirm the progression of these stages (Martens and Eckert 2007) or whether performing a skill with high rates of fluency results in retention, response, or stimulus generalization (Binder 1996). VanDerHeyden and Burns (2008) provided preliminary support for the link between retention of skills and generalization by indicating that students with performance rates of 17 (grades 2 and 3) and 29 (grades 4 and 5) DCPM on retention probes also achieved proficiency on the Stanford Achievement Test (9th edition; Harcourt Brace 1997). Mayfield and Chase (2002) identified a link between retention of computation skills and application to novel math tasks. Lin and Kubina (2005) also found strong correspondence ($r = 0.74$) between single- and multiple-digit multiplication fluency, providing evidence that increases with fact fluency early in the skill hierarchy can lead to improvements in more advanced skills. Taken together, these studies suggest that skill proficiency or mastery may evoke some forms of generalization; however, additional research is needed linking treatment options to generalized outcomes.

Purpose of the Present Study

The purpose of this study was to replicate and extend the literature on incremental rehearsal. In addition to examining the impact of the intervention on the accuracy and DCPM of three sets of unknown problems, we examined retention of facts from session to session, and stimulus generalization to subskill mastery (SSM) multiplication probes (facts 3–9), fractions, and word problems. Additionally, we examined pre-post performance on the grade level general outcome measure. We expected the participant to improve on the generalized measures once she achieved proficiency in each of the problem facts presented during incremental rehearsal.

Method

Participant and Setting

The participant, Sarah, was a 12-year old, seventh grade, Hispanic girl from a mid-sized northeast suburban middle school (grades 6–8). Sarah was referred in January of the academic year to the grade level student support team for difficulty in basic multiplication facts. Sarah participated in general education classes, was not diagnosed with a learning disability, and English was her primary language. As part of the standard curriculum in seventh grade, Sarah participated in two

mathematics courses (Interactive Mathematics, Number Sense & Geometry). The ethnic composition of the middle school was as follows: (a) 51% African American, (b) 22% Caucasian, (c) 16% Asian, (d) 9% Hispanic, and (e) 2% Multi-Race. Approximately 45% of students enrolled in the middle school were from low income backgrounds and 41% were English Language Learners. According to a state assessment of seventh graders' mathematics performance, 30% of students in the district performed at or above proficient levels (Massachusetts Department of Elementary and Secondary Education 2009).

The interventionist met individually with Sarah in an empty room within the library during school hours twice weekly for 20 min at a time designated as appropriate by Sarah's math teacher and did not interfere with regular academic instruction time.

Materials and Measures

Pre- and post-intervention general outcome measures (GOM)

In order to assess Sarah's grade level computation performance, one GOM mixed skill probe from AIMSweb (Shinn 2004) was administered before and after treatment. Reliability for computation GOMs is adequate ($r = .83-.93$; Shinn 2004). Criterion validity ranges from .36 to .62 (Thurber et al. 2002). Each probe was administered for 4 min consistent with standard administration procedures and scores represent total digits correct.

Subskill mastery measures (SSM)

During the study, Sarah's performance was assessed using SSM probes that were created according to the procedures described by Shapiro (2004) and administered using standard directions (Shinn 1989). Each probe was administered for 2 min. A random numbers simulation program was used, thereby randomizing the selection and order of the problems for each probe. Each SSM probe contained 56 problems (8 rows, 7 columns) and included any basic multiplication facts ranging from 3 to 9. Research has supported both the reliability and validity of SSM in mathematics. The test-retest ($r = .79$) and parallel forms ($.61 < r < .79$) of reliability for multiplication SSM are adequate (Foegen et al. 2007).

Problem sets

Thirty multiplication facts were divided into three sets (A, B, C) containing 10 problems each. Problem sets are commonly used to evaluate performance when they are consistent with the sequential teaching of computation facts (e.g., Skinner et al. 1989; McCullum et al. 2006). For each set of problems, three different probes were constructed containing 56 problems (8 rows, 7 columns) with each problem repeated 5 or 6 times across the problem set. Standard administration procedures (Shinn 1989) were used. These problem sets included 9 inverse facts.

