

Generalization of Sight Word Accuracy using a Common Stimulus Procedure: A Preliminary Investigation

Eric M. Mesmer · Gary J. Duhon · Karen Hogan ·
Brandi Newry · Savannah Hommema ·
Courtney Fletcher · Maggie Boso

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

Abstract The need to investigate strategies promoting the generalization of reading skills is well documented. The current study builds upon previous literature by examining the impact of a common stimulus procedure on accuracy of reading generalization words in isolation. The common stimulus procedure used in this study intended to enhance the salience of the orthographic and phonemic features between taught and generalization words by highlighting the common rime structure in each word with a specific color. Taught and generalization words representing four word families were used (e.g., end, et, ell, & en). Using a multiple-baseline design across three participants with a replication of procedures for a fourth student, the impact of the common stimulus on accuracy of reading generalization words was examined. Results demonstrated increased accuracy and stability of responding when common stimulus procedures were implemented.

Keywords Generalization · Reading · Common stimulus

Students who struggle to learn how to read often have deficits in basic reading skills (Snow et al. 1998). Specifically, the development of basic decoding skills is one proficiency essential for successful reading (Snow et al. 1998). Goswami (1986) has suggested that children may use a process of orthographic analogy as one strategy

E. M. Mesmer (✉)
Department of Psychology, Radford University, Radford, VA 24161, USA
e-mail: emesmer@radford.edu

S. Hommema · C. Fletcher · M. Boso
Radford University, Radford, VA, USA

G. J. Duhon · K. Hogan · B. Newry
Oklahoma State University, Stillwater, OK, USA

for decoding unknown words. According to Goswami (1986), beginning readers will recognize the similarities in words that are closely constructed, for example “hat” and “mat”, and subsequently will use their recognition of the known word to read the unknown word. Thus, if the beginning reader knew the word “hat” he would be able to also read the word “mat” by recognizing similarities in the rime/at/. Research has confirmed the effectiveness of this decoding strategy in typically developing first- and second-grade students (Bowey et al. 1998). According to the National Reading Panel (2000), this process of acquiring phonics skills is reflected in analogy-based phonics. This approach emphasizes that children focus on parts of words that they know in order to read similar yet unknown words.

Studies examining strategies based on this approach with students with reading difficulties have demonstrated limited generalization effects. Berends and Reitsma (2006) examined the effects of an analogy-based reading strategy on reading fluency and generalization to generalization words. Seventy-four second-grade Dutch students experiencing reading difficulties were included in the study. The students were described as accurate readers but not fluent. Students were assigned to one of two intervention conditions. In Condition 1, students viewed a computer-based reading program in which 20 words were repeatedly presented on 20 occasions. The 20 words were orthographically similar with common consonant–vowel structures. In Condition 2, students were exposed to the same reading program but instead viewed 400 different words including the 20 words from group 1. Following training, students in both groups were expected to read two lists of words. The first list included words that were orthographically similar (i.e., either common onset or rime sequences) to the set of 20 words in group 1. The second list included words that were not orthographically similar to the training words. Results of the study demonstrated that students exposed to the 20 similar words in a repeated fashion made significant gains in accurately and fluently reading those 20 words while those exposed to 400 different words did not improve on the target 20 words. Unfortunately, when measuring generalized reading improvement on words closely resembling the target words (i.e., orthographically similar), even the students in Condition 1 failed to greatly improve in their accurate and fluent reading of these words.

In a similar study, Thaler et al. (2004) studied the impact of two reading interventions on 20 German students’ accuracy and fluency rates in reading generalized word sets. Both intervention groups were exposed to an intervention that highlighted the orthographic similarity between groups of words sharing a common onset pattern. One intervention group was required to read each phoneme in a presented word as the particular letter corresponding to the phoneme was color-highlighted in a computer-based program. The second group was merely required to follow along as common words were presented and highlighted on the computer screen. Following each session, students were then presented with 32 closely related words (same onset as the training words) and 32 non-related words in order to examine accurate and fluent reading with these generalized word sets. Results indicated no improvements in reading accuracy for either group for closely related and non-related words. Regarding fluency, both groups demonstrated a dramatic increase on the taught words, increasing performance by approximately 40%.

Unfortunately, performance on the closely related words and non-related words showed similar improvements of only 18%.

