Teaching Functional Sight Words to Students with Developmental Disabilities: A Comparison of Whole Word and Constructed-Response Instruction

Shari Daisy, M.Ed.

Simmons College

Send all correspondence to:
Shari Daisy, M.Ed.
16 Chestnut Circle
Merrimack, NH 03054
thedaisys@yahoo.com
(603) 429-2073
The purpose of this study was to compare the effectiveness of two instructional methods on the learning of functional sight words by school-age children. Three students with developmental disabilities participated in the study as part of a classroom-based implementation of whole word instruction using a popular commercially available curriculum (Edmark) and a teacher designed constructed-response curriculum. An alternating treatments design was used to assess the effectiveness of the two approaches. The dependent variables were number of sight words read/spelled correctly, latency of response, and duration of response. In addition, trials to criterion, retention of sight words, duration of lessons, and stimulus generalization were measured. Comprehension of sight words was directly trained. The constructed-response lessons also addressed issues of student choice, student preferences, and instruction through the use of social games.
THE PROBLEM

According to some, reading is unnatural in that although most children learn to speak through observational learning, they require instruction from others in order to learn how to read (Gough & Hillinger, 1980; Stanovich, 1994). Along with reading instruction, children also require an ability to analyze phonemes (often called phonological awareness) in order to become successful readers (Stanovich, 1994). Students with developmental disabilities may have particular difficulties with analyzing and processing words into their smaller units or parts (Stanovich, 1994). This may be why much of the sight word research geared towards instruction of students with developmental disabilities utilizes techniques involving whole word strategies.

However, this focus on the whole word does not address issues that students with developmental disabilities may have with stimulus overselectivity. According to Lovaas, Schreibman, Koegel, and Rehm (1971), stimulus overselectivity (also called stimulus blocking, restricted stimulus control, or selective stimulus control) occurs when “attention to one stimulus in a complex stimulus situation blocks or inhibits the attention to another cue also present” (p. 220). When students have difficulties with stimulus overselectivity, it may negatively impact their ability to generalize to reading words in context.
Problems with stimulus overselectivity while reading sight words are characterized by attending to certain topographical features of the word (e.g. the first letter, the length, the shape) and then making reading errors based on an inability to attend to all of the relevant stimuli (i.e. all the letters in the word and the order of the letters). An example of stimulus overselectivity while reading sight words would be the substitution of “fun” for “fat” or “chicken” for “children.”

Along with the problems of reading being unnatural and students with developmental disabilities having difficulties with phonological awareness and stimulus overselectivity, in a meta-analysis of 48 sight word research studies published between 1980 and 1997, only 46% measured comprehension and only 4 of the 48 studies measured stimulus and response generalization to functional use of the sight words taught (Browder & Xin, 1998). Although 87% of the studies trained to mastery, the essential problem for the student’s functional application is, “What good is it knowing how to verbalize a word if I don’t know what it means, and I don’t know how or when to use it?”

PREVIOUS RESEARCH

In the meta-analysis and review of sight word research conducted by Browder and Xin (1998), there were several...
variables identified related to effectiveness in teaching sight words. These included: postresponse interventions (prompts, feedback, students repeating words after correction), larger word sets, and opportunities for generalization or application. There was no significant difference regarding effectiveness for type of reinforcement used, time delay procedures, error level, or instructional format (Browder & Xin, 1998). As stated previously, a major weakness in the sight word literature is a lack of focus on comprehension, generalization, and functional use outcome measures.

A variety of whole word instructional strategies have also been identified in the research as effective in teaching sight words including: flashcards, picture-to-text or text-to-picture matching, text-to-text matching, repetition, errorless learning, interspersal training, prompting strategies including constant time delay, stimulus fading, and reinforcement (Belfiore, Skinner & Ferkis, 1995; Browder & Xin, 1998; Collins & Griffen, 1996; Conley, Derby, Roberts-Gwinn, Weber, & McLaughlin, 2004; Cuvo & Klatt, 1992; Fossett & Mirenda, 2006; Kirby, Holborn & Bushby, 1981; Neef, Iwata & Page, 1977; Winterling, 1990).