Fraction probes

Multiplication fraction probes without reducing (e.g., $8/9 \times 3/6$) were created using the 30 unknown facts (using only digits 3 through 9) with Microsoft[®] Excel so that each probe contained 32 problems (8 columns, 4 rows). Six different probes were created. In order to provide consistency in administration, Sarah had 2 min to complete each probe. Each fraction answer resulted in 3 or 4 digits and no improper fractions were included.

Word problem probes

Using procedures from Stein et al. (2006), simple word problem probes were created for multiplication facts 3 to 9. Two problem types were included: (a) those with the word *each* or *every* (i.e., Keisha bought 6 bags of balloons. Each bag costs \$4. How much money did she spend?); and (b) those with the word *per* or a phrase using *a* (i.e., The grocery store makes and packages brownies in their store. The bakers put 6 brownies in a package. How many brownies would be in 8 packages?). Six different word problem worksheets were created containing 25 problems each. In order to provide consistency in administration, each probe was presented for 2 min.

Procedural integrity and intervention protocols

A typed intervention protocol consisting of scripted instructions for each step contained in the incremental rehearsal procedure was created according to procedures outlined by Burns (2005). An identical protocol was provided to an independent observer. At the end of each protocol, spaces were provided for the independent observer to mark the total number of steps completed during the session and to calculate the percentage of steps completed.

Dependent Measures

The primary dependent variables were the number of digits computed correctly per minute (DCPM) and the percentage of digits computed correctly across the problem sets. A secondary dependent variable was used to examine the retention of known facts between treatment sessions. The following dependent variables were used to examine generalization: (a) total digits correct on grade-level GOM probes (pre-post only), (b) DCPM and percentage of correct digits for SSM multiplication (facts 3–9) probes, (c) DCPM on fraction probes, and (d) DCPM on word problem probes. The following responses were scored as correct: (a) individual digits (even if the number was reversed or rotated), (b) “place holder” numbers, and (c) digits below the line. The following responses were scored as incorrect: (a) incorrect digits, (b) digits that were correct but appeared in the wrong place value, and (c) omitted digits (Shapiro 2004).

Procedures

General procedures

The second author was trained in CBM and intervention procedures by the first author and conducted the experiment. Sessions occurred twice weekly for 12 weeks and lasted approximately 20 min. Only 22 of 24 sessions were completed due to one school absence and a shortened school day.

Pre- and post- assessments

First, seventh grade GOM math probes were administered before and after treatment implementation to measure Sarah's performance on seventh grade material. Second, a fact assessment was conducted to determine Sarah's known and unknown facts. All multiplication facts from 0 to 12 were printed vertically on 3×5 index cards with black numbers. The cards were shuffled in order to permit random presentation of the facts. Known facts were those whose answer was identified correctly within 2 s (Burns 2005). No answers, incorrect answers, or correct answers that were provided after more than 2 s had lapsed were scored as incorrect. Once the unknown facts were identified, each was presented to Sarah a second time to confirm these facts as unknown. This assessment was conducted twice across two sessions to ensure accurate identification of known and unknown facts. All unknown facts were missed across both sessions. Thirty unknown facts were identified through these procedures and consisted of facts 3–9. Sarah did not know the '12' facts; however, given the number of unknown facts occurring earlier in the hierarchy, we did not include these.

Baseline procedures

Problem set probes for the target skill were employed according to Shinn's (1989) instructions and no intervention components were added. At least one baseline probe was administered across generalization measures.

Incremental rehearsal procedures

Standard procedures were followed according to Burns (2005). Prior to each session, all 30 unknown cards were presented to Sarah in order to provide a measure of retention from session to session. Any cards answered incorrectly, not at all, or correctly after 2 s had lapsed were counted as unknown. Known and unknown facts were recorded prior to each intervention session. A random selection of nine known facts was identified from each pile. Working in problem sets beginning with A, the first unknown fact with the answer was read aloud to Sarah by the interventionist after which Sarah was asked to restate the fact with the answer. Next, the interventionist presented Sarah with the first known fact. This sequence was repeated so that the first unknown fact was presented followed by the first and second known facts until all nine known facts were practiced after the first unknown

fact (see Burns 2005). Once this sequence was completed, the first unknown fact became the first known fact (replacing the last known fact) for the presentation of the second unknown fact. Unknown facts were presented and rehearsed individually as described above until the sequence was completed or three errors occurred (Burns 2005). The average number of items rehearsed each session was six. Following each session two or three of the following were presented: (a) problem set A probe, (b) problem set B probe, (c) problem set C probe, (d) fraction probe, (e) word problem probe, or (f) SSM multiplication probe facts 3–9.

