JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2005, **38,** 511–527 NUMBER 4 (WINTER 2005)

# *ANALYSIS OF RESPONSE REPETITION AS AN ERROR-CORRECTION STRATEGY DURING SIGHT-WORD READING*

APRIL S. WORSDELL

SOUTHERN ILLINOIS UNIVERSITY

AND

BRIAN A. IWATA, CLAUDIA L. DOZIER, ADRIENNE D. JOHNSON, PAMELA L. NEIDERT, AND JESSICA L. THOMASON

UNIVERSITY OF FLORIDA

A great deal is known about the effects of positive reinforcement on response acquisition; by contrast, much less research has been conducted on contingencies applied to errors. We examined the effects of response repetition as an error-correction procedure on the sight-word reading performance of 11 adults with developmental disabilities. Study 1 compared single-response (SR) repetition and multiple-response (MR) repetition, and results showed that all 6 participants acquired more sight words with the MR procedure. Study 2 compared MR error correction following every incorrect response (continuous) and following one third of incorrect responses (intermittent), and results showed that all 6 participants acquired more sight words when error correction was continuous. Study 3 compared MR error correction in which errors required practice of the training word (relevant) versus a different word (irrelevant), and results showed that 3 of 9 participants showed better performance under the relevant condition; however, all participants showed improvement even under the irrelevant condition. Findings are discussed in terms of the behavioral processes by which error correction may enhance performance during acquisition.

DESCRIPTORS: sight-word reading, error correction, negative reinforcement, stimulus control

---

When the desired outcome of intervention is skill acquisition, two changes in performance must occur: Correct responses must increase, and incorrect responses must decrease. Typical methods of instruction include explicit consequences for both types of responses and have been applied with success across a wide range of performances, such as manual signing (Carr & Kologinsky, 1983; Remington & Clarke, 1983), self-care skills (Horner & Keilitz, 1975; Nutter & Reid, 1978), spelling (Birnie-Selwyn & Guerin, 1997; Neef, Iwata, & Page, 1980), and reading (Pany, McCoy, & Peters, 1981; Rosenbaum & Breiling, 1976).

Increases in correct responding observed during acquisition have been attributed primarily to the influence of positive reinforcement (Cooper, Heron, & Heward, 1987; Miltenberger, 2003); as a result, much less research has been conducted on consequences for errors. It is possible, however, that consequences for errors not only decrease incorrect responses but also contribute to increases in correct responses. Therefore, more detailed analyses of the effects of error correction are warranted.

The most common error-correction procedures involve the instructor prompting the learner to emit a correct response following the occurrence of an error. Researchers have

This research is based on a dissertation submitted by the first author to the University of Florida in partial fulfillment of requirements for the PhD degree and was supported in part by a grant from the Florida Department of Children and Families. We thank Timothy Hackenberg, Cecil Mercer, Scott Miller, and Timothy Vollmer for their helpful comments on a previous version of the manuscript.

Reprints may be obtained from April Worsdell, Rehabilitation Institute, Southern Illinois University, Mailcode 4609, Carbondale, Illinois 62901 (e-mail: worsdell@siu.edu).

doi: 10.1901/jaba.2005.115-04

shown improvements in correct performance with the use of vocal prompts to continue the task (Cowley, Green, & Braunling-McMorrow, 1992; Koegel & Egel, 1979), imitative models of correct performance (Carr & Kologinsky, 1983; Wheeler & Sulzer, 1970), and manual prompts to produce the correct response (Carey & Bucher, 1983; Remington & Clarke, 1983).

In reading instruction, emphasis has been placed on prompting strategies that involve exposure to the misread word following an error. In whole-word or word-supply correction, for example, the instructor recites the entire word after an error is made and requires the learner to repeat it. Results of several studies have shown that response repetition of this type, occurring once or several times per error, facilitates performance during sight-word reading tasks (Barbetta, Heward, & Bradley, 1993; Espin & Deno, 1989; Rose, McEntire, & Dowdy, 1982; Singh & Singh, 1986). It is not clear, however, the degree to which amount of repetition influences sight-word reading performance, and this question was addressed in Studies 1 and 2 of the present research.

To the extent that error correction actually contributes to increases in correct responding, it may do so in at least two ways. First, error correction may simply provide additional opportunities for the correct response to occur under the appropriate stimulus conditions, thereby enhancing stimulus control over correct responding on subsequent trials. Alternatively, error correction could function as punishment for incorrect responses; if so, the only way to avoid correction trials is to emit a correct response, which is strengthened through negative reinforcement. The avoidance characteristics of correct responding become more apparent when one considers that error correction typically follows not only errors but also the absence of a response. Thus, it is not simply the omission of an error that avoids error correction, but the occurrence of a correct response. The relative influence of these two processes—additional practice and avoidance—in facilitating stimulus control are difficult to identify in error-correction procedures that involve response repetition because both features are present. Three studies, however, illustrated a method for separating these effects by including a correction procedure that did not involve practice of correct responses.

Axelrod, Kramer, Appleton, Rockett, and Hamlet (1984) compared the effects of "relevant" and "irrelevant" error-correction procedures on spelling performance. During the relevant condition, children were required to correctly spell all misspelled words from a weekly pretest, write their parts of speech and phonetic spelling, and use them in five sentences. The irrelevant condition was identical to the relevant condition except that the words were not taken from the weekly pretest (i.e., the words were not part of the weekly spelling lesson). Both procedures produced increases in spelling accuracy on weekly tests. A similar procedure was used by Rodgers and Iwata (1991) to examine acquisition during a visual discrimination (match-to-sample) task. During the relevant (practice) condition, errors were followed by repetition of the trial until a correct response occurred. In the irrelevant (avoidance) condition, errors resulted in the presentation of trials on an unrelated task (color matching). Results showed that 4 participants performed better when relevant stimuli were presented during error correction, whereas the other 3 participants performed better when irrelevant stimuli were presented. Finally, Cuvo, Ashley, Marso, Zhang, and Fry (1995) compared the effects of relevant and irrelevant error correction on the spelling and reading performance of high school special education students. During the relevant condition, the participant was required to write (or verbalize) the misspelled (or misread) word. During the irrelevant condition, the participant was required to practice spelling (or verbalizing) a more difficult nontraining word. Results

Table 1
Demographic Information

| Participant | Age | Diagnosis | Reading level | Studies |
|---|---|---|---|---|
| Ernie | 23 | Moderate MR,[a] Prader-Willi syndrome | Grade 1 | 1, 2, 3 |
| Seth | 23 | Moderate MR, Down syndrome | Grade 3 | 1, 2, 3 |
| Robin | 37 | Moderate MR | Grade 1 | 1 |
| Ariel | 23 | Mild MR, Prader-Willi syndrome | Grade 2 | 1, 3 |
| Hayley | 33 | Mild MR, Prader-Willi syndrome | Grade 3 | 1, 3 |
| Justin | 21 | Mild MR | Grade 2 | 1 |
| Becky | 20 | Moderate MR, Prader-Willi syndrome | Grade 3 | 2, 3 |
| Tess | 27 | Mild MR, Prader-Willi syndrome | Grade 3 | 2, 3 |
| Mark | 36 | Moderate MR | Grade 2 | 2, 3 |
| Kara | 29 | Mild MR, Prader-Willi syndrome | Grade 5+ | 2, 3 |
| Maisy | 48 | Moderate MR, cerebral palsy, seizure disorder | Grade 3 | 3 |

[a] MR = mental retardation.

showed that 9 of the 10 participants acquired spelling and reading skills equally quickly during relevant and irrelevant practice conditions (1 student's performance was slightly better during the relevant condition). Study 3 of the present research attempted to systematically replicate the findings of Axelrod et al., Rodgers and Iwata, and Cuvo et al. and to extend them to the sight-word reading performance of adults with mental retardation.

