Word Prediction Programs with Phonetic Spelling Support: Performance Comparisons and Impact on Journal Writing for Students with Writing Difficulties

Anna S. Evmenova, Heidi J. Graff, Marci Kinas Jerome, and Michael M. Behrmann

George Mason University

This investigation examined the effects of currently available word prediction software programs that support phonetic/inventive spelling on the quality of journal writing by six students with severe writing and/or spelling difficulties in grades three through six during a month-long summer writing program. A changing conditions single-subject research design was used and replicated across the participants. Using a daily writing prompt, students alternated between Co:Writer, WordQ, and WriteAssist word prediction programs. The results provided evidence for the effectiveness of various word prediction programs over word processing, and demonstrated improvements in spelling accuracy across conditions. Relative gains in the total number of words and composition rate were modest for the majority of the participants and should be interpreted with caution due to several methodological issues. The social validity interviews revealed that all students enjoyed the word prediction programs and found them beneficial. Study limitations and recommendations for future research are discussed.

In recent years there has been an increasing interest in technology applications for students with high-incidence disabilities, including students who struggle with writing. Several applications have been discussed in the literature (Edyburn, 2005; Higgins & Raskind, 2004; Lewis, 1998, MacArthur, Ferretti, Okolo, & Cavalier, 2001). Writing is a complex skill and students may experience difficulties with a variety of aspects, including mechanics and written content expression; however, computer-related technologies can enable students to bypass their deficits and support them through all stages of the writing process (Behrmann & Jerome, 2002; Lewis, 1998; Williams, 2002; Zhang, 2000). To date, accumulated research has shown the effectiveness of word processors for ease of text alteration and manipulation (Lewis, Ashton, Haapa, Kieley, & Fielden, 1998; MacArthur & Graham, 1987; MacArthur & Schwartz, 1990; Zhang, 2000), as well as spell checkers and other aids for ease of editing (Ashton, 1999; MacArthur, Graham, Haynes, & DeLaPaz, 1996; McNaughton, Hughes, & Ofeshe, 1997; Montgomery, Karlan, & Coutinho, 2001). Text-to-speech software programs that allow users to hear written products have been found to be effective for accuracy monitoring (MacArthur, 1998, 1999; Raskind & Higgins, 1995). Outlining and brainstorming programs that allow visual representation of ideas have been shown to support users in planning and organizing their writing (Anderson-Inman & Ditson, 1999; Blair, Ormsbee, & Brandes, 2002; Sturm & Rankin-Erickson, 2002). In addition, the use of speech recognition programs, which transform spoken words into text, have resulted in longer, more complex and accurate writing passages by students with learning disabilities and/or writing difficulties (DeLaPaz, 1999; Higgins & Raskind, 2000; MacArthur & Cavalier, 2004; Quinlan, 2004; Raskind & Higgins, 1999).

Despite the available research, the area of assistive technology for students with mild disabilities is still not fully developed (Edyburn, 2005). While some evidence of the effectiveness of the technology exists, few studies have investigated its use by students with high-incidence disabilities in authentic learning environments, such as schools or other educational programs (Blackhurst, 2005; Edyburn, 2001). This is true for the utilization of word prediction software for students with learning disabilities. Word prediction programs were originally created for users with physical disabilities, and were designed to increase typing rate and decrease spelling errors (Tumlin & Heller, 2004). While the reduction in keystrokes addresses the needs of students with physical disabilities, additional features of word prediction software can be helpful in compensating for word recall, spelling, and handwriting difficulties faced by students with learning disabilities (Lewis, 1998). With a word prediction program, the user is offered a list of word choices as she/he begins to type the word. Word suggestions appear before or after the letter of the word is entered. The word “prediction” feature allows for the words to be generated based...
on the lexical and grammatical context (MacArthur, 1999; Sitko, Laine, & Sitko, 2005). As with other assistive technology solutions, word prediction programs may yield writing products of higher quality if the features of the program are coordinated with the user’s abilities and needs (Ashton, 2005; Sitko et al., 2005). However, despite being a promising application, research on the use of word prediction software for students with writing difficulties is limited (MacArthur et al., 2001).