In the aforementioned studies by Berends and Retesma (2006) and Thaler et al. (2004), efforts were made to control for the orthographic and phonologic patterns represented within target and generalization words. Despite this commonality, results from both studies failed to result in impressive generalized improvements to generalization words. Other approaches to improve accuracy in reading non-taught words have also been attempted. For example, DeRose et al. (1996) conducted a study to teach reading of words with seven children from Brazil who were considered non-readers. Students were taught to match a set of printed words to dictated words via discrimination trials. Through discrimination of comparisons requiring the matching of pictures with printed and dictated words, the researchers proposed that the formation of stimulus equivalence classes would assist in promoting accurate identification of similar words (i.e., recombined words using the same syllables as those previously taught). However, despite the training, the average percentage of correctly read generalization words was 40.6% (no baseline for comparison).

Greater levels of success in producing generalized reading outcomes have been documented. For example, Joseph and Nist (2008) were able to dramatically increase accurate generalized reading of unknown words taught in isolation to reading those words in connected text using an incremental rehearsal procedure. By utilizing a pre-post-design, these researchers demonstrated that post-intervention accuracy increased above 90% for all 6 participants. Similarly, Daly et al. (2004) were successful in increasing the number of generalization words read accurately using a procedure intended to teach phoneme segmenting and blending. Specifically, these researchers found that participants who were taught to identify phonemes within nonsense words were able to accurately read these same phonemes within real words. Focusing teaching efforts at the phonemic level resulted in substantially higher accuracy rates compared to an intervention focused upon sight word acquisition.

Others have attempted to study the generalized effects of interventions on oral reading fluency within connected text (i.e., reading passages). Results of these studies have been equivocal, with improvements in generalized reading performance from pre- to post-intervention ranging in fluency gains of less than 10% to as high as 100%. In a meta-analysis examining the impact of repeated readings on generalized fluency gains, Therrien (2004) found a mean effect size of .50. However, substantially greater results were found when the repeated reading intervention was facilitated by an adult. These inconsistent results make it difficult to convincingly identify those practices that are likely to result in the greatest gains in generalization and may suggest that generalization is best understood at the idiographic level. Despite this possibility, the literature includes several common features that are useful in informing our current understanding of generalization of fluency outcomes. First, the majority of studies have utilized a repeated reading intervention to improve fluency. It is likely that the beneficial effects of repeated readings interventions on fluency have made it particularly attractive for investigating whether this intervention results in generalization (Ardoin et al. 2008).

However, studies suggest that repeated reading of a passage, in and of itself, is likely to produce only limited to moderate gains in fluency on novel passages (Therrien 2004; Valleley and Shriver 2003; Yurick et al. 2006). Second, fluent reading in generalization passages is most successful when a high amount of word overlap (i.e., 80% or higher) is present across intervention and generalization passages. High word overlap may provide a level of stimulus control that increases the likelihood of fluent responding in generalization passages (Daly et al. 2007). Third, instructional practices including adult facilitation, matching of skill instruction to materials read, error correction, and positive reinforcement have been found to enhance generalization effects (Ardoin et al. 2008; Daly et al. 2005, 2006; Therrien 2004).

Despite gains in our knowledge regarding generalization of reading fluency, the limited number of studies examining generalization of reading accuracy has left a void in our understanding of the methods that will promote accurate reading across generalized contexts (Ardoin et al. 2008). In many of the previous studies intending to promote generalization, efforts were made to ensure similarity between training words/passages and generalization words/passages. While such a strategy should exploit stimulus control, the results of these studies suggest that large improvements in generalized responding are often inconsistent. One possible explanation of these results is that the stimuli intended to elicit desired generalized responding are not sufficiently salient. It has been demonstrated that the mere presence of common stimuli may not be enough to produce generalized occurrences of behavior (Stokes and Baer 1976). Instead, as was the case in the Daly et al. study (Daly et al. 2004), these stimuli must not only be noticeable but greater outcomes may be more likely when these stimuli are functionally related to the desired response (Stokes and Baer 1977).

The purpose of this study was to examine the impact of a common stimulus procedure on accurate reading of unknown words that were orthographically similar to a set of taught words. Similar to the Berends and Retisma study (2006), the current study attempted to incorporate common stimulus conditions that saliently highlighted critical (i.e., those functionally related to an accurate response) word structures during intervention. However, in addition to using these common stimuli during intervention, this study was unique in that these stimuli were also present during assessment trials with generalization words.