## GENERAL METHOD

### *Participants and Setting*

Eleven adults with developmental disabilities participated in one or more of three studies (see Table 1 for demographic information). All were adults employed at a day vocational program, and all but 3 (Ernie, Seth, and Maisy) were involved in a continuing education program on the premises of the workshop. The education program consisted of instruction in daily living skills (e.g., money management, safety skills), although some academic instruction was provided. Individuals were selected for participation based on caregiver or teacher report of a need for improvement in sight-word reading skills. None of the participants received explicit sight-word instruction throughout the study other than during experimental sessions.

Graduate and undergraduate psychology students served as experimenters. Sessions were conducted 3 to 5 days per week in quiet areas adjacent to the participants' work sites. Session length varied both within and across the three studies, although an individual session never lasted more than 25 min (range, 2 to 24 min). All training and testing sessions were conducted individually, with the experimenter and participant seated across from one another at a table and observers seated close enough to hear the participant's responses.

### *Pretesting and Sight-Word Selection*

A pool of words was obtained from lists of grade-level sight words and was assembled by the first author to serve as potential training stimuli. Words were grouped in lists and were designated either by a specific grade level (1 to 5), or by a 5+ if the words were above the fifth-grade reading level. A pretest similar to that described by Barbetta et al. (1993) was conducted with each person to identify novel (i.e., unknown) words. Word lists selected for pretesting were at least one grade level above a participant's current reading level. During the pretest, participants were instructed to try their best to read each word without making a mistake. Words were presented one at a time on index cards, and participants were given 5 s to respond. No consequences were delivered for either correct or incorrect responses, but noncontingent praise (e.g., "nice reading") was delivered intermittently to maintain attention to the task. Words read correctly during the first pretest were omitted from the word list, and the remaining words were pretested a second time. Words read incorrectly during both pretests were considered unknown and were used as training stimuli. Based on results of the pretest, participants were designated at a specific reading level. For example, if a participant correctly read a high percentage of Grade 2 words but was unable to read the majority of Grade 3 words, then he or she was considered to be at a Grade 2 reading level.

An additional pretest was conducted to identify acceptable alternative word pronunciations by Ariel and Jeff, who had minor speech difficulties. The experimenter read unknown words aloud (without showing any word cards), and the participant was prompted to repeat each word. If the participant was unable to emit a close approximation to an unknown word, then the word was omitted from the list of training words. However, if the participant's pronunciation of a word sufficiently resembled the correct pronunciation (e.g., saying "ceweal" after hearing "cereal"), then the word was kept in the pool of training words, and the alternative pronunciation was regarded as correct.

*Reinforcer Identification*

A multiple-stimulus assessment (DeLeon & Iwata, 1996) was conducted with all participants to identify reinforcers for correct responses during training. Edible items were used for all participants except those with dietary restrictions; leisure activities or calorie-free drinks were used for the latter group. The number of items assessed ranged from 7 to 16 across participants, and the assessment was conducted once (Ernie, Ariel, Hayley, Justin, Becky, Tess, and Kara) or was repeated three times (Seth, Robin, Mark, and Maisy).

Based on results of the assessment, one or more stimuli were selected for use during training. If the assessment included only edible items, the most highly preferred item was used as a consequence for correct responses. If the assessment was limited to leisure items, a token system was developed to dispense reinforcers. A reinforcer menu was created from stimuli identified as moderately to highly preferred during the assessment. Each correct response made during baseline or training resulted in the delivery of one ticket, and tickets could be exchanged at the end of a session for one or more of the activities listed on the menu.

*Training Sessions*

Each of the three studies involved a baseline (no error correction) condition followed by a comparison between two error-correction procedures. An initial set of eight unknown words was assigned to the baseline condition and to each of the two training conditions. Words within a set were presented three times per session in a semirandom order (yielding 24 trials). In an attempt to equate word difficulty across conditions, an equal number of one- or multisyllable words were included within each word set. Words were printed by hand on index cards (7.6 cm by 12.7 cm), and different colored cards were used to designate the baseline set and the two training sets. To ensure some familiarity with the correct pronunciation of words, whenever a new word was introduced into a word set, the card was shown to the participant and the word was read aloud prior to daily training sessions.

A mastery criterion was developed so that correctly read words in each set were continually replaced by new, unknown words. A word was considered to be mastered if it was read correctly on every presentation within a training session (i.e., three consecutive times during one session). Once a word was mastered, it was removed from the array and was replaced with a new, unknown word. Words that were mastered during baseline were excluded from the rest of the experiment (i.e., they were not used during training conditions).

*Retention Tests*

Two types of tests were administered to determine whether words mastered during training were retained. The first test (short term) was administered before daily training sessions and included all words that were mastered during the previous day's training sessions. On most days, a short-term retention test was conducted approximately 24 hr after one or more words were mastered during training. However, weekends and unscheduled absences from the work site sometimes increased the time between word mastery and short-term testing to a maximum of 72 hr. If more than 72 hr elapsed between word mastery and testing, a short-term retention test was not administered. The second test (long term) was administered once per week and included all words that had been mastered during the previous week's training sessions.

Prior to each short- or long-term retention test, the participant was instructed, "Here are some words that you've learned in our reading sessions. You won't get any [reinforcer], but please try your best to read each word without making a mistake." The experimenter then shuffled the mastered word cards and presented each one for 5 s. No consequences were provided for either correct or incorrect responses during testing.

*Response Measurement and Reliability*

Graduate and undergraduate psychology students used laptop computers to record data on (a) the frequency of correct responses, incorrect responses, reinforcer deliveries, and error-correction presentations and (b) the percentage of intervals during which error correction occurred during training. A response was scored as correct if the participant accurately pronounced a word within 5 s of its presentation. A response was scored as incorrect if (a) the participant's vocal response did not match the word displayed on the card (including mispronunciations, omissions, or substitutions) or (b) the participant said "I don't know" or failed to respond within 5 s of the presentation of a card.

Across all three studies, the primary dependent variable was the cumulative number of words meeting the mastery criterion. Observers kept a running tally of the words mastered (i.e., pronounced correctly three consecutive times during a given training session). Retention test data also were collected using a paper-and-pencil data sheet, and observers recorded whether a word was correct or incorrect by placing a check mark in the appropriate column.

Interobserver agreement was assessed by having a second observer simultaneously but independently collect data with the primary observer during 33% of all training sessions (35% of sessions in Study 1, 32% of sessions in Study 2, and 33% of sessions in Study 3), 43% of all short-term retention tests (47% in Study 1, 38% in Study 2, and 43% in Study 3), and 47% of all long-term retention tests (50% in Study 1, 44% in Study 2, and 47% in Study 3).