Out of the literature on stimulus equivalence, research in constructed-response has also shown promising results in teaching students to read and spell sight words (de Rose & de Souza, 1996; Dube, McDonald, McIlvane, & Mackay, 1991; Lee-
Vieira, Mayer, & Cameron, 2006). In a comprehensive article published in 1992, Stromer, Mackay and Stoddard reviewed a variety of classroom applications of stimulus equivalence techniques. Included was a thorough description of constructed-response, a technique that focuses on visual discrimination, sequencing, and direct manipulation of the letters in the word. During a constructed-response lesson, the student uses letters (typically letter tiles) to construct words by matching-to-sample with systematic fading of the model (Stromer et al., 1992). The authors hypothesized that since constructed-response involves the student attending to individual letters in a word, it might also address problems related to stimulus overselectivity (Stromer et al., 1992).

THE SOLUTION

To date, despite the large research base indicating a variety of effective techniques in teaching sight words, there has not been a study directly comparing the whole word approach to the constructed-response approach. Due to the issues raised previously in regards to difficulties students with developmental disabilities may have with stimulus overselectivity, it is crucial that if one instructional approach is more effective than another it be identified. Also, any research conducted on functional sight word instruction must address concerns with functional use outcome measures.
The purpose of this study was to:

1. Attempt to identify any significant differences in the learning outcomes of functional sight words when taught using a popular commercially available whole word approach (Edmark) compared to instruction using a constructed-response approach.

2. Assess learning of sight words using a variety of outcome measures: number of sight words read/spelled correctly, latency of response, duration of response, trials to criterion, retention, and stimulus generalization to functional text (e.g. menus).

3. Directly include instruction in comprehension by using the actual sight word items during all lessons (whole word and constructed-response).

4. Compare the efficiency of instruction by keeping durational time measures of all lessons taught.

5. Include components of social validity and the identification of behavioral cusps.

6. Address issues of student choice, student preferences, and instruction through the use of social games (during constructed-response sessions).

METHOD

PARTICIPANTS AND SETTING
Three middle school students with developmental disabilities participated in this study: Kristy a 15-year-old female, Tom a 13-year-old male, and Felicia a 12-year-old non-verbal female. All three students had received instruction in sight words using the Edmark reading curriculum for several years and had made gains in their sight word recognition base. However, their special education teacher was concerned with difficulties they were having with stimulus generalization and stimulus overselectivity. All aspects of the study were carried out in a separate special education classroom in the students’ public middle school.

SOCIAL VALIDITY

This study addressed two of the three levels of social validation identified in Wolf’s article on social validity: selection of goals valued by the society and selection of procedures that are socially acceptable (1978). In preparing for the study, parents of the students were surveyed regarding their views on the importance of their child learning how to read a menu, learning through the use of games, and being exposed to the actual foods/drinks represented by the sight words.

All of the parents stated that it was important for their child to learn how to play games and supported the use of games in instructional sessions. Two of the three parents stated
learning how to read a menu was an important skill that would improve their child’s autonomy and independence. Although one parent stated that, “any chance to read is a good thing” she also felt that her child might not need the skill since she would not be going to a restaurant alone. Although all of the parents stated that exposure to different foods/drinks would be helpful, one parent was doubtful that her child would expand her food repertoire, where another was hopeful that it may encourage her child to try different foods.

**BEHAVIORAL CUSPS**

Functional sight words that had the potential to serve as a behavioral cusp were selected for instruction in this study. Rosales-Ruiz and Baer (1997) defined a behavioral cusp as “a behavior change that has consequences for the organism beyond the change itself, some of which may be considered important” (p. 434). It was determined that by teaching the students words they may encounter while reading menus at restaurants and fast food establishments, additional opportunities could become available to them including communication with others in the community, independence, and choice and decision making.