Methods

Participants and Setting

Participants included 4 students between the ages of 7 years—8 months and 8 years—6 months, and all four participants were in the second grade. Three of the participants attended a rural Southwestern school in a small community, and one participant attended an urban school in a mid-size city in the Southeast. Two of the participants (Jenny and Madison) were Caucasian females, while one participant (Linda) was an African American female. The remaining participant, John, was a

Caucasian male. At the time of the study, Jenny, Madison, and John were all receiving special education services due to identified learning disabilities in reading. In addition, these 3 students all attended the same general education classroom and received the same reading instruction. Linda was not receiving special education services but was identified by her teacher as struggling in reading. The reading instruction provided to Linda in the general education classroom was similar to instruction provided to the other three participants with emphasis placed on whole word instruction and use of leveled text materials.

Procedures took place within each participant's school in an area separate from each student's typical instructional setting. For three of the students, activities occurred at a small table in a computer lab and for the fourth student, procedures were carried out at a student's desk in the hallway directly outside of the classroom.

Materials

Materials for this study consisted of two duplicate sets of 35 words each presented individually on 3 inch \times 4 inch laminated flashcards. The 35 words represented four word families including 9 words ending with *en*, 10 words ending with *et*, 10 words ending with *ell*, and 6 words ending in *end*. All words within each of these four word families were included in the study, thus resulting in 35 total words. The four word families were selected to ensure some degree of consistency across each family. First, each family chosen included the *e/vowel* sound. Second, an effort was made to select word families that had roughly the same number of words within each family.

One set of the 35 words simply included the words typeset in black and bolded 14 inch font. The second set of 35 words was identical to the first set, with one exception, the common word ending for each of the four word families was highlighted with a unique color. Specifically, the color green was used to highlight the rime portion of *en* ending words, while blue was used to highlight *et* words, yellow was used with *ell* words, and red was used with *end* words.

A small kitchen timer or stopwatch was used during each session to record the fluency with which words were read (however, fluency data were not used for purposes of this study). In addition, an 8.5 \times 11 inch score sheet was used to record student performance. The score sheet included check boxes indicating the phase of the study, checkboxes indicating which of the 35 words were read correctly, and a place to record the date and the participants' initials.

Dependent Variables

The dependent variables were the percentage of *taught* and *generalized* words read correctly. Taught words included 11 words representing the 4 word families (i.e., 3/*en*/words, 3/*et*/words, 3/*ell*/words, and 2/*end*/words) and were used during training procedures. Within each word family, an effort was made to ensure that one word was correctly modeled by the researcher, followed by the presentation of

2 practice words that participants received feedback to ensure they understood the task. Due to the fewer number of words present in the/end/family, however, only 1 practice word was provided in addition to the modeled word. These procedures resulted in the 11 taught words. Generalized words included the remaining 24 non-taught words representing the 4 word families.

Percentage correct was determined based upon the number of words read accurately divided by the total number of words presented during each data collection session. A word was considered to be correct if the word was orally pronounced in a manner consistent with Standard English. A word was counted as correct if pronunciation reflected a regional accent.

Experimental Design

A multiple-baseline design across 3 participants with a replication of procedures across a fourth participant was used to evaluate the impact of the intervention and the generalization strategy on percentages of taught and generalization words read correctly. The data collection phases of the study included baseline, training (intervention), and generalization. Importantly, data collection and intervention integrity procedures were the same for all 4 participants despite the fact that data were collected in more than one location. For the three participants from the same school, data were collected concurrently by two school psychology graduate students. Data were collected from the last participant, Linda, approximately 12 months later by two school psychology graduate students and one psychology undergraduate student. All five students involved in data collected were trained in study procedures by the primary researcher and were required to demonstrate accurate use of intervention and assessment procedures prior to data collection.

Procedures

Pre-Experimental Procedures

Prior to inclusion in the study, each participant was screened to determine if they would be suitable candidates. Each participant was administered a reading screener to determine whether they could accurately identify the letters of the alphabet as well as consonant and vowel sounds. Participants were required to have 90% accuracy or higher on these measures. In addition, each participant was asked to attempt to read each of the 35 words used in the study. This first exposure to the study words served as the first baseline data point for each of the participants. If a potential participant read more than 70% of the words accurately, they were not included in the study. All four participants selected for this study met the selection criteria.