Percentage agreement was calculated in three ways. Agreement for the paper-and-pencil measures (mastered words and correct responses during testing) was calculated by dividing the smaller number of responses by the larger number of responses and multiplying by 100%. Mean agreement was 99% (range, 97% to 100%) for mastered words, 96% (range, 86% to 100%) for correct responses during short-term retention tests, and 94% (range, 78% to 100%) for correct responses during long-term retention tests.

Agreement for frequency measures (correct and incorrect responses during training, reinforcer deliveries, and error-correction presentations) was calculated by first dividing each session into 10-s intervals. The smaller number of responses in each interval was divided by the larger number of responses, and these proportions were averaged across intervals and multiplied by 100%. Mean agreement was 95% (range, 91% to 99%) for correct responses, 93% (range, 89% to 97%) for incorrect responses, 96% (range, 92% to 99%) for reinforcer deliveries, and 95% (range, 89% to 98%) for error-correction presentations.

Finally, agreement for the partial-interval measure (percentage of intervals during which error correction occurred) was calculated by dividing the total number of agreement intervals by the total number of intervals in a session and multiplying by 100%. Mean agreement for error-correction occurrence was 93% (range, 87% to 97%).

## STUDY 1: NUMBER OF REPETITIONS DURING ERROR CORRECTION

Foxx and Jones (1978) presented data indicating that spelling performance was enhanced by increasing the amount of error correction following the occurrence of an incorrect response. By contrast, Cuvo et al. (1995) found that neither spelling nor reading skills were acquired more quickly when error-correction amount was increased. The purpose of this study was to further examine the speed of sight-word acquisition as a function of the number of response repetitions following an error. Ernie, Seth, Robin, Ariel, Hayley, and Justin participated in Study 1.

### *Method*

Baseline lengths were staggered to conform to a multiple baseline across subjects design. Following baseline, two error-correction conditions with differing amounts of response repetition were alternated in a multielement design.

*Baseline (no error correction).* A set of eight words was presented three times per session in random order (24 word presentations). Word cards were presented one at a time, and participants were instructed to "read this word." Correct responses resulted in the delivery of a preferred edible item (Seth and Robin) or a ticket that was exchangeable for postsession access to preferred items or activities (Ernie, Ariel, Hayley, and Justin). Incorrect responses were ignored and resulted in presentation of the next word card. No other instructions or prompts were delivered.

*Single-response (SR) repetition.* Both the presentation of cards and the consequences for correct responses were identical to the baseline condition. Rather than ignoring incorrect responses, however, the experimenter implemented an SR repetition procedure following each error. SR consisted of the experimenter modeling the correct pronunciation of a word and requiring the participant to repeat the word once. When a participant made an incorrect response, the experimenter said, "No, the word is —. Say —." The card was displayed to the participant for the duration of SR repetition. Following the participant's correct or incorrect vocal imitation of the word, the experimenter presented the next card.

*Multiple-response (MR) repetition.* The procedures were similar to those in the SR condition, except that errors were corrected using an MR repetition procedure. MR consisted of the experimenter modeling the correct pronunciation of a word and prompting the participant to repeat the word five times in the presence of the card.

### *Results and Discussion*

Figure 1 shows the cumulative number of words mastered by each participant in each condition. Participants either mastered very few words (Ariel, Ernie, and Hayley) or ceased to show progress (Seth, Justin, and Robin) during baseline. Their rate of word mastery increased when SR and MR procedures were implemented, with the best performance being observed during MR in all cases. Ernie showed the largest difference in performance between the two error-correction conditions, mastering 70% more words in MR than in SR. In the other cases, differences were not as large. Across the remaining participants, an average of 33% more words were mastered during the MR condition relative to the SR condition. It is interesting to

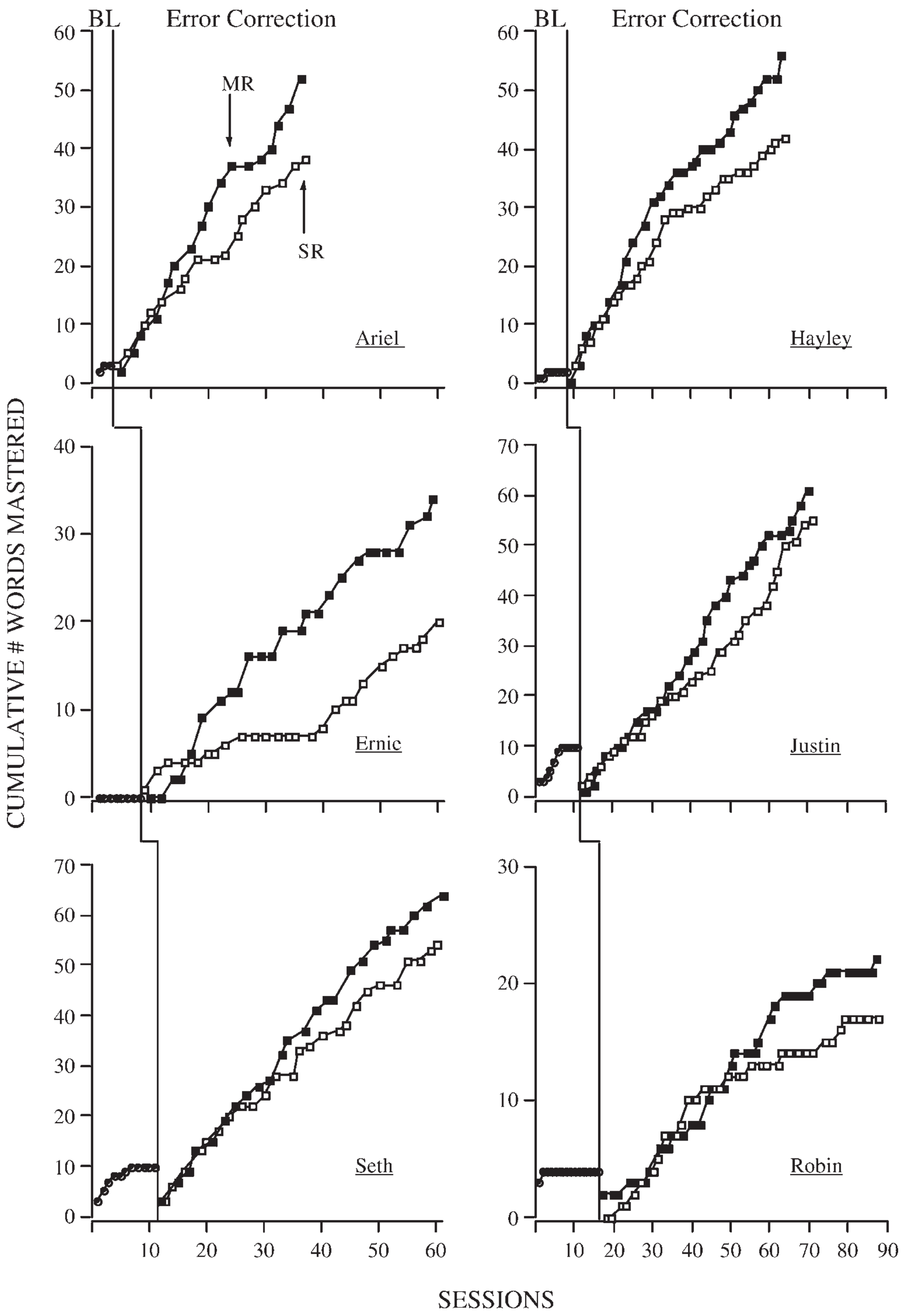


Figure 1. Cumulative number of words mastered during baseline (BL), single-response (SR) repetition, and multiple-response (MR) repetition.