**RESPONSE DEFINITIONS AND DATA COLLECTION**

Number of sight words read/spelled correctly, duration of response (for words spelled), and latency of response (for words read) were the primary dependent variables used to assess the
effectiveness of the interventions. A correctly read/spelled word was defined as a word that was pronounced or spelled exactly as written within 10 seconds of the presentation of the discriminative stimulus. For Kristy and Tom, the discriminative stimulus was the visual presentation of the sight word written in lower case text on an index card. For Felicia, the discriminative stimulus was the verbal cue “spell” along with the verbal pronunciation of the sight word. Latency of response was defined as the duration of time between the presentation of the discriminative stimulus and the beginning of the student’s verbal response. Duration of response was defined as the duration of time between the presentation of the discriminative stimulus and the completion of the student’s written response. Duration and latency times were rounded to the nearest second.

Secondary dependent variables included: trials to criterion, retention of sight words, duration of lessons, and stimulus generalization. Trials to criterion was defined as the number of times the student identified the sight word (e.g. verbalizing, pointing, circling, etc.) during the whole word instructional session or the number of times the student constructed the sight word during the constructed-response instructional session. Retention of sight words was defined as the percentage of previously taught sight words correctly read/spelled within 10 seconds of the presentation of the
discriminative stimulus. Duration of lessons was defined as the total elapsed time from the beginning of the lesson to the end of the lesson. Stimulus generalization was defined as number of previously taught sight words correctly read/pointed to within 10 seconds of the presentation of the discriminative stimulus (a menu paired with the verbal cue “read this” or “find the word _____”).

DESIGN

An alternating treatments design was used to compare the effectiveness of sight word learning using a whole word instructional approach and a constructed-response instructional approach (Barlow & Hayes, 1979). Sight words selected for the study came from The Edmark Functional Word Series: Fast Food/Restaurant Words. The words in the series were divided in half, with words 1-50 designated for whole word instruction as described in the Edmark Teacher’s Guide and words 51-100 designated for constructed-response instruction. Instructional sessions were randomly alternated between the two methods. Outcome sessions were conducted every 10 words taught, with 5 words taught using whole word instruction and 5 words taught using constructed-response instruction. Inter-observer reliability checks were conducted during 20% of the outcome sessions.

MATERIALS
Functional Sight Words

Materials used during the whole word instructional sessions were taken from *The Edmark Functional Word Series: Fast Food/Restaurant Words* and included a word recognition text, a vinyl display mask, and worksheets.

During constructed-response sessions, materials included: a small white board, 3 different colors of dry erase markers, a dry eraser, and games/letters used to construct the target word. Four games were available for students to choose from: unifix cubes labeled with individual lower case letters, Spill and Spell (a game with upper case letter dice), Nerdy Wordy (a game that included upper case letter dice and a grid to place the letters in), and Letter Perfect (a game that involved placing upper case letter tiles in a plastic grid).

An item representing the target word for the lesson (typically a food item) was used during all instructional sessions, both whole word and constructed-response.

PROCEDURES

BASELINE

Kristy and Tom were presented with the 100 sight words selected for the study one at a time. Words were written in lower case text on index cards and the students were directed to read them. Kristy correctly read 5 of the 100 words and Tom correctly read 8 of the 100 words. During the baseline measure,
Kristy made numerous errors based on the topography of the words.

Felicia was initially presented an array of 6 of the 100 sight words at a time and directed to point to a target word. With this initial assessment, Felicia was able to identify 39 of the 100 words, however it was hypothesized that she was discriminating using an exclusion technique (based on the first letter of the word). Therefore, a second test was administered where using pencil and paper Felicia was directed to spell the 39 words she had initially correctly identified. Using spelling as an outcome measure, Felicia correctly spelled 7 of the 100 words. It was determined that spelling would be an appropriate outcome measure for Felicia since it would also allow her to communicate with others in the community who were not familiar with her signs and gestures.

WHOLE WORD SIGHT WORD INSTRUCTION

Whole word instruction procedures were designed using the Edmark curriculum and teacher’s guide, with one exception: at the beginning of the lesson, students were primed with the item representing the target word for the lesson. During priming, students were told the word that they would be learning and informed that at the end of the lesson they could try the target food/drink. This step was intended to directly train for comprehension of the word. It was hypothesized that by directly
experiencing (through sight, smell, touch, and taste) the word that they were learning to read, students’ comprehension would be positively affected.