Baseline

During baseline, experimenters presented each of the 35 color-highlighted flashcards to participants. The 35 flashcards were randomly presented to ensure that words within one family were not consecutively shown as one block. Prior to presenting the flashcards, students were told that they would be shown a series of words and were asked to read each word as it was presented. Because fluency data were also collected (although not presented as part of this study), students were asked to read as accurately and quickly as possible. Students were told that they would not receive any assistance and to skip a word if they did not know it. The first flashcard was then presented and the timer was started. Each of the 35 words was shown to the participants and correct and incorrect words were sorted into two piles. At the end of the session, the student was asked to return to his/her classroom and the flashcards were recorded as being read either correctly or incorrectly on the daily score sheet.

Training

During the training phase, an intervention was delivered with the goal of increasing participants' accuracy in reading the presented words. The intervention utilized an analogy-based phonics approach in which the rime within each word set was emphasized. On each day of training, the same 11 taught colored flashcards were placed in one pile. These flashcards were then presented to participants using the following procedures. First, the experimenter presented a model word from the first word family followed by either two (/en/,/et/&/ell/) or one (/end/) practice words from the corresponding word family. For example, the model word for the *en* family was *ten* and as the word was presented the examiner stated, "This word is ten, ten". Next a second word from the *en* family was presented (i.e., the first practice word), and the similarity between the model word and the practice word was pointed out. Specifically the researcher stated, "This word is pen, pen. Ten and pen sound alike, they both end with the/en/sound". The color cue common to both of these words (in this case green) was then pointed out to the participant as a way to make the association between the words more salient. This procedure was then repeated with a second practice word for the/en/,/et/, and/ell/word families. Finally, the participant was then told that they would be asked to read each of the taught words just presented to them as part of the training procedure. If the word was read correctly no feedback was provided. If a word were read incorrectly the experimenter stated, "The word is _____. See this word ends with the/en/sound and has the color green". The incorrectly read word was presented to the participant again and the participant was asked to read the word correctly. Following presentation of the last *en* taught word, this process was repeated with each of the three remaining word families.

Following the aforementioned training procedure, an assessment was conducted in which the participant was then shown each of the 11 taught words used in the intervention in order to determine how many were correctly read. Importantly, during this assessment, although the 11 words were the same words used during training, the flashcards on which these words were contained *did not include* the

color coding. This was done to see if participants would accurately read the taught words in the absence of color highlighting. The 11 words were randomly ordered and represented all four word families.

The assessment procedure detailed previously continued throughout the training phase each day following implementation of the intervention. In addition, beginning on day two and each day thereafter during the training phase, each participant was also presented with the set of 24 flashcards, which contained generalized words that were not used during intervention procedures. Again, during this phase, the 24 generalization flashcards did not contain the color coding, highlighting the common rime of a word family. In doing so, this allowed for a daily comparison of how the student performed on the 11 taught words and the 24 generalized words.

Generalization

The same procedures used during training were replicated during the generalization phase with two critical exceptions. When data were collected at the end of each session using the 11 training words, *the color cue* associated with the word family to which that word belonged *was included*. Second, when the 24 generalization words were presented the day following the previous training session, the 24 words were presented on the set of flashcards that contained *the color cue* for each respective word family. Thus, to summarize, during generalization the intervention was continued and participants were assessed using the colored flashcards, but during the training phase the non-colored flashcards were used for both taught and generalized words.

The intent of using the non-colored flashcards during training to assess accuracy with taught and generalized words was to (1) demonstrate that gains would be demonstrated for taught words in the absence of the color cues and (2) that such gains would not be seen for generalized words. The primary rationale for including the color cues during the generalization phase was to demonstrate that (1) minimal additional gain would be evidenced when assessing taught words and (2) accuracy gains would be evidenced on the generalized words once the color cue was presented. Finally, the use of color flashcards during baseline assessment was to demonstrate that use of color in and of itself had little impact on accurate reading of the generalized words and that improvement would not be evidenced until the color cues were subsequently associated with intervention procedures.