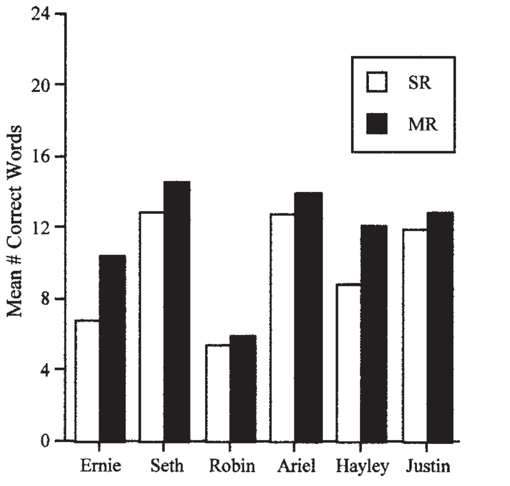


Figure 2. Mean number of correct words read per session during SR and MR.

Table 2
Percentage of Words Retained on Short- and Long-Term Tests during Single-Response (SR) Repetition and Multiple-Response (MR) Repetition

| | Short-term retention test | | Long-term retention test | |
|---|---|---|---|---|
| Participant | SR | MR | SR | MR |
| Ernie | 51 | 72 | 30 | 63 |
| Seth | 78 | 87 | 67 | 77 |
| Robin | 53 | 64 | 53 | 73 |
| Ariel | 85 | 81 | 68 | 74 |
| Hayley | 85 | 80 | 58 | 62 |
| Justin | 79 | 92 | 69 | 81 |

note that participants' acquisition rates under the two error-correction conditions were similar initially; MR produced superior effects only after repeated exposure to the two procedures (over 40 sessions for Robin).

Figure 2 shows the mean number of words read correctly per session by each participant during SR and MR error correction (out of a total of 24 words). More correct responses were made in the MR condition than in the SR condition in every case. Thus, not only were more words mastered during MR but more total words were read correctly in this condition compared to SR.

In terms of efficiency, the SR and MR procedures were roughly equivalent. MR sessions took more time to complete, although differences in session length were small (SR $M$ = 4.9 min; MR M = 6.4 min). In addition, a slightly larger percentage of session intervals was devoted to correcting errors during MR than during SR for 4 of the 6 participants (SR $M$ = 52%; MR $M$ = 55%).

Table 2 shows results of the retention tests. On short-term tests, 4 of the 6 participants retained more words under the MR condition. On long-term tests, all participants retained more MR words than SR words, although in Hayley's case, the percentage difference in long-term retention was marginal (SR $M$ = 58%; MR $M$ = 62%).

These results replicate findings from a number of studies indicating that error correction (SR) facilitates performance. The MR procedure was superior to SR along a number of dimensions. MR resulted in (a) more cumulative words mastered, (b) more words read correctly per session, (c) and more words retained both short and long term. These findings extend the work of Foxx and Jones (1978) by showing that a larger amount of error correction resulted in better performance, but they are inconsistent with other research showing no enhanced effects when amount of error correction was increased (Cuvo et al., 1995).

A possible limitation of MR is the amount of time required to complete the procedure. More training time was consumed by MR relative to SR, although the time differences were small. MR error correction (five repetitions) necessarily took longer to administer than SR error correction (one repetition). However, negligible differences were observed between the percentage of intervals devoted to error correction during SR and MR. Because a nearly equivalent amount of time was spent correcting errors during SR and MR, the difference in training time may have been due to the increased delivery (and consumption) of reinforcers for correct responses.

Finally, although the MR procedure consistently produced better performance than did the SR procedure, the relation between number of repetitions and extent of improvement is unknown. For example, Cuvo et al. (1995) observed no differences in acquisition rate with comparisons of 5, 10, and 15 repetitions (Experiment 2) and one and five repetitions (Experiments 3 and 4). We selected five repetitions for the MR procedure somewhat arbitrarily as an amount clearly larger than one yet not too large to be disruptive. It is possible, therefore, that similar effects would have been observed if the MR condition consisted of only two repetitions.

Because research on error correction has focused primarily on comparisons between error correction and no correction, experimental conditions have involved correction following every error. However, consistent implementation may not be practical or even possible outside the laboratory setting. Teachers often do not have the support staff available to deliver consistent feedback to all students during instruction. Under typical training conditions, it may be feasible to implement error-correction procedures for a portion of a student's mistakes but not for all of them. Therefore, it may be important to consider the effects of intermittent error correction, which was done in Study 2.

## STUDY 2: SCHEDULE OF ERROR CORRECTION

In Study 1, we manipulated amount of error correction by varying the number of word repetitions required per incorrect response (one repetition in SR vs. five repetitions in MR). Varying the consistency with which consequences are arranged for errors may be seen as an alternative way to manipulate the amount of error correction. In Study 2, we compared the effects of consistent fixed-ratio (FR) 1 and intermittent variable-ratio (VR) 3 schedules of error correction. Seth, Ernie, Tess, Kara, Mark, and Becky participated in Study 2.

### *Method*

During the two training conditions, an MR repetition procedure was used in which five repetitions of the correct word were required; the conditions differed only with respect to the proportion of errors corrected per session.

*Baseline (no error correction).* A set of eight words was presented three times each (24 words per session). The experimenter presented one word card at a time and instructed the participant to read the word. When a correct response occurred, either an edible item (Seth and Mark) or a ticket (Ernie, Tess, Kara, and Becky) was delivered. When an incorrect response occurred, it was ignored, and the next card was presented. No other instructions or prompts were delivered.

*MR (continuous).* Both the presentation of cards and the consequences for correct responses were identical to the baseline condition. Contingent on each occurrence of an incorrect response, the MR procedure was implemented (i.e., the experimenter prompted the participant to repeat the word five times in the presence of the card). Prompts to repeat the word continued regardless of the participant's correct or incorrect pronunciation of the word. Following the participant's fifth repetition of the word, the next card was presented.

*MR (intermittent).* Procedures used in this condition were similar to those of the continuous MR condition, except that error correction was not implemented after every incorrect response. Instead, an average of every third error was corrected with the MR procedure (i.e., a VR 3 schedule of error correction). Errors that were not corrected were ignored.