**Previous Research**

A majority of word prediction studies were conducted in the 1990s. Several recent studies have provided limited support for the positive effects of word prediction programs on the writing readability/legibility and spelling of students with learning disabilities (Handley-More, 2003; MacArthur, 1998, 1999; Williams, 2002). In his original study, MacArthur (1998) compared the effects of speech synthesis and word prediction to word processing on the composition rate, spelling, and legibility of the written dialogue journal entries. For four out of five students, the features offered by My Words word prediction program resulted in improved legibility and spelling in dialogue journal entries. The composition rate was not affected by the treatment. Later, MacArthur (1999) extended his original study by using a more sophisticated Co:Writer word prediction program. In this latter study, students alternated between handwriting, word processing, and word prediction during free journal writing activities. Use of the word prediction program yielded improvement in the proportion of correctly spelled words for one out of three students and decreased the composition rate for two students. Technology had no effect on the proportion of legible words during free journal writing.

Furthermore, in the same article MacArthur (1999) reported the results of the second study conducted with the same participants and in the same conditions but engaging in a more demanding writing task. Students wrote from dictation, thus increasing vocabulary demands. The results of the second study demonstrated improvements across legible and correctly spelled words as well as the reduced composition rate for two out of three students. Similarly, Williams (2002) and Handley-More (2003) reported relative improvements in the number and variety of words, as well as in the percentage of legible and correctly spelled words. Thus, while existing research is somewhat ambiguous about the value of word prediction on composition rate and legibility of writing by students with learning disabilities, this technology has been found to improve students’ spelling accuracy (Mirenda, Turoldo, & McAvoy, 2006).

MacArthur (1998, 1999) identified the lack of prediction from phonetic spelling as a contributing factor to the mixed results of his studies. Users had to know the exact beginning letter of the word in order for it to appear in the prediction list. Students with severe spelling problems did not benefit from the word prediction programs since they often did not know the correct initial letters of words (MacArthur, 1998). Word prediction technology has continued to progress so that software programs now recognize phonetic spelling in addition to conventional spelling. Relevant predictions are now possible based on the phonetic spelling of a word. Consequently, students with learning disabilities may experience more success using current, more advanced word prediction programs than was found in previous studies (MacArthur, 1998, etc.).

**Research Questions**

The purpose of this study was to examine the effects of the new generation of word prediction software programs that support phonetic spelling on the length, spelling accuracy, and composition rate of journal writing by students with severe writing and/or spelling difficulties. This study compared three different word prediction programs to include student preferences. The study was designed to replicate and extend the work of previous researchers (MacArthur, 1998, 1999). It addresses the following research questions:

1. Are there differences in length, spelling accuracy, or rate of journal writing for students with writing difficulties across different word prediction programs?
2. What were students’ reactions to using different word prediction programs?

A changing conditions single-subject research design was used and replicated across six students.

**METHOD**

**Participants**

Participants were students in grades 3 through 6 attending a 4-week long technology-based summer writing camp located at a major northeastern university. The camp uses technology and innovative computer software programs to enhance the writing skills of students identified by their parents and teachers as having difficulties with the writing process. The majority of the participants had been identified by their schools as having learning disabilities. Six of the 15 campers were identified as candidates for word prediction intervention based on informal writing assessments as well as writing samples collected prior to the study. The writing assessments and samples provided data related to students’ phonetic spelling, vocabulary, word recall, keyboarding skills, and reading ability to differentiate between words on a prediction list.

It was critical to consider the effects of word processing and word prediction in light of students’ keyboarding skills and familiarity with the Word system. For this reason, parents were asked to report students’ level of familiarity with the computer and various software programs in the Parent Questionnaire that was distributed prior to the beginning of camp. According to this information, Students B, M, D, and J were sufficiently familiar, while Students C and R were comfortable enough with word processing to perform basic writing activities. Throughout the camp, due to its technological focus, all campers received more training and extended practice in basic word processing features.