Procedural Integrity

Observations of intervention/data collection sessions occurred to ensure that study procedures occurred as intended. Twenty-five percent of sessions for each participant were randomly selected and observed by the primary researcher. A procedural checklist was used to record observation data. During baseline, the checklist contained seven steps. In summary, these steps included whether word cards were presented in randomized order, whether verbal directions were provided as indicated, whether the timer was used, whether correct and incorrect responses were sorted into separate piles, and whether correct and incorrect responses were

recorded on the score sheet. The same assessment procedures were observed and recorded during training and generalization phases (the exception being the use of non-colored and colored word cards, respectively). In addition, during the training phase, seven additional steps were observed to ensure that instructional procedures were followed. These seven steps were aligned with the procedures detailed in the training phase. Results of these integrity checks revealed that procedures were implemented as intended, as 93 to 100% of steps were included for each observed session.

Results

One baseline data point was collected for Linda with accuracy rates of 45 and 25% on taught and generalization words, respectively. On the first day of intervention, no gains were evidenced with the taught words. However, dramatic gains in level of the data were demonstrated quickly thereafter as Linda was able to read the taught words with 100% accuracy by day 5 of the intervention. Similar to the other participants in the study, the success of the intervention had minimal, if any impact on accuracy with generalization words. In Linda's case, the generalization words remained at a level comparable to baseline conditions, with accuracy rates fluctuating between 25 and 0% correct. When generalization procedures were implemented, a change was evidenced in the level of accuracy rates with generalization words. That is, when the materials used to assess reading accuracy were altered and contained the highlighted word endings used in the phonics building training procedures a change was evidenced. During the first generalization session, accuracy rates increased to 33%, with more dramatic increases evidenced during sessions 2–4 of the generalization phase.

As seen in Fig. 1, Jenny's performance during baseline was very stable, accurately reading 36% of the taught words and 29% of the generalization words on both days. Substantial increases in level and slope were evidenced with the training words during the training phase. Specifically, accuracy rates gradually increased with scores ranging between 63 and 100%. Similarly, an increase in level of accuracy was demonstrated by Jenny on generalization words as scores increased to 45% on the first data collection session during the training phase. After 3 days of an increasing trend that included a plateau of 70% accuracy in reading the generalization words, Jenny's performance began to decrease with a final score of 33% accuracy. Thus, while generalization effects were initially seen when presented with generalization words, Jenny's performance became unstable and downward trending over time. Finally, during the generalization phase, Jenny's performance with training words remained stable as each reading performance exceeded 90%. More importantly, Jenny's performance with generalization words stabilized at an accuracy rate of 75%. This level of performance generally exceeded her accuracy rates with generalization words when compared to the training phase.

Data for John can also be found in Fig. 1. John's reading accuracy during baseline was relatively stable but somewhat higher than the other participants in the study. John's accuracy in reading training words during baseline varied between 70 and 45% and his performance represented a negative trend. John's performance with