Experimental Design

A multiple probe across problem sets design was employed (Barlow and Hersen 1984). Following baseline each problem set was targeted sequentially beginning with problem set A. Three follow-up sessions were conducted for problem sets A and B. Follow-up data could not be obtained for problem set C as the school year ended.

Interscorer Agreement

Interscorer agreement was assessed for DCPM by having a second experimenter independently score each probe. Comparisons between the two experimenters were conducted on a digit-by-digit basis. Reliability data were collected for 50% of sessions. Agreement was calculated by dividing the number of agreements per digit by the number of agreements per digit plus disagreements per digit and multiplying by 100 [(Agreements per minute/Agreements per minute +Disagreements per minute) \times 100]. Mean percent agreement for DCPM was 100%.

Procedural Integrity

Procedural integrity was assessed by an independent observer during 23% of the sessions. A checklist, which was identical to the protocols used by the interventionist, was created to assess procedural integrity for each intervention. The checklist contained the following information: (a) the intervention steps, (b) dependent variable administration steps, and (c) the materials required. Every intervention and dependent variable administration step was described in detail and included scripts that the interventionist was to follow. The independent observer was required to record a checkmark for presence of the required materials and when the intervention and CBM probe steps were correctly implemented by the experimenter. The number of steps checked by the independent observer was divided by the total number of steps listed for the procedure and then multiplied by 100. Procedural integrity across observed sessions was 99%.

Results

Sarah retained all 30 facts during the last three sessions of the 19 total intervention sessions implemented. Sarah's performance yielded an average gain of 2 problems