### *Results and Discussion*

Figure 3 shows the cumulative number of words mastered during baseline and error-correction conditions. Five of the 6 participants showed little or no improvement during baseline (Tess was the exception). The introduction of both error-correction procedures resulted in increases in all participants' word

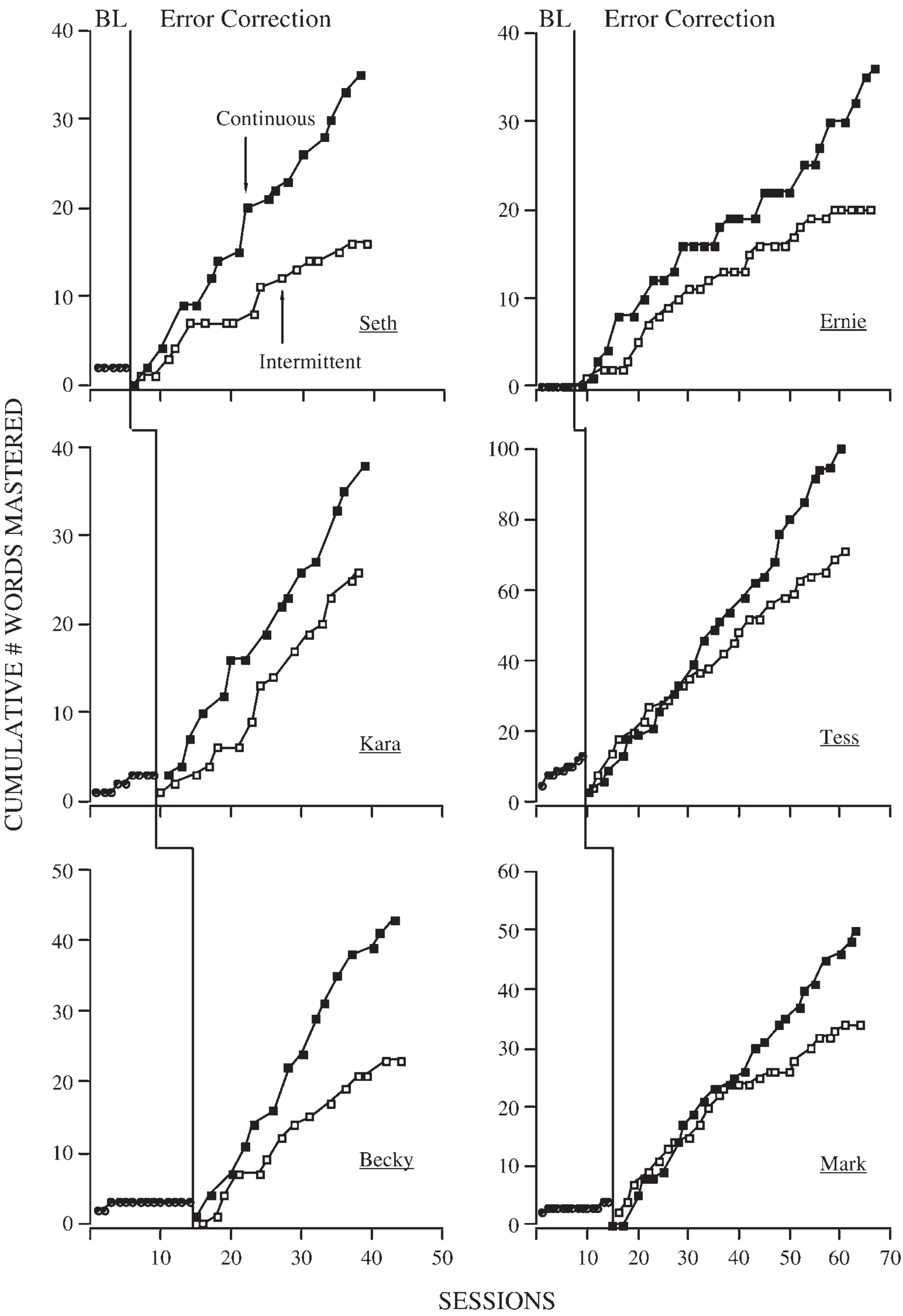


Figure 3. Cumulative number of words mastered during BL, continuous MR, and intermittent MR.

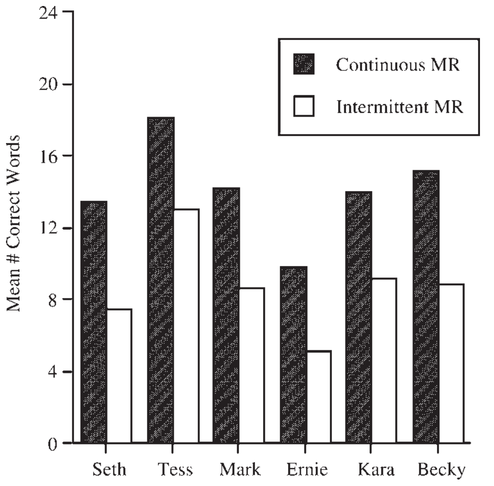


Figure 4. Mean number of correct words read per session during continuous MR and intermittent MR.

Table 3
Percentage of Words Retained on Short- and Long-Term Tests during Continuous MR and Intermittent MR

| | Short-term retention test | | Long-term retention test | |
|---|---|---|---|---|
| Participant | Continuous | Intermittent | Continuous | Intermittent |
| Seth | 85 | 69 | 82 | 82 |
| Tess | 94 | 94 | 86 | 84 |
| Mark | 73 | 75 | 58 | 50 |
| Ernie | 69 | 58 | 49 | 58 |
| Kara | 90 | 90 | 83 | 75 |
| Becky | 86 | 91 | 91 | 91 |

mastery, with the best performance observed under the continuous MR condition in all cases. Seth showed the largest difference in acquisition between the two conditions, mastering 35 words in the continuous MR condition and 16 words in the intermittent MR condition, which represented a 119% improvement. Similar although less pronounced improvements also were observed for Kara (46%), Becky (87%), Ernie (80%), Tess (38%), and Mark (47%).

Figure 4 shows the mean number of correct responses emitted per session during continuous and intermittent MR error correction. All 6 participants read more words correctly when every error was followed by the MR procedure (continuous MR) rather than when a varying proportion of errors was corrected (intermittent MR).

Continuous MR sessions took longer to complete than did intermittent MR sessions (continuous $M$ = 6.5 min; intermittent $M$ = 5.2 min) for 5 of the 6 participants; Tess's session lengths were the same across the two conditions. For all 6 participants, a larger percentage of session time was devoted to correcting errors in the continuous MR condition than in the intermittent MR condition (continuous $M$ = 46%; intermittent $M$ = 28%).

Table 3 shows results of the retention tests. On short-term tests, Seth and Ernie retained more words mastered during the continuous MR condition, whereas Mark and Becky retained a slightly larger percentage of words mastered during the intermittent MR condition. Tess and Kara retained an equal percentage of words across the two conditions. On long-term tests, Tess, Mark, and Kara retained a slightly larger percentage of words during the continuous MR condition, whereas Ernie retained more words during the intermittent MR condition. Seth and Becky showed no difference in percentage retention across the two conditions. Overall, results of the retention tests were inconsistent in that there was little correspondence between training performance and test results. This perhaps was due to the fact that participants mastered many more words under the continuous MR condition and subsequently had to read more words from that condition on retention tests.