Students keyboarding skills were determined via the Type-to-Learn 3 program. This software reports students’ typed words per minute (wpm), accuracy percentage score, and the
but putting them down on paper was a challenge for him. Inative ideas that came from his well-developed imagination, to do and for others to read” the finished product. He had cre-
mother, the writing process was “physically difficult for him
education resource and inclusion settings. According to his
range for written expression.
and organizing his writing. It was observed that he fixated on
practiced their typing skills with TypetoLearn
Student R was considered a candidate for word prediction use to
provide him with additional word choice support.
Student J was a rising sixth-grade boy from a medium-
income family. He received special education services in inclus-
Student M was a rising fifth-grade boy from a high-income
family. Student M was recommended for the summer camp
because of his reluctance toward writing. He was also tested
and found to have traits for autism spectrum disorder with
a very high-functioning level. Student M was reported to
be easily frustrated with the writing process when he could
not think of the correct spelling. He found it hard to focus,
organize, and convey his ideas in a cohesive document. He
was considered a study participant for his phonetic spelling
and his deficiencies of finding the right word to convey his
ideas.
At the age of 9, student D was a rising fourth-grade boy
from a middle-class family. He received up to 8 hours a week
extra help in the general education classroom for difficulties
with writing. His mother shared with the researchers that
Student D “did not like to write and did so as little as possi-
ble.” His biggest problem was attending to a task; therefore,
he was considered a candidate for word prediction use to
provide him with additional word choice support.
Student J was a rising sixth-grade boy from a medium-
income family. He received special education services in inclus-
ive settings. Student J was reported to have severe
spelling difficulty; therefore, he was considered a good can-
didate for the use of word prediction programs. His mother
requested all writing assignments to be completed on a com-
puter with additional help for “spelling issues.” Student J’s
performance on TOWL (Test of Written Language) could not
be scored because his paragraph could not be deciphered. His
reading and comprehension were within the average range,
slightly above grade level. His writing was reported to be
well below grade level (early fourth-grade level).
Student R came from a middle-income family. This rising
fourth-grade student received special education services in a
resource classroom. After careful observations and consul-
tations with his family, school teachers, and camp instruc-
tor, Student R was determined as a study candidate to help
him overcome his hesitancy to write. In addition, utilizing
computer programs for writing addressed his fine motor and
handwriting challenges.

Materials

All Conditions

In all conditions students were asked to write daily for 20
minutes in response to the journal entry prompt. The purpose
of such journal writing is to provide students with more writ-
ing opportunities and daily practice (Reagan, Mastropieri, &
Scruggs, 2005). It usually is free of any kind of evaluation
(Williams, 2002). Personal narrative prompts were randomly
assigned to students from a pool of 30 preselected prompts.
They were designed to be interesting and unbiased based

<table>
<thead>
<tr>
<th>Participants</th>
<th>Gender</th>
<th>Age</th>
<th>Ethnicity</th>
<th>Disability</th>
<th>IEP Case</th>
<th>Writing Accuracy*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student B</td>
<td>M</td>
<td>12</td>
<td>WH</td>
<td>SLD</td>
<td>WE, M</td>
<td>57%</td>
</tr>
<tr>
<td>Student C</td>
<td>M</td>
<td>9</td>
<td>WH</td>
<td>SLD</td>
<td>WE, RC</td>
<td>54%</td>
</tr>
<tr>
<td>Student M</td>
<td>M</td>
<td>10</td>
<td>AS</td>
<td>AST</td>
<td>N/A</td>
<td>71%</td>
</tr>
<tr>
<td>Student D</td>
<td>M</td>
<td>9</td>
<td>WH</td>
<td>ADD</td>
<td>N/A</td>
<td>67%</td>
</tr>
<tr>
<td>Student J</td>
<td>M</td>
<td>11</td>
<td>WH</td>
<td>SLD</td>
<td>WE</td>
<td>41%</td>
</tr>
<tr>
<td>Student R</td>
<td>M</td>
<td>9</td>
<td>WH</td>
<td>SLD</td>
<td>WE, M</td>
<td>49%</td>
</tr>
</tbody>
</table>

Note: Age = at the beginning of the study; WH = White; AS = Asian; SLD = specific learning disabilities; ADD = attention deficit disorder; AST = autism spectrum tendencies; WE = writing expression; RC = reading comprehension; M = math; * – writing spelling accuracy level is based on informal writing assessment prior to the beginning of the camp.
on gender, ethnicity, and socioeconomic status (e.g., “What is your favorite part of the day?,” “What is something that makes you feel happy or sad?,” etc.)

Baseline Condition

In the baseline condition, students used Microsoft Word for journal writing. Students were not able to use spell checkers or grammar checkers during writing. All camp participants were characterized by poor handwriting skills that severely interfered with their ability to write. Besides that, handwriting on paper was avoided as a baseline measure to control for the novelty of technology-medium integration effect. Clark (1983) noticed an increased level of effort and focus in research subjects as they were introduced to novel media. This increased attention seemed to diminish as they became more familiar with the technology medium. Thus, it was critical to compare students’ writing performance using word prediction programs to a similar technology-based instrument such as word processing.