Using the vinyl display mask and the word recognition book from the Edmark curriculum materials, students were presented with isolated lines of text. Each line of text varied, but could include the target word, nonsense words, distracter words, and other fast food/restaurant words. Depending on the line of text (noted by a key for the instructor in the margin), students were directed to point to or verbalize specific words. Since Felicia is non-verbal, lessons were modified so that she could choose from two alternatives when the lesson required verbalization (e.g. “Is this bacon or burger?” combined with the instructor’s palms presented for Felicia to point to). Praise was provided on a CRF schedule. Criterion to move on to the next lesson was no more than four errors during the word recognition lesson.

Word recognition was followed by a series of worksheets reinforcing the target word and reviewing previously taught words. Worksheets varied, but typically involved students drawing a line from a picture of the word to text form of word, circling the word, and writing the word. Criterion was 80% accuracy. If students did not meet this criterion, errors were reviewed and a clean copy of the worksheet was given.
After all students had completed the lesson, the group prepared and sampled the item of the day.

CONSTRUCTED-RESPONSE SIGHT WORD INSTRUCTION

As in whole word instruction, at the beginning of the lesson students were primed with the item representing the target word for the lesson. Students were told the word that they would be learning and informed that at the end of the lesson they could try the target food/drink. As stated previously, this step was intended to directly train for comprehension of the word. Students then chose which game they would like to use for the lesson out of four possible options. Once the game was chosen, the instructor selected the letters that would be needed to construct the target word, along with three distracter letters. All other letters were removed from the learning area.

Using a white board, students were initially provided with a full written model of the target word and directed to “spell ____” by matching-to-sample. Students used their letter cubes/tiles/dice to construct the target word, racing to see who could finish the fastest. Praise for correct responses was provided on a CRF schedule. If students made an error, verbal and gestural prompts were provided to ensure a correct response on the second try. After construction of the word, students were cued to read the word and then spell it. For the non-
verbal student, the instructor read and spelled the word as she tracked with her finger. Criterion to move on the next step in the fading process was 100% accuracy.

Subsequent trials faded the model via backward chaining. (Backward chaining was chosen as the fading procedure after surveying regular education teachers and probing the method with both regular education students and the target students. It was determined that fading via backward chaining allowed for additional ordinal terminology instruction when cues such as “what’s first?” and “what’s next?” were used. Backward chaining also facilitated successful completion of the word, which may be reinforcing for the learner.) Once students could construct the word without a model, one distracter letter at a time was included in the letter pool. Criterion to move on to the next word in the program was 100% accuracy with three distracter letters present.

To address issues of stimulus overselectivity that arose during the study (it was noted that Kristy was selectively attending to the size and placement of the model), trials were randomly varied in regards to the written model that was provided. Each trial, the instructor used different colored markers, different size text, lower case letters, upper case letters, varied placement on the white board, and different letter shapes.
As in whole word instruction, after all students had completed the lesson, the group prepared and sampled the item of the day.

DAILY REVIEW

Each school day students reviewed all previously taught words, regardless of the method used to teach them. Kristy and Tom were presented with the words in a variety of formats (flash cards, white board, written on paper) and asked to read them out loud. Felicia was told the words and asked to spell them. If students were incorrect, they were told the correct answer and asked to repeat it.

OUTCOME TESTS

Outcome tests were conducted every 10 words taught, with 5 words taught using whole word instruction and 5 words taught using constructed-response instruction. Outcome tests were conducted utilizing the same procedures as in baseline, with Kristy and Tom verbally pronouncing the target words and Felicia spelling the target words. Stimulus generalization tests were conducted after 30 words were taught. During stimulus generalization tests, students were given menus obtained from restaurants in their community and asked to find or read specific target words in designated areas of the menu. The instructor used blocking techniques when the menus were visually overwhelming.
INTER-OBSERVER AGREEMENT AND TREATMENT INTEGRITY

Inter-observer agreement was measured during 50% of Felicia’s outcome tests and 33% of Tom and Kristy’s outcome tests. Inter-observer agreement was calculated for both response accuracy and time measures (duration or latency depending on the participant). Treatment integrity was measured during 13% of the instructional sessions. Checklists were developed for both instructional methods and a trained observer collected data regarding accuracy on the part of the instructor.