Results of this study were consistent with those from Study 1 indicating that larger amounts of error correction enhance performance. Nonetheless, all participants (perhaps Tess was an exception) in this study showed improvement over baseline even in the intermittent MR condition. Thus, the results also suggest that instructors should take advantage of opportunities to correct errors even occasionally.

## STUDY 3: THE FUNCTIONS OF ERROR CORRECTION

Error-correction procedures typically involve practice of correct performance. In this sense, the error-correction response is relevant to the task at hand because it is topographically similar to the response being taught (Foxx & Jones, 1978; Jenkins & Larson, 1979; Ollendick, Matson, Esveldt-Dawson, & Shapiro, 1980). A logical and widely held assumption is that error correction improves performance by increasing the opportunities for a correct response to occur in the presence of the relevant stimulus accompanied by the trainer's cue (instruction, presentation of the sample, etc.). As noted previously, however, negative reinforcement also may influence behavior during error correction. To the extent that practice requires effort, delays reinforcement, or increases the duration of training, correct responding may increase as a function of avoidance.

Studies 1 and 2 were not designed to determine the extent to which discriminative control during training developed as a function of repeated practice of a relevant response versus repeated practice per se because all error-correction conditions involved practicing relevant responses. Moreover, the superiority of multiple over single (Study 1) and continuous over intermittent (Study 2) error correction also could be interpreted from either perspective in that the superior conditions could be described as producing better stimulus control through either more practice or a stronger negative reinforcement contingency. The purpose of Study 3, therefore, was to compare rates of sight-word acquisition when error correction consisted of the emission of topographically similar (relevant) versus dissimilar (irrelevant) responses. Rates of acquisition under the irrelevant condition might reveal the influence of negative reinforcement in the development of stimulus control over sight-word reading; any superiority under the relevant condition would reveal the additional influence of repeated practice of the correct response. Tess, Ariel, Ernie, Becky, Maisy, Hayley, Seth, Kara, and Mark participated in Study 3.

### *Method*

Baseline lengths were staggered to conform to a multiple baseline across subjects design. Following baseline, relevant and irrelevant error-correction conditions were alternated in a multielement design. MR error correction was used across both training conditions, and the procedures differed only with respect to the word that was repeated after an incorrect response was made.

*Baseline (no error correction).* A set of eight words was presented three times per session (24 word presentations). Word cards were presented one at a time, and participants were instructed to "read this word." Correct responses resulted in the delivery of a preferred edible item (Maisy, Seth, and Mark) or a ticket exchangeable for a preferred item or activity (Tess, Ariel, Ernie, Becky, Hayley, and Kara). Incorrect responses resulted in presentation of the next card. No additional instructions or prompts were given.

*MR (relevant).* Both the presentation of word cards and the consequences for correct responses were identical to the baseline condition. However, contingent on each incorrect response, the experimenter implemented MR error correction using the relevant word (i.e., the word that was misread). Following five correct or incorrect repetitions of the relevant word, the next card was presented.

*MR (irrelevant).* Procedures were identical to the relevant condition described above, except that a nontraining word was repeated during MR error correction. One word was selected for each participant by removing it from the participant's pool of unknown words (i.e., it never appeared later as a training word and was therefore irrelevant). When an incorrect response occurred, the experimenter said, "No, the word is [correct word]. Say [irrelevant word]." The card was displayed for the entire duration of the error-correction procedure.

After the participant correctly or incorrectly repeated the irrelevant word five times, the next word card was presented.

*Results and Discussion*

Figure 5 shows the cumulative number of words mastered during baseline and error-correction conditions. All participants' sight-word performances increased during error correction relative to baseline. Performance by 3 of the 9 participants (Tess, Ariel, and Ernie) during the relevant condition clearly exceeded that during the irrelevant condition. Performance by 5 other participants (Hayley, Becky, Kara, Maisy, and Seth) was similar during relevant and irrelevant conditions. Mark was the only participant who mastered more words during the irrelevant condition, and this discrepancy was large: He mastered 50% more words in the irrelevant condition.

Figure 6 shows each participant's mean number of correct words during the relevant and irrelevant conditions. More words were read correctly during relevant error-correction sessions by 7 of the 9 participants. The most pronounced difference was seen with Ariel, who made an average of 14.2 correct responses during relevant sessions compared to 9.2 correct responses during irrelevant sessions. Becky and Mark read more words correctly in the irrelevant condition.

Table 4 shows the retention test results, which were inconsistent both within and across participants. On short-term tests, 5 participants' retention was higher for words mastered in the relevant condition (Tess, Ariel, Becky, Maisy, and Hayley). The remaining 4 participants retained more words from the irrelevant condition, with the most dramatic difference observed with Mark (63% of relevant words retained, 81% of irrelevant words retained). On long-term tests, 6 participants (Becky, Maisy, Hayley, Seth, Kara, and Mark) retained more words from the relevant condition, and 2 (Ernie and Ariel) retained more words from the irrelevant condition.

The most consistent and interesting finding of this study was that error correction of any type enhanced sight-word acquisition (as in Study 2, Tess may have been an exception). In particular, results from the irrelevant condition suggest the influence of negative reinforcement during error correction because the only difference between the baseline and irrelevant MR condition was a contingency in which participants were required to repeat irrelevant words following errors.

We should note that the nontraining words presented during the irrelevant condition may not have been irrelevant in every possible respect. Because sight words were used as stimuli during irrelevant error correction, there may have been some shared properties (e.g., similar letters or phonemes) between the training word and the irrelevant word. Future research might examine the effects of irrelevant stimuli that have no common properties with the training stimuli (e.g., completion of math problems contingent on incorrect sight-word reading).

The comparison of relevant and irrelevant error correction produced mixed results. In most cases, sight-word acquisition was equivalent under relevant and irrelevant conditions; relatively large improvements were seen in the performance of only 3 participants during relevant error correction (Tess, Ariel, and Ernie). Nevertheless, the use of relevant stimuli during error correction might be recommended for several reasons. First, small improvements, if maintained over an extended period of time, may yield a significant advantage in terms of overall amount of learning. Second, although the relevant condition showed only modest superiority over the irrelevant condition in this study, it is quite possible that larger differences would be observed with other types of performance. Finally, it seems reasonable to take advantage of the fact that relevant error correction capitalizes on two learning processes.