Treatment Conditions

Students used three word prediction programs: Co:Writer, WordQ, and WriteAssist. With these programs, text is entered either in a separate program application or in Microsoft Word. As each letter is typed the list of predicted words appears in a small window located by the cursor. If the intended word appears in the list, a student can select the word by clicking on it or typing the number for that word. The selected word is automatically inserted into the sentence. If the intended word does not appear in the predicted list, a student continues to type. All three programs provide speech feedback, so students have the option to hear predicted words before selecting one of them. These programs also have an option for the teacher to decide how many words will appear in the prediction list (usually between one and nine). While the number of predicted words is usually based on an individual student’s needs, in this research project the number of predicted words was limited to five for all programs. All three programs have spell checkers built into them. However, for the sake of this study, the spell checker option was disabled in all word prediction programs as well as in the word processing. While these three programs are similar in their features, they are slightly different in the level of sophistication and the size and diversity of the dictionary.

Co:Writer

Co:Writer SOLO Edition, at the time of the study, was the most recent version of the program developed by Don Johnston Inc. For the purposes of this study, the word prediction feature was used when students were offered word choices even before any letters were typed. With the FlexSpell feature Co:Writer provides prediction choices based on phonetic, inventive spelling typical of students with learning disabilities and writing difficulties. This feature considers the most common letter-patterns students may try in order to spell a word (e.g., phonetic substitutions, common letter confusions, letter reversals, letter omissions, letter additions, letter doubling, and singling, etc.). Thus, the word prediction options do not depend on the correct first letters. In addition, teachers can customize the spelling support for their students allowing for flexible spelling to be taken into consideration always, after three letters, or encouraging exact spelling only. In this study, the FlexSpell feature was set to always provide word prediction based on phonetic spelling. Co:Writer also utilizes Linguistic Word Prediction intelligence. This function ensures that the word prediction list offers grammatically correct word choices based on the context and previously used words (e.g., predicted verbs match plural nouns, etc.). Capitalization and punctuation are also reinforced within this program (e.g., predicted words are automatically capitalized). Thus, its prediction choices facilitate improvements in correct subject–verb agreement, proper spelling, capitalization, and appropriate pronoun and article use. Co:Writer SOLO Edition provides multiple ways to customize the program to students’ individual needs. For the purposes of this research study, all students used the intermediate dictionary template, which consists of 12,000 words covering many school subjects such as history, geography, and science. The intermediate dictionary supported the age and writing level of the campers without overwhelming them. In addition to the speech feedback, Co:Writer offers functions such as eWordBank and Topical Dictionary. These features support student’s writing on different topics and in different genres, predicting the most appropriate words for the selected topic and/or genre. However, eWordBank and Topical Dictionary features were not used during journal writing in this research study.

WordQ

WordQ by Quillsoft is a word prediction tool used with a standard word processor. This program appears as a simply four-button overlay toolbar for a standard word processor. This program bases word predictions on students’ creative writing and context, offers examples for commonly confused words, and provides speech feedback. It offers prediction options even before the starting letter is entered based on what is most likely to be the next word. This program automatically capitalizes the next word after the period. Unlike other two-word prediction programs, by default WordQ does not provide grammar-based predictions, so correct syntax fully depends on students. It is possible to manipulate the program to see different word endings; however, this additional option was not used in this study. Several user vocabularies are available (e.g., U.S., U.K., Canadian, blank, starter [5,000 words], intermediate [10,500], and advanced [15,000]). U.S. intermediate vocabulary was used in this study to match students’ grade level. It included 10,500 words commonly used by second- to eighth-grade writers. The program learns the user’s writing style and improves the prediction options. WordQ also offers an opportunity to expand the vocabulary by adding words one-by-one or creating new topic dictionaries. As with Co:Writer, this option was not used for the purposes of this study.
WriteAssist

*WriteAssist* by Second Guess software is a dyslexia-oriented word predictor. Program features include context-dependent prediction, which ensures that a student is offered suggestions even without typing anything. The program makes a prediction of the possible next word based on grammatical patterns and the context. It also provides optional automatic capitalization of the first word in a sentence and proper nouns. *WriteAssist* includes a 30,000-word vocabulary pretrained from more than 30 million words of English text. Users are also supported with the speech-feedback feature. Just like the other two programs, *WriteAssist* constantly learns new words adding them to the prediction list based on the user’s writing style and word choices. However, in comparison with other two programs, *WriteAssist* does not provide as many additional features and can be considered to be the simplest program of the three.