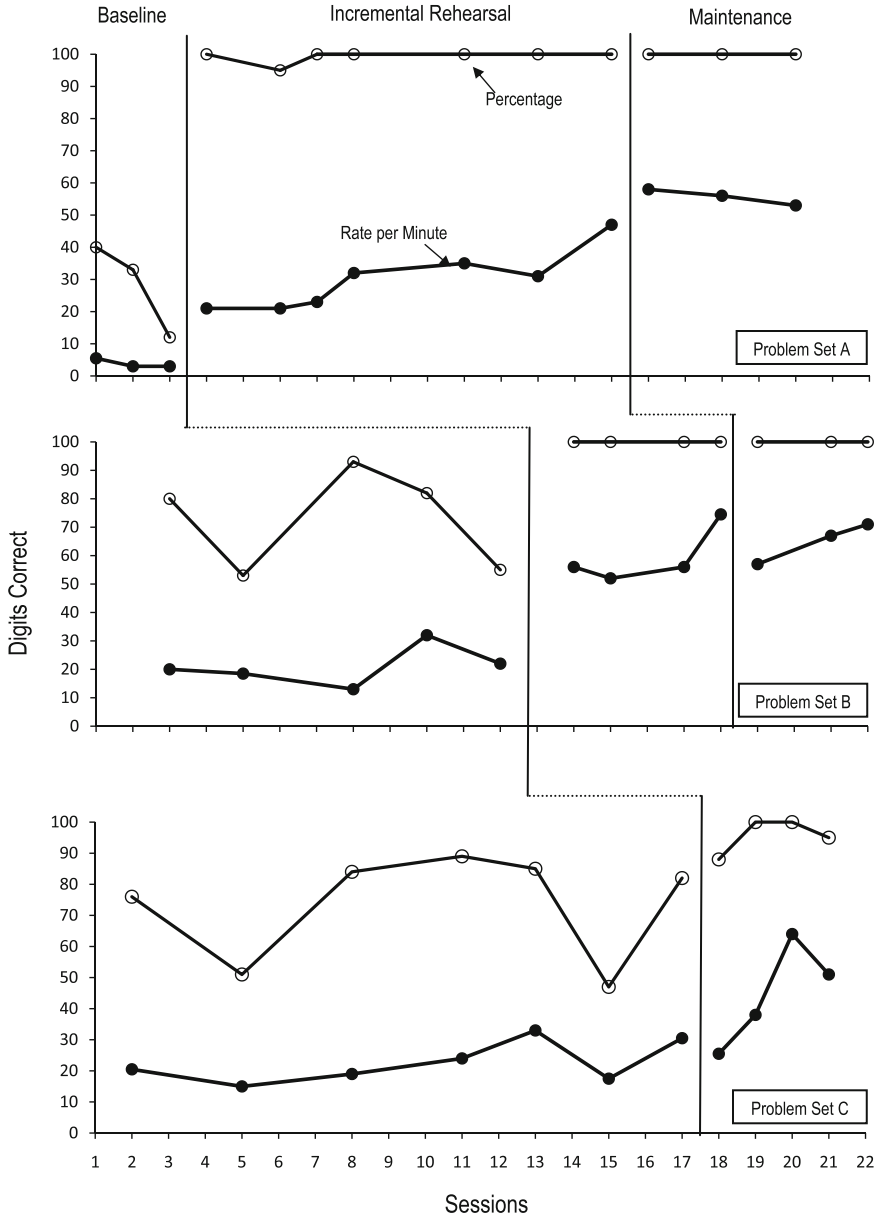


Fig. 1 Percentage of digits correct and digits correct per minute for the target skill across baseline, intervention, and follow-up phases

per session (range, 0–6). Figure 1 presents the percentage accuracy and DCPM results across problems sets. As can be seen from Sarah’s performance on the problem set probes, these probes did not present equal levels of difficulty. Sarah’s