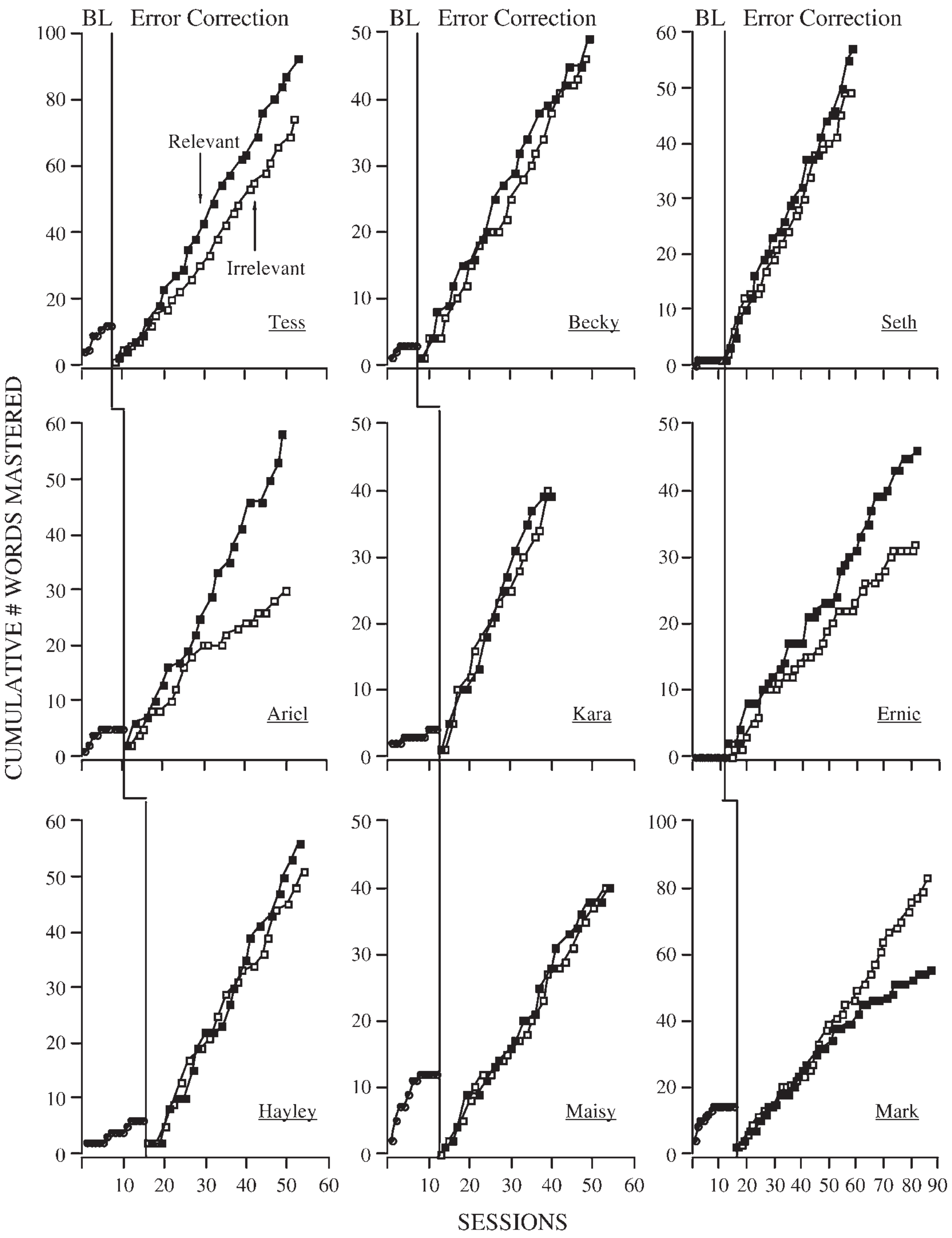


Figure 5. Cumulative number of words mastered during BL, relevant MR, and irrelevant MR.

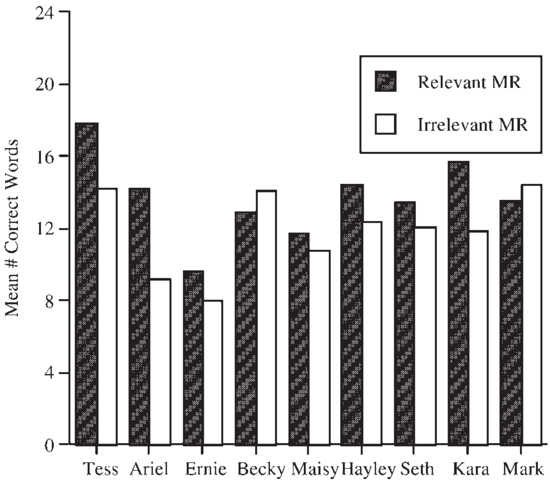


Figure 6. Mean number of correct words read per session during relevant MR and irrelevant MR.

Table 4
Percentage of Words Retained on Short- and Long-Term Tests during Relevant MR and Irrelevant MR

| | Short-term retention test | | Long-term retention test | |
|---|---|---|---|---|
| Participant | Relevant | Irrelevant | Relevant | Irrelevant |
| Tess | 94 | 89 | 94 | 94 |
| Ariel | 91 | 82 | 74 | 87 |
| Ernie | 72 | 75 | 57 | 66 |
| Becky | 92 | 85 | 82 | 79 |
| Maisy | 81 | 78 | 96 | 86 |
| Hayley | 82 | 72 | 75 | 57 |
| Seth | 79 | 84 | 75 | 62 |
| Kara | 89 | 91 | 91 | 81 |
| Mark | 63 | 81 | 59 | 58 |

## GENERAL DISCUSSION

Positive reinforcement is the most fundamental principle of behavior and is incorporated in virtually every learning-based approach to instruction. Furthermore, numerous studies have examined the influence of various characteristics of positive reinforcement, including type, amount, schedule, and delay. Another common component of instructional procedures is a contingency for the occurrence of an incorrect response, usually in the form of a correction trial; it is rare to find a study in which an incorrect response simply produced the next trial (i.e., extinction). Only a handful of studies, however, have systematically examined the effects of error correction, and the present research extended this literature with three comparative analyses.

The most striking finding, replicated across the three studies, was that every error-correction condition produced improvement in sight-word performance over a baseline condition consisting of positive reinforcement. Sight-word acquisition was low during baseline when preferred edible items or tickets were available contingent on correct responses. This result raises a question about the relative contributions of positive and negative reinforcement during behavioral acquisition, which cannot be answered with any of the present data because all error-correction conditions also included a positive reinforcement contingency. Although it is unlikely that anyone would recommend a teaching strategy devoid of rewards, a comparative analysis of the effects of positive and negative reinforcement seems warranted, if only to provide a more complete account of how learning occurs.

Studies 1 and 2 provided parametric analyses of the amount of error correction. Study 1 defined amount in a more traditional way: the number of response repetitions (one or five) on a given correction trial. Study 2 defined amount as the consistency with which error correction was applied (following every incorrect response or about every third incorrect response). Results of both studies indicated that more error correction produces better results. As is true of all parametric analyses, however, conclusions must be limited to the values assessed. We have already noted that performance differences observed in the SR and MR conditions of Study 1 may not have been obtained had the MR condition consisted of two rather than five repetitions; conversely, it is possible that error correction applied more consistently than in the intermittent (VR 3) condition of Study 2 would have shown results similar to those seen in the continuous condition. Our purpose in selecting particular values was to determine whether the parametric manipulation per se would influence behavior and not to determine the point at which differences emerge.

Another limitation of Studies 1 and 2 was the absence of data on responses that may interfere with acquisition. For example, Carey and Bucher (1983) compared the effects of short and long overcorrection on task performance and off-task behavior. No differences were observed between the conditions for these variables, but the authors noted an increase in problem behavior associated with the longer duration. We did not see such effects on the behavior of our participants, but these were only anecdotal observations. Finally, because the present studies used sight-word reading as a target behavior, the degree to which error-correction amount influences other skills is unknown. Thus, a number of questions remain about the effects of amount of error correction that could serve as the basis for future research.