**Experimental Design**

Changing conditions single-subject design was employed to investigate the effects of various word prediction programs on improving students’ writing (Alberto & Troutman, 2006). Prior to treatment, students’ baseline level of writing was collected for a minimum of three data points across the first week of camp. Following the stabilization of baseline, the first treatment was introduced. Students were randomly assigned to one of the three logically preestablished orders of word prediction programs so that each student had an opportunity to try three programs by the end of the study. During each following week students wrote using different programs, alternating the order across the participants. The random assignment to a program for each particular week was used to control the influence of the increasing mastery and familiarity with word prediction skills (Figures 1–3). Changing treatments sequentially allowed the examination of various programs before finding one that was the most beneficial for each particular student. While changing conditions single-subject design is adequate for deciding which intervention works the best for each individual student, it cannot be used for establishing a functional relationship between the baseline and each of the conditions without the return to baseline (Kennedy, 2005). Since there was no return to baseline in this study, the results should be interpreted with great caution, especially in those conditions when the improvements are inconclusive. In turn, the replication across students was used to support the possibility of stronger functional relationships between various word prediction programs and improvements in students’ writing as well as control for confounding variables such as novelty of treatment (Clark, 1983; Weller, 1996) and acquisition of necessary skills (Alberto & Troutman, 2006).

**Dependent Measures**

The dependent variables examined included: total number of words, the proportion of correctly spelled words, and composition rate by written words per minute.

**Total Number of Words**

Total number of words was calculated for each of the students’ writing samples. Proper nouns were counted as words while numerals (unless they were spelled out) were not counted as words. When scoring the total number of words, raters counted both correctly and incorrectly spelled words. The differences in length of writing were compared between word processing and word prediction, as well as among the three different programs.

**Correctly Spelled Words**

The proportion of correctly spelled words was calculated by dividing the number of correctly spelled words by the total number of words. The raters marked each word that was spelled correctly (e.g., one word misspelled three times was counted as three mistakes). Words written in the incorrect tense or form were counted as spelling errors. Homonyms were considered misspelled. Words that were spelled correctly but were inappropriately used within the context of the sentence (e.g., “peanut butter sandbox” vs. “peanut butter sandwich”) were also counted as spelling errors. Capitalization and punctuation were not considered to be errors, as these skills were not taught directly.

**Composition Rate**

Composition rate of typed words per minute was calculated by dividing the total number of words in students’ writing by the total minutes of composition time. Throughout the study, the beginning and ending writing times were recorded for each student.

**Procedures**

Once students’ and parents’ permissions were obtained, participants were asked to engage in journal writing for approximately 20 minutes at the beginning of each camp session over 4 weeks. The majority of the sessions lasted for 15–20 minutes across the participants and conditions. However, a few students tended to write for 10 minutes or less (Student M in sessions 2, 3, 6, 8, 12, 14; Student J in session 15; and Student R in session 5), while some others wrote for 25 minutes or more (Student B in sessions 1, 4; Student C in sessions 4, 5, 7, 10, 17; and Student J in session 10). Students received personal narrative writing prompts daily. Later, during the day, students were engaged in other writing activities, including brainstorming, drafting, revising, editing, and production. Thus, the purpose of the journal writing activity was to provide another opportunity to write without spending time on editing. However, if students wrote fewer than three sentences, they were asked to elaborate by providing more details on the topic. Students were also encouraged to spell a word or choose it from the prediction list independently, without any help.
Baseline Condition

During the first week of camp, students wrote their journal entries using a word processor without access to spell check. Depending on students’ typing and computer skills, they received instruction in typing and using the word processor if needed. Such instruction included one-on-one training from the researcher as well as practice with the Type2Learn 3 software program, which provided interactive lessons to teach typing and improve speed. One of the prerequisites for participation in this study was that students had to pass the Young/Easily frustrated level associated with the typing speed goal of six words per minute. Handwriting was substituted with word processing for the baseline condition, thus fewer changes occurred when students were introduced to treatment conditions (Kennedy, 2005).