RESULTS

LEARNING OUTCOMES

Felicia’s learning acquisition outcome results are displayed in Figure 1. During the baseline measure, Felicia accurately spelled 1 word selected for the Edmark condition and 2 words selected for the constructed-response condition. During the first phase of the intervention, which included a daily review, Felicia accurately spelled more words taught in the constructed-response condition in the first two outcome tests. In the third outcome test she accurately spelled more words taught in the Edmark condition. Words 31-40 were taught without a daily review. This resulted in Felicia being unable to spell any of these words accurately in the final spelling outcome test. However, when given an outcome test based solely on identification, she was able to identify all of the words taught
in the Edmark condition and 4 out of 5 words taught in the constructed-response condition.

Although not indicated on Figure 1, comparable outcomes for Felicia’s duration of response were obtained by the completion of the study. During the last outcome test, Felicia spelled words taught in the Edmark condition with an average of .61 letters per second and words taught in the constructed-response condition with an average of .65 letters per second.

Figure 1.

During the three retention tests (see Figure 2), Felicia showed improvement in her ability to accurately spell words taught in the Edmark condition as demonstrated by jumps from 60%
accuracy to 90% accuracy to 100% accuracy. Retention of words taught in the constructed-response condition varied. Felicia spelled 100% of the constructed-response words correctly in the first retention test, 80% in the second test, and 87% in the final test.

Felicia’s stimulus generalization outcomes (identifying target words in local menus) are also presented in Figure 2. Felicia demonstrated the ability to generalize to a varied stimulus by accurately identifying 15 of 15 words taught in the Edmark condition and 14 out of 14 words taught in the constructed-response condition.

Figure 2.
Tom’s learning acquisition outcome results are displayed in Figure 3. During the baseline measure, Tom accurately read 2 words selected for the Edmark condition and 1 word selected for the constructed-response condition. During the first phase of the intervention, which alternated between Edmark and constructed-response instruction, Tom accurately read more words taught in the Edmark conditions than in the constructed-response conditions. The second phase of the intervention eliminated constructed-response instruction and substituted Edmark with an observing response (verbally spelling the target words while pointing to the letters). This resulted in Tom reading all of the words taught in the Edmark with observing response condition and 4 out of 5 of the words taught in the traditional Edmark condition.
During the first retention test (see Figure 4), Tom accurately read 100% of the words taught in both the Edmark and constructed-response conditions. His performance decreased during the second retention test, with 80% accuracy reading words taught in the Edmark condition and 90% accuracy reading words taught in the constructed-response condition. Retention tests were not administered for words taught in the final phase of the intervention (Edmark and Edmark with observing response).

Tom’s stimulus generalization outcomes (identifying target words in local menus) are also presented in Figure 4. Tom demonstrated the ability to generalize to a varied stimulus by accurately identifying 14 of 15 words taught in the Edmark
condition and 12 out of 14 words taught in the constructed-response condition.

Figure 4.

Kristy’s learning acquisition outcome results are displayed in Figure 5. During the baseline measure, Kristy accurately read none of the words selected for the Edmark condition and 1 word selected for the constructed-response condition. During the first phase of the intervention, which alternated between Edmark and constructed-response instruction, Kristy accurately read the same number of words in the first outcome test for words taught in the two different conditions. In the second outcome test, Kristy accurately read more words taught in the
Edmark condition than in the constructed-response condition. The second phase of the intervention eliminated constructed-response instruction and substituted Edmark with an observing response (verbally spelling the target words while pointing to the letters). This resulted in Kristy reading all of the words taught in the traditional Edmark condition and 2 out of 5 words taught in the Edmark with observing response condition.

During the first retention test (see Figure 6), Kristy accurately read 80% of the words taught in the Edmark condition and 100% of the words taught in the constructed-response condition. Her performance increased during the second
retention test for words taught in the Edmark condition and decreased for words taught in the constructed-response condition. Retention tests were not administered for words taught in the final phase of the intervention (Edmark and Edmark with observing response).