Study 3 examined the extent to which avoidance (negative reinforcement) contributed to increases in correct responding during error correction by including a condition in which correction trials did not provide practice of the target (i.e., relevant) response. Results suggest that avoidance may, in fact, account for increases in correct responding. As noted previously, however, the potential benefits of relevant practice favor its inclusion as an error-correction component. Although this series of experiments was designed to evaluate the contribution of consequences for correct and incorrect responses on sight-word acquisition, the potential role of antecedents to the results obtained cannot be discounted. Prior to the introduction of any novel word into a training set, the experimenter displayed the word on an index card and read the word out loud to the participant. In the few cases in which sight words were acquired during baseline, increases in correct responses may have been due to the effects of positive reinforcement, antecedent modeling, or some combination of the two. Also, all error-correction procedures involved a vocal model of the correct response immediately following an error. That is, when an error occurred, the experimenter always recited the correct pronunciation of the word before proceeding with the specific correction procedure. Therefore, it is possible that sight-word acquisition during error correction was due in part to the presentation of a vocal model of the correct response and that any contribution of avoidance to acquisition may be a result of its interaction with these antecedent and consequent models (stimulus prompts).

Finally, results of the retention tests were not always consistent with performance observed during training sessions. Although 10 of the 12 comparisons in Study 1 (6 participants, two types of tests) favored MR over SR error correction, test results in Studies 2 and 3 did not correlate well with training performance, and several factors may have accounted for those discrepancies. First, mastery of more words in a given training condition resulted in more test words. Second, the mastery criterion may have not been sufficiently stringent to show differences in the retention tests; a more stringent mastery criterion may have resulted in higher overall retention percentages or distinct differences in retention between training conditions. In fact, it is almost certain that an across-session mastery criterion would have been reflected in data from short-term retention tests. Third, it is possible that improvements in performance during error correction were temporary.

In summary, results of the present studies provide additional information on some ways in which error-correction contingencies influence behavior. However, the data raise as many questions as they answer and illustrate the fact that error correction is a complex process that is perhaps subject to as many qualitative and quantitative variations as is positive reinforcement. Given the ubiquity of error correction as a basic instructional component and the relatively small amount of research that has been done on it, the present data might best be viewed as an example of how future research might be designed.

## REFERENCES

Axelrod, S., Kramer, A., Appleton, E., Rockett, T., & Hamlet, C. C. (1984). An analysis of the relevance of topographical similarity on positive practice of spelling errors. *Child and Family Behavior Therapy*, *6*, 19–31.

Barbetta, P. M., Heward, W. E., & Bradley, D. M. (1993). Relative effects of whole-word and phonetic-prompt error correction on the acquisition and maintenance of sight words by students with developmental disabilities. *Journal of Applied Behavior Analysis*, *26*, 99–110.

Birnie-Selwyn, B., & Guerin, B. (1997). Teaching children to spell: Decreasing consonant cluster errors by eliminating selective stimulus control. *Journal of Applied Behavior Analysis*, *30*, 69–91.

Carey, R. G., & Bucher, B. (1983). Positive practice overcorrection: The effects of duration of positive practice on acquisition and response reduction. *Journal of Applied Behavior Analysis*, *16*, 101–109.

Carr, E. G., & Kologinsky, E. (1983). Acquisition of sign language by autistic children II: Spontaneity and generalization effects. *Journal of Applied Behavior Analysis*, *16*, 297–314.

Cooper, J. O., Heron, T. E., & Heward, W. E. (1987). *Applied behavior analysis*. Columbus, OH: Merrill.

Cowley, B. J., Green, G., & Braunling-McMorrow, D. (1992). Using stimulus equivalence procedures to teach name-face matching to adults with brain injuries. *Journal of Applied Behavior Analysis*, *25*, 461–475.

Cuvo, A. J., Ashley, K. M., Marso, K. J., Zhang, B. L., & Fry, T. A. (1995). Effects of response practice on learning spelling and sight vocabulary. *Journal of Applied Behavior Analysis*, *28*, 155–173.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, *29*, 519–533.

Espin, C. A., & Deno, S. L. (1989). The effects of modeling and prompting feedback strategies on sight word reading of students labeled learning disabled. *Education and Treatment of Children*, *12*, 219–231.

Foxx, R. M., & Jones, J. R. (1978). A remediation program for increasing the spelling achievement of elementary and junior high school students. *Behavior Modification*, *2*, 211–231.

Horner, R. D., & Keilitz, I. (1975). Training mentally retarded adolescents to brush their teeth. *Journal of Applied Behavior Analysis*, *8*, 301–309.

Jenkins, J. R., & Larson, K. (1979). Evaluating error correction procedures for oral reading. *Journal of Special Education*, *13*, 145–156.

Koegel, R. L., & Egel, A. L. (1979). Motivating autistic children. *Journal of Abnormal Psychology*, *88*, 418–426.

Miltenberger, R. G. (2003). *Behavior modification: Principles and procedures* (3rd ed.). Belmont, CA: Wadsworth/Thomas Learning.

Neef, N. A., Iwata, B. A., & Page, T. J. (1980). The effects of interspersal training versus high density reinforcement on spelling acquisition and retention. *Journal of Applied Behavior Analysis*, *13*, 153–158.

Nutter, D., & Reid, D. H. (1978). Teaching retarded women a clothing selection skill using community norms. *Journal of Applied Behavior Analysis*, *11*, 475–487.

Ollendick, T. H., Matson, J. L., Esveldt-Dawson, K., & Shapiro, E. S. (1980). Increasing spelling achievement: An analysis of treatment procedures utilizing an alternating treatments design. *Journal of Applied Behavior Analysis*, *13*, 645–654.

Pany, D., McCoy, K. M., & Peters, E. E. (1981). Effects of corrective feedback on comprehension skills of remedial students. *Journal of Reading Behavior*, *13*, 131–143.

Remington, B., & Clarke, S. (1983). Acquisition of expressive signing by autistic children: An evaluation of simultaneous training and sign-alone training. *Journal of Applied Behavior Analysis*, *16*, 315–328.

Rodgers, T. A., & Iwata, B. A. (1991). An analysis of error-correction procedures during discrimination training. *Journal of Applied Behavior Analysis*, *24*, 775–781.

Rose, T. L., McEntire, E., & Dowdy, C. (1982). Effects of two error-correction procedures on oral reading. *Learning Disability Quarterly*, *5*, 100–105.

Rosenbaum, M. S., & Breiling, J. (1976). The development and functional control of reading-comprehension behavior. *Journal of Applied Behavior Analysis*, *9*, 323–333.

Singh, N. N., & Singh, J. (1986). Increasing oral reading proficiency. *Behavior Modification*, *10*, 115–130.

Wheeler, A. J., & Sulzer, B. (1970). Operant training and generalization of a verbal response form in a speech-deficient child. *Journal of Applied Behavior Analysis*, *3*, 139–147.

*Received July 28, 2004*
*Final acceptance June 30, 2005*
*Action Editor, Edward Daly*