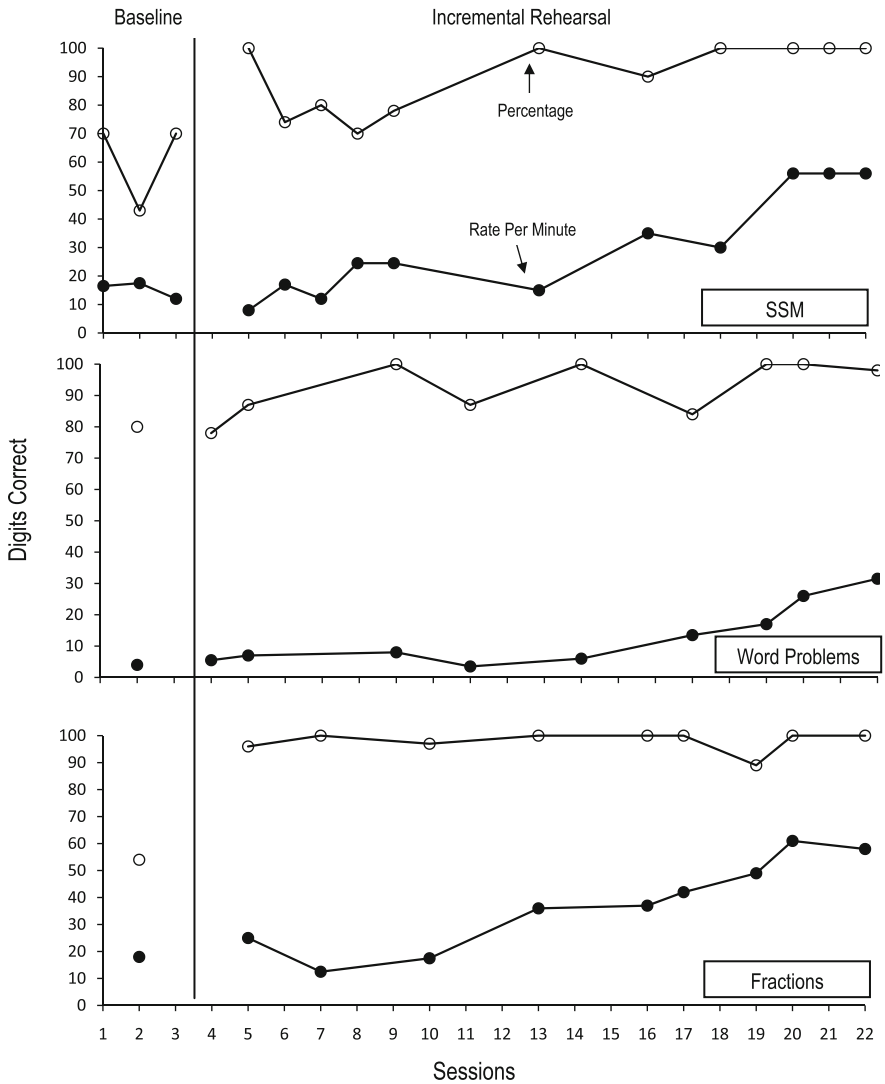


Fig. 2 Digits correct per minute and percentage of digits correct for the generalization SSM, word problem, and fraction probes across baseline and intervention phases

baseline performance was decreasing for problem set A ($M = 28.3\%$), but was variable for problem sets B ($M = 72.6\%$) and C ($M = 75.2\%$). Introduction of incremental rehearsal produced immediate level changes for problem sets A and B and a gradual increase for problem set C. Performance was maintained for problems sets A and B at 100% following treatment termination.

Baseline performance for DCPM fell in the frustrational range (Burns et al. 2006) for problem sets A ($M = 3.8$), B ($M = 21.1$), and C ($M = 22.8$). Baseline was stable for problem set A but yielded slightly increasing trends for problem sets B

and C. Consistent with the percentage accuracy data, introduction of incremental rehearsal produced level changes for problem sets A ($M = 30.0$) and B ($M = 48.7$) and a gradual increase in trend for problem set C ($M = 44.6$). Performance across all three problem sets met mastery level criteria (>49 DCPM; Burns et al. 2006). This performance level was maintained across problem sets A ($M = 55.6$) and B ($M = 65.0$).

Figure 2 illustrates generalization of practice on problem sets to SSM computation probes across all single-digit multiplication facts 3–9 as well as word problem and fraction probes. Mean baseline performance for the SSM computation probes (top panel) was 61.0% and for DCPM was 15.5. Although an immediate increase was observed in percentage accuracy with the introduction of incremental rehearsal, rates were variable following the first probe. During the last four intervention sessions, percentage accuracy was 100%. For DCPM, a gradually increasing trend was observed with the final sessions of the intervention corresponding to performance falling in the mastery range. Stable performance in the final three sessions reflects Sarah's completion of every problem on the probe within the time limit.

With the introduction of incremental rehearsal, percentage accuracy for word problems (top panel) was consistent with that of baseline but eventually increased to 100% by the end of the treatment sessions. A gradual increase was observed for DCPM, with the highest performance corresponding to the end of treatment. For the fraction probes an immediate increase was observed for percentage accuracy following introduction of treatment and these rates were maintained. An immediate level change was observed for DCPM followed by a decrease in performance and then an increasing trend.

Pre-post performance on the seventh grade GOM probe increased from 30 digits correct during the last session of baseline to 51 digits correct on the same probe administered during the 22nd session.

Discussion

The purpose of this study was to extend the research on incremental rehearsal for mathematics computation as well as examine response maintenance and stimulus generalization. This study extended previous research by examining pre-session retained cumulative gains and post-session problem set performance. Consistent with Burns (2005) we included SSM multiplication fact probes as a measure of generalization but also examined target skill generalization on probes consisting of fractions and word problems. Findings from this study are consistent with those from Burns (2005) and support the use of incremental rehearsal to increase fluent computation performance. This study also illustrates that incremental rehearsal increases the percentage accuracy of computation performance. Cumulative retention data taken prior to each session illustrated that a gradually increasing trend with several plateaus was evident and generalization across similar stimulus conditions was achieved following fluent performance across each set of problems taught.