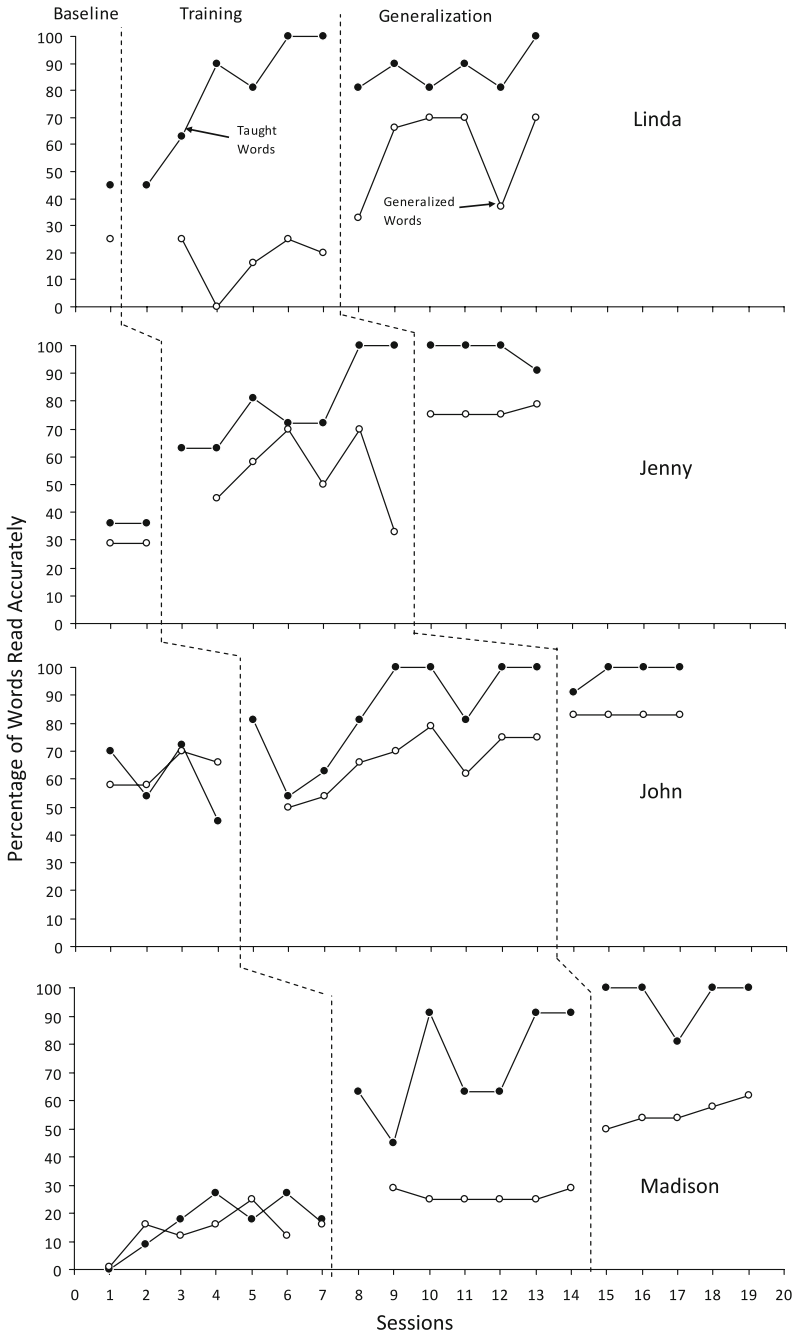


Fig. 1 During training phase assessment of taught and generalization words occurred using non-colored flashcards. During generalization phase, assessment of taught words occurred using non-colored flashcards. Assessment of generalization words occurred using colored flashcards

generalization words trended slightly in a positive direction but level of performance was relatively stable ranging between 58 and 70% accuracy. Once the training phase began, John's level of accuracy quickly jumped to 81% with training words but then fell dramatically during the second session. From that point on, John's performance gradually increased and four of the last five sessions included reading accuracy rates of 100%. Regarding his accuracy with generalization words, his performance dipped to slightly below baseline levels but gradually increased over the course of this phase. During the last five sessions of this phase, John's performance with generalization words leveled off generally between 70 and 80% accuracy. Thus, only slight increases were demonstrated by John with generalization words indicating that minimal generalization of intervention effects to these words occurred. Finally, during the generalization phase, John's performance on both training and generalization words was highly stable. Accuracy with training words remained at 100%, while accuracy with generalization words showed a slight increase to 83% across all sessions.

Baseline data for Madison were relative stable, showing a slight positive trend for both training and generalization words. However, once the training phase began, Madison demonstrated a substantial gain in the percentage of training words read accurately. Her initial performance increased to 63% accuracy with 5 of the next 6 data points ranging between 63 and 91% accuracy. Thus, although her performance was somewhat inconsistent, it represented a substantial change in level and trend. Unfortunately, only slight positive gains were demonstrated with generalization words during the training phase. Madison's level of accuracy with these words was highly stable ranging between 25 and 29%. Finally, during generalization, Madison's accuracy in reading taught words was stable and slightly improved compared to the intervention phase as 4 of 5 data points showed 100% accuracy. Positive gains were seen with generalization words as her performance showed an initial increase in accuracy to 50% once generalization techniques were implemented. Her performance reading the generalization words during remaining sessions showed a slight increasing trend with scores ranging between 54 and 62% accuracy.

Discussion

Students who struggle to learn how to read often have difficulty applying developing skills across response contexts, thus they may struggle when asked to read materials that they have had minimal exposure to or practice with. Explicit strategies or interventions that attempt to promote generalization of reading skills are somewhat limited within the behavioral education literature. Studies that do exist have met with mixed results and possibly suggest that the stimuli (letters) intended to cue desired behavior (accurate reading) may not be sufficient.