Treatment Condition

Prior to beginning the treatment, students received instruction on how to use each of the word prediction programs. The training session started with a short PowerPoint presentation with the basic information about word prediction. Then, all main features (e.g., checking the list, having each word from the list pronounced in order to make a word choice, speech feedback during writing, etc.) were demonstrated and practiced.
Participants were randomly assigned to the order in which they used programs weekly (see Figures 1–3). Each week the researcher modeled the particular program assigned to students for that time. Students learned how to start the program, enable the word prediction feature, utilize the speech feedback feature, and where to look for the prediction list. The journal writing activity was simulated for students in order to address specific functions of the particular program. Each session followed a preestablished script to ensure consistency in how every participant was introduced to each program. After the training, students had time to practice using the software.

**Interrater Reliability and Fidelity of Treatment**

Interrater reliability was determined using printed writing samples as permanent products. Random writing samples (33 percent) were distributed to an independent observer to ensure scoring reliability. The interrater agreement on each of three dependent variables was calculated using the total agreement method, the smaller total recorded by each observer was divided by the larger total recorded by each observer was divided by the larger total and multiplied by 100 percent (Kennedy, 2005). Interrater agreement averaged at 99 percent (ranging from 89 percent to 100 percent) for the total number of words,
Fidelity of treatment data were collected during 33 percent of all sessions. The randomly observed activities were compared to the checklist of expected researcher's actions. Those actions included introducing the topic choices; introducing word prediction programs according to the script; asking students to elaborate if they wrote less than three sentences; and encouraging and allowing students to spell words or select a word from the prediction list without any additional help. Thus, it was ensured that the researcher and students were doing what they were supposed to do. Such observations occurred both during trainings and journal writing. The number of correct behaviors performed by the researcher and the participants was divided by the number of planned behaviors and multiplied by 100 percent. Sessions during which technical difficulties occurred were excluded from the data analysis. Mean procedural reliability was 100 percent across all training and journal writing sessions.

Social Validity

The social validity of each intervention, each individual word prediction program, was examined through student interviews conducted during and at the end of the study.
Social validity as defined by Kennedy (2005) is “the estimation of the importance, effectiveness, appropriateness, and/or satisfaction various people experience in relation to a particular intervention” (p. 219). As with any assistive technology device or program, it is very important to ensure a person’s willingness to use it. Student preferences for a program and its technological features often play an important role in its effectiveness. A large percentage of assistive technologies are abandoned because they do not meet a person’s needs and expectations (Scherer, 2005). It is critical to seek students’ input when technology is selected (Parette, Wojcik, Peterson-Karlan, & Hourcade, 2005).

RESULTS

Student performance on the total number of words, proportion of words spelled correctly, and composition rate was determined and analyzed through the visual analysis of mean lines in the data and the percentage of nonoverlapping data (PND) scores (Scruggs, Mastropieri, & Casto, 1987). While the changing conditions single-subject design used in this study is apt for the comparisons between the programs described in detail later, the following results on word processing versus word prediction are only preliminary and should be considered with caution, especially in those cases when the changes are less noticeable and/or inconclusive. However, some obvious differences between the baseline and various word prediction programs were noted as follows.

Word Processing Versus Word Prediction

Figures 1–3 illustrate the relative effectiveness of various word prediction programs over word processing, revealing evident improvements in the proportion of correctly spelled words by all the participants with each of the programs. While gains in the total number of words and composition rate were smaller, most students performed better on these two measures with at least one of the three word prediction programs as compared to word processing alone. Thus, the visual analysis demonstrates that participants increased in the total number of words averaging in gains from 21 words (SD = 6) to 36 words (SD = 6.9) across all students and all word prediction programs. In particular, the amount of writing doubled for Student B with all three programs, for Students C, J, and R with WordQ, for Student D with WriteAssist software. The gains were quite modest and Student M only slightly increased the amount of writing with all three programs. The overall performance of the participants on each measure can be seen from the descriptive statistics in Table 2.

The positive effects of word prediction on students’ writing were more obvious from the analysis of the proportion of correctly spelled words. On average students increased their spelling accuracy from 58 percent (SD = 3.5) to 96 percent (SD = 4) across all the programs. In addition, with the composition rate, students wrote relatively faster with all the programs improving from 1.1 wpm (SD = 0.3) to 2.4 wpm (SD = 0.7) when using WordQ; to 1.9 wpm (SD = 0.4) when using CoWriter; and to 2.1 wpm (SD = 0.6) when using WriteAssist. However, the individual performance of each participant with different word prediction programs on the composition rate measure should be noted, as changes were more evident for some students than others. The overall gains in students’ writing from word processing to word prediction were corroborated by the PND scores across students and across all three programs with the average 80 percent improvement in the total number of words, 100 percent in the proportion of correctly spelled words, and 84 percent in the composition rate. However, the replication of this study using a different single-subject research design is needed to suggest the effectiveness of word prediction over word processing.