Data examining generalization suggests SSM probes initially yielded an inverse relationship between percentage accuracy and DCPM in that when percentage accuracy was high DCPM was lower and vice versa, perhaps demonstrating the transition between acquisition and fluent performance. This finding might be consistent with Sarah learning to discriminate between which number combinations produce which specific responses and illustrate that practicing these responses among other known responses in the form of the SSM worksheet was initially challenging (Haring and Eaton 1978). This finding seems to be consistent with Haring and Eaton's distinction between drill and composite practice. Drill may be an important step towards building fluency but composite practice is likely needed to yield response maintenance and generalization. The DCPM data for fractions appeared to follow a similar trend; however, DCPM for word problems increased more slowly. This is likely a function of the stimulus presented to Sarah; that is, the fraction problems were more similar to the problem sets probes than the word problem probes were. It is notable that DCPM performance across generalization measures was highest during the last 4 sessions, consistent with the cumulative retention data which demonstrated that Sarah had responded quickly and accurately when presented with nearly all targeted 30 facts. Although Sarah demonstrated mastery performance on problem set probes according to the criteria developed by Burns et al. (2006), fluent performance on the SSM probe did not achieve proficiency according to this criterion until Sarah also performed fluently on problem set C.

There are several possible reasons for our results. These data appear to be consistent with those from VanDerHeyden and Burns (2008) which suggests that proficient performance leads to generalization on other related tasks. In the case of this study the stimuli were very similar, placing the same targeted math facts in slightly different contexts. That is, the SSM probes contained the identical unknown problems practiced to fluency along with previously known facts. Likewise the fraction probes contained the same 30 practiced facts but presented horizontally in the form of fractions. Therefore the format under which Sarah was expected to perform these facts were changed, however the facts remained the same. The more similar the stimuli the more likely generalization will occur (Alberto and Troutman 2006). Less proficient generalization was observed with the word problem probes, which requires that students recognize the math fact in the context of language. It may be that this type of transfer is more difficult to achieve given the increasing distance between the practiced and presented stimuli (Fuchs et al. 2009). These data may suggest that stimulus generalization is possible following increases in fluency on the target skill, the explanation for which can be found within the heuristic provided by the instructional hierarchy (Haring and Eaton 1978). It is also possible that, despite the probe nature of the experimental design, simply administering these generalization probes offered opportunities to practice on multiple types of exemplars (Cooper et al. 2007), thereby serving as specific strategy that promoted generalization.

Although this study contributes to our understanding of incremental rehearsal and generalization on computation performance, there are several limitations that require attention. First, no replications of these data were included. Therefore, future

researchers should replicate and extend this study by including more participants and different target skills. Second, although a functional relationship between treatment and the target behaviors was established such a relationship was not established with the generalization data. Third, given the probe nature of these data it is difficult to determine why percentage accuracy increased substantially on the fraction probes between baseline and treatment. It is possible that the probes, although randomly assigned to worksheets using Microsoft[®] Excel, sampled more items at the start of the probe from the first set of 10 facts exposed to treatment. Fourth, this study only examined treatment impact on stimulus generalization. It is possible that response generalization is more difficult to achieve and requires targeted programming. It would be useful to examine various forms of generalization that can be achieved following fluent performance with a target skill in future research.

In summary, these data provide additional evidence that incremental rehearsal improves accurate and fluent performance on target skills that can be maintained following treatment and will generalize to similar stimulus conditions, illustrated herein through probes consisting of word problems and fractions. Furthermore, these data may support the notion that generalization can be evoked following fluent performance as posited in the instructional hierarchy and demonstrated in previous work (Lin and Kubina 2005; Mayfield and Chase 2002; VanDerHeyden and Burns 2008). Taken together these findings do not necessarily suggest that there are two options for promoting generalization of academic skills, spontaneous or programmed, but perhaps also recognize that achieving skill mastery to an empirically identified criterion can lead to generalization, at least on closely related stimulus tasks. However, more research is needed to develop the relationship between skill proficiency and generalization in the area of mathematics.