Kristy’s stimulus generalization outcomes are also presented in Figure 6. Despite significant blocking of extraneous text, Kristy was unable to identify any of the target words in menus from local restaurants. It was suspected that the combination of too much distracting text and small font size resulted in her being unable to generalize.

Figure 6.
INSTRUCTIONAL EFFICIENCY

Figures 7 through 12 show the results of the instructional efficiency measures for all three participants. For all participants, the Edmark condition involved more trials to reach criterion. However, constructed-response instructional lessons had significantly higher time durations. During the final phase of the interventions for Tom and Kristy, instructional efficiency was comparable between the traditional Edmark condition and the Edmark with observing response condition.

Figure 7.

![Instructional Efficiency in Trials to Criterion - Felicia](image-url)
Figure 8.

![Instructional Efficiency in Trials to Criterion - Tom](image)

Figure 9.

![Instructional Efficiency in Trials to Criterion - Kristy](image)
Figure 10.

**Instructional Efficiency in Lesson Duration - Felicia**

![Bar graph showing lesson duration for Felicia with different word counts and instructional methods.]

Figure 11.

**Instructional Efficiency in Lesson Duration - Tom**

![Bar graph showing lesson duration for Tom with different word counts and instructional methods.]

- EM and CR
- EM and EMO
INTER-OBSERVER AGREEMENT AND TREATMENT INTEGRITY

Inter-observer agreement for Felicia’s responses was recorded during 2 of her 4 outcome tests. For both tests there was 100% agreement on the accuracy of Felicia’s responses. There were slight differences in agreement on duration of response (no more than one second), which resulted in 77% and 93% agreement on duration.

Inter-observer agreement for Tom and Kristy’s responses was recorded during 1 of their 2 outcome tests. There was 95% agreement on the accuracy of Tom’s responses and 80% agreement on his latency of response. There was 100% agreement on the
accuracy of Kristy’s responses and 92% agreement on her latency of response. As in Felicia’s IOA, the differences in agreement during the time measures were never more than one second.

Out of 100 total lessons taught, 13 were observed to measure procedural integrity. Three different instructors participated in the Edmark condition. Instructor 1 was observed 4 times and her accuracy ranged from 93% to 100% with an average of 98%. Instructor 2 was observed 3 times and her accuracy ranged from 84% to 94% with an average of 90%. Instructor 3 was observed 2 times and her accuracy ranged from 82% to 92% with an average of 87%. Only one instructor participated in the constructed-response condition (the author) and she was observed 4 times. Her accuracy ranged from 96% to 100% with an average of 98%.

DISCUSSION

The purpose of this study was to compare the effectiveness of two instructional methods on the learning of functional sight words by school-age children. Results varied for each of the three participants. For Felicia, the component that impacted her learning the most was not the instructional method, but rather the daily review. For Tom, Edmark instruction was more effective than constructed-response; but Edmark with an observing response was even better. For Kristy, traditional Edmark instruction was the most effective method.
In terms of instructional efficiency, Edmark was clearly superior. On average, Edmark instructional lessons took half the time of constructed-response lessons despite requiring more trials to reach criterion. Culturally, Edmark was also the more accepted method by the teaching staff in the classroom.

A secondary purpose of this study was to directly train for comprehension and assess for stimulus generalization, areas that are currently lacking in published research. Although comprehension was not assessed in the current study, it was directly trained through the use of priming before each lesson. In addition, promising results were obtained for two of the participants in regards to their ability to generalize their learning to different stimuli (identification of target words in real menus). As a result of this study, instruction in restaurant/fast food words was terminated for Kristy. It was determined that due to her unsuccessful attempts at stimulus generalization, most likely due to the overwhelming textual features on menus, this was not an appropriate area to focus on. Instead, her functional sight word program began to focus on community signs (which typically are limited in their extraneous text).

As in the meta-analysis and review of sight word research conducted by Browder and Xin (1998), this study also found no significant differences in terms of instructional format as most
effective to teach sight words. Instead, students benefited
most from systematic instruction with opportunities for daily
review and practice. It is hoped that the results obtained in
this study will encourage teachers to not only choose an
efficient instructional method to teach sight words, but also
one that is maximally effective for each individual student,
with the understanding that each student may require a different
method.
References