Results of this study showed that three of the four participants demonstrated slight improvement accurately reading the generalization words when the intervention phase was implemented. In other words, these three students (Jenny, John, and Madison), showed some spontaneous generalization as reading

performance improved with words not included in the intervention. The fourth student, Linda, demonstrated no improvement with generalization words during the intervention phase. Despite the slight gains for 3 of the 4 participants, these gains are not likely to be considered sufficient. However, when a programmed generalization procedure was added (use of the color cue on the flashcards), 3 participants (Linda, Jenny, and Madison) showed improvement in their accurate reading of generalization words beyond that gained by the intervention alone. Moreover, performance with the generalization words more closely approximated the level of accuracy that Jenny had in reading taught words (i.e., words used as part of the intervention). Therefore, even when spontaneous generalization occurred, these results demonstrated that further improvements could be made when an explicit generalization strategy was employed.

If interventions fail to effectively promote generalization then student success is likely to be limited to those conditions highly associated with the intervention context/stimuli. While continuous implementation of an intervention across contexts may produce excellent results, it is not reasonable to believe that this level of implementation would be feasible in most educational settings. Thus, low effort alternatives that produce improvements in generalized student performance would seem important to identify. Current results suggest that a common stimulus procedure using a simple cue (i.e., colors highlighting part of a word) was an effective strategy for increasing reading accuracy. Through the training procedure, the color cue resulted in increased stimulus control such that reading generalization words was performed at a higher level when the cue was present. This finding is consistent with previous research demonstrating the success of common stimuli in producing generalization (Mesmer et al. 2007). It should not be concluded, however, that a color cue would be more effective than other cues in enhancing generalization effectiveness. Other types of cues (visual or auditory) may be just as effective.

With regard to reading similar words within word families, it appears that cues that highlight critical stimulus features such as orthographic and phonological similarities may assist in facilitating the generalization process. Importantly, however, use of the highlighting alone during training procedures was not sufficient to produce substantial improvements with generalized words. Such improvement was only noted once the color cue was added to the generalized words during assessment procedures. Finally, while accurate reading of generalized words improved when the color cue was added, it ultimately would be important to fade the color cue in order to ensure accurate reading in a variety of authentic reading contexts.

Finally, results from this study add to a scant literature base examining the use of behaviorally based strategies to facilitate generalization with academic responses. Only a few instances exist illustrating use of the various types of generalization strategies initially articulated by Stokes and Baer (1977) over 30 years ago. It would seem that an emphasis on exploration of these strategies is timely considering the current context of public education. Specifically, the burgeoning use of response to intervention (RTI) models within the schools places the need for effective intervention programming at the forefront of work by researchers and practitioners.

Struggling learners have difficulty generalizing newly acquired skills and they will be the prime recipients of RtI services. It should not be assumed that these students will spontaneously generalize skills acquired during the RtI process across academic contexts. Therefore, these students would benefit from a research base that has more thoroughly explored generalization strategies.

Limitations and Future Research

It is important to note several limitations of this study. First, for 2 of the 4 participants the number of baseline data points was not sufficient to establish stability in the data. While the data points that were collected confirm reported difficulties in reading, it is possible that the performance of these two participants may have fluctuated with additional data collection. As a result of not collecting additional baseline data points, it is possible that other variables may have contributed to the change in data once the intervention was implemented including variation in performance that may have more accurately reflected baseline levels of reading accuracy. Second, because the same limited number of words (35 in total) were used throughout the duration of the study, it is possible that the gains seen for both taught and generalization words were simply a result of continuous exposure and practice with these words. Evidence indicating significant increases in accuracy levels immediately following implementation of either training or generalization procedures for 3 of the 4 participants suggests that this was unlikely the case. However, this limitation might be addressed in future studies by using a larger number of words or nonsense words thereby increasing the likelihood of novel exposure or by using a large number of nonsense words. Another option might be to use an alternating treatment design so that a discrepancy in the effects could be detected more quickly prior to the possibility of a practice effect occurring.

Although not a methodological limitation, another weakness of this study, from a practical perspective, was that the generalization cue was never faded. It is unlikely that practitioners would be comfortable incorporating any cue for an extended period of time. Thus, the results would be more socially valid if the positive effects were to continue in the absence of the generalization stimulus. A cue such as the one used in this study could be easily faded in practice.

A fourth limitation of this study relates to the small and relatively homogenous sample. Although the strategies used in this study had a positive impact on study participants, it would be erroneous to conclude that these results would necessarily be replicated with larger and broader populations of students. Future research replicating the results of this and like studies will be necessary to address this weakness.