References

- Alberto, P. A., & Troutman, A. C. (2006). *Applied behavior analysis for teachers* (7th ed.). Upper Saddle River: Pearson Merrill Prentice Hall.
- American Psychological Association. (2002). Criteria for evaluating treatment guidelines. *American Psychologist*, *57*, 1052–1059.
- Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change* (2nd ed.). Boston: Allyn and Bacon.
- Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *Behavior Analyst*, *19*, 163–197.
- Burns, M. K. (2005). Using incremental rehearsal to practice multiplication facts with children identified as learning disabled in mathematics computation. *Education and Treatment of Children*, *28*, 237–249.
- Burns, M. K., VanDerHeyden, A. M., & Jiban, C. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review*, *35*, 401–418.
- Codding, R. S., Hilt-Panahon, A., Panahon, C., & Benson, J. (2009). Addressing mathematics computation problems: A review of simple and moderate intensity interventions. *Education and Treatment of Children*, *32*, 279–312.
- Cohen, J. D., Servan-Schreiber, D., & McClelland, J. C. (1992). A parallel distributed processing approach to automaticity. *American Journal of Psychology*, *2*, 239–269.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River: Merrill.

- Daly, E. J., Bonfiglio, C. M., Mattson, T., Persampieri, M., & Foreman-Yates, K. (2006). Refining the experimental analysis of academic skills: Part II use of brief experimental analysis to evaluate reading fluency treatments. *Journal of Applied Behavior Analysis, 39*, 323–331.
- Daly, E. J., Martens, B. K., Barnett, D., Witt, J. C., & Olson, S. C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement instruction and intensity. *School Psychology Review, 36*, 562–581.
- Evidence-based Intervention Work Group. (2005). Theories of change and adoption of innovations: The evolving evidence—based intervention and practice in school psychology. *Psychology in the schools, 42*, 475–494.
- Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *The Journal of Special Education, 41*, 121–139.
- Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., et al. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology, 101*, 561–576.
- Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *Journal of Special Education, 33*, 18–28.
- Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*, 93–304.
- Haring, N. G., & Eaton, M. D. (1978). Systematic instructional procedures: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, & C. L. Hansen (Eds.), *The fourth r: Research in the classroom* (pp. 23–40). Columbus: Charles E. Merrill Publishing Co.
- Johnson, K. R., & Layng, T. V. J. (1996). On terms and procedures: Fluency. *The Behavior Analyst, 19*, 281–288.
- Lin, F., & Kubina, R. M., Jr. (2005). A preliminary investigation of the relationship between fluency and application for multiplication. *Journal of Behavioral Education, 14*, 73–87.
- Martens, B. K., & Eckert, T. L. (2007). The instructional hierarchy as a model of stimulus control over student and teacher behavior: We're close but are we close enough? *Journal of Behavioral Education, 16*, 83–91.
- Massachusetts Department of Elementary and Secondary Education. (2009). MCAS annual comparisons. Retrieved October 14, 2009 from <http://profiles.doe.mass.edu/mcas/mcascharts2.aspx?linkid=33&orgcode=02440000&fycode=2009&orgtypecode=5&>
- Mayfield, K. H., & Chase, P. N. (2002). The effects of cumulative practice on mathematics problem solving. *Journal of Applied Behavior Analysis, 35*, 105–123.
- McCullum, E., Skinner, C. H., Turner, H., & Saecker, L. (2006). The taped-problems intervention: Increasing multiplication fact fluency using a low-tech, classwide, time-delay intervention. *School Psychology Review, 35*, 419–434.
- National Mathematics Advisory Panel. (2008). Foundations for Success: The Final Report of the National Mathematics Advisory Panel. Washington, DC: U.S. Department of Education.
- Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention* (3rd ed.). New York: Guilford Press.
- Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. New York: The Guilford Press.
- Shinn, M. R. (2004). *Administration and scoring of mathematics computation curriculum-based measurement (M-CBM) and math fact probes for use with AIMSweb*. Minneapolis: NCS Pearson Inc.
- Skinner, C. H., Turco, T. L., Betaty, K. L., & Rasavage, C. (1989). Cover, copy, and compare: A method for increasing multiplication performance. *School Psychology Review, 18*, 412–420.
- Stein, M., Kinder, D., Silbert, J. & Carnine, D. W. (2006). Designing effective mathematics instruction: A direct instruction approach. Merrill.
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349–367.
- Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31*, 498–514.
- VanDerHeyden, A. M., & Burns, M. K. (2008). Using curriculum-based assessment and curriculum-based measurement to guide elementary mathematics instruction: Effect on individual and group accountability scores. *Assessment for Effective Intervention, 30*, 15–31.