Finally, it is important to note that while data were collected on procedural integrity, including whether participants responses were sorted into “correct” or “incorrect” piles and whether words in each of these piles were subsequently recorded on the score sheet, integrity checks did not include whether the word cards in these two piles were accurately recorded on the score sheet (e.g., whether words in the correct/incorrect piles were accurately marked as correct/incorrect on the

score sheet). As a result inter-scorer agreement could not be calculated and therefore any scoring errors that may have occurred were not noted or addressed.

As suggested earlier, future research might examine ways of enhancing intervention effectiveness through the examination of generalization strategies. While comparison of various types of stimuli that are most effective in producing target responses across academic contexts would be useful, many other generalization techniques require exploration as well. Moreover, a comparison of different classes of generalization strategies might be useful, particularly as they relate to different types of academic difficulties. For example, the use of common stimuli largely reflects a strategy that changes the generalization context after instruction. In other words, some change is made in the materials or environment (i.e., addition of a cue) following instruction, with the idea that the change will improve the response. Other strategies, however, might employ cues prior to instruction or might incorporate changes within instructional procedures. Finally, future research might focus upon methods for reliably quantifying the change that occurs as a result of implementing a generalization strategy. Because most strategies will require some level of extra effort it might be important to determine whether the effort is worth the subsequent level of improvement. Within this study, it certainly could be argued that the minimal effort required to use a color cue during generalization conditions was worth the increase in generalization that was evidenced.

References

- Ardoin, S., Eckert, T., & Cole, C. (2008). Promoting generalization of reading: A comparison of two fluency-based interventions for improving general education student's oral reading rate. *Journal of Behavioral Education, 17*, 237–252.
- Berends, I., & Reitsma, P. (2006). Remediation of fluency: Word specific or generalized training effects? *Reading and Writing, 19*, 221–234.
- Bowey, J. A., Vaughan, L., & Hansen, J. (1998). Beginning readers' use of orthographic analogies in word reading. *Journal of Experimental Child Psychology, 68*, 108–133.
- Daly, E., III., Bonfiglio, C., Mattson, T., Persampieri, M., & Foreman-Yates, K. (2005). Refining the experimental analysis of academic skills deficits: Part I. An investigation of variables that affect generalized oral reading performance. *Journal of Applied Behavior Analysis, 38*, 485–497.
- Daly, E., Chafouleas, S., III., Persampieri, M., Bonfiglio, C., & La Fleur, C. (2004). Teaching phoneme segmenting and blending as critical early literary skills: An experimental analysis of minimal textual repertoires. *Journal of Behavioral Education, 13*, 165–178.
- Daly, E., III., 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.
- Daly, E., III., Shroder, J., & Robinson, A. (2006). Testing treatments for oral reading fluency problems: Two case studies. *Journal of Evidence Based Practices in the Schools, 7*, 4–23.
- DeRose, J., DeSouza, D., & Hanna, E. (1996). Teaching reading and spelling: Exclusion and stimulus equivalence. *Journal of Applied Behavior Analysis, 29*, 451–469.
- Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology, 42*, 73–83.
- Mesmer, E., Duhon, G., & Dodson, K. (2007). The effects of programming common stimuli for enhancing stimulus generalization of academic behavior. *Journal of Applied Behavior Analysis, 40*, 553–557.

- National Reading Panel (NRP). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- Nist, L., & Joseph, L. (2008). Effectiveness and efficiency of flashcard drill instructional methods on urban first-graders word recognition, acquisition, maintenance, and generalization. *School Psychology Review, 37*, 294–308.
- Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stokes, T., & Baer, D. (1976). Preschool peers as mutual generalization facilitating agents. *Behavior Therapy, 7*, 549–556.
- Stokes, T., & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349–367.
- Thaler, V., Ebner, E., Wimmer, H., & Landerl, K. (2004). Training reading fluency in dysfluent readers with high reading accuracy: Word specific effects but low transfer to generalization words. *Annals of Dyslexia, 54*, 89–113.
- Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated readings. A meta-analysis. *Remedial and Special Education, 25*, 252–261.
- Valleley, R., & Shriver, M. (2003). An examination of the effects of repeated readings with secondary students. *Journal of Behavioral Education, 12*, 55–76.
- Yurick, A., Robinson, P., Cartledge, G., Lo, Y., & Evans, T. (2006). Using peer-mediated repeated readings as a fluency-building activity for urban learners. *Education and Treatment of Children, 29*, 469–506.