Total Number of Words With Different Word Prediction Programs

As can be seen from Figure 1, almost all students demonstrated a relative increase in the total number of words from the baseline to the word prediction condition. However, the contrast was more evident for students using some programs than others. According to the visual analysis and mean values, Students C, J, and R demonstrated clear improvements in the total number of words when using the WordQ word prediction program (improvements from an average of 33

<table>
<thead>
<tr>
<th>Students</th>
<th>BL</th>
<th>WQ</th>
<th>CW</th>
<th>WA</th>
<th>BL</th>
<th>WQ</th>
<th>CW</th>
<th>WA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student B</td>
<td>22.7 (8.5)</td>
<td>58 (13.8)</td>
<td>57.2 (6.4)</td>
<td>48.8 (5.1)</td>
<td>61 (4)</td>
<td>91.8 (4.4)</td>
<td>92 (7)</td>
<td>93.2 (3.1)</td>
</tr>
<tr>
<td>Student C</td>
<td>33.3 (15.9)</td>
<td>63.8 (6.6)</td>
<td>39 (5.4)</td>
<td>37.8 (4.7)</td>
<td>61 (7)</td>
<td>98.4 (1.5)</td>
<td>96.6 (2.6)</td>
<td>97.4 (2.4)</td>
</tr>
<tr>
<td>Student M</td>
<td>16.3 (3.1)</td>
<td>26.4 (6.3)</td>
<td>20 (5.5)</td>
<td>24 (5.3)</td>
<td>75.3 (1.2)</td>
<td>98.2 (2.5)</td>
<td>95.6 (7.4)</td>
<td>98 (2.7)</td>
</tr>
<tr>
<td>Student D</td>
<td>14 (6.1)</td>
<td>23.8 (5.8)</td>
<td>20.6 (4.4)</td>
<td>31.2 (8.6)</td>
<td>59 (2)</td>
<td>99.2 (1.8)</td>
<td>97.8 (3.2)</td>
<td>100 (0)</td>
</tr>
<tr>
<td>Student J</td>
<td>27.7 (12)</td>
<td>59 (18.7)</td>
<td>34.2 (3.8)</td>
<td>33 (10)</td>
<td>47 (5.6)</td>
<td>92 (5.7)</td>
<td>91.4 (5.3)</td>
<td>92.2 (5)</td>
</tr>
<tr>
<td>Student R</td>
<td>12.7 (1.5)</td>
<td>29.8 (4.3)</td>
<td>22 (5.7)</td>
<td>16.4 (3.9)</td>
<td>44.7 (1.5)</td>
<td>98 (1.9)</td>
<td>92.4 (10.8)</td>
<td>95 (4.7)</td>
</tr>
<tr>
<td>Total:</td>
<td>21.1 (6)</td>
<td>43.5 (9.3)</td>
<td>32.2 (6.2)</td>
<td>31.9 (6.3)</td>
<td>58 (3.5)</td>
<td>96.3 (3)</td>
<td>94.3 (5.4)</td>
<td>96 (3.7)</td>
</tr>
</tbody>
</table>

**Note:** BL = baseline, word processing condition; WQ = WordQ; CW = CoWriter; WA = WriteAssist; M (SD).

TABLE 2

Descriptive Statistics (M(SD)) for Each Dependent Variable across Conditions and Participants
(SD = 15.9) words to 64 (SD = 6.6) for Student C; from 28 (SD = 1.2) to 59 (SD = 18.7) for Student J; and from 13 (SD = 1.5) to 30 words (SD = 4.3) for Student R. At the same time, Student C and J’s performance when using CoWriter and WriteAssist, as well as Student R’s writing when using WriteAssist demonstrated little if any improvement over word processing (e.g., Student C did not write more with either CoWriter or WriteAssist).

While Students B and M showed similar progress with all three programs, they appeared to produce a slightly larger number of words, particularly with the WordQ program (M = 58; SD = 13.8 and M = 26.4; SD = 6.3, respectively). However, while Student B’s amount of writing increased sufficiently with all three programs, changes in Student M’s writing were only minor. Student D’s graphed data indicated effectiveness of the WriteAssist software only (M = 31.2; SD = 8.6).

The percentage of nonoverlapping data corroborates the conclusions drawn from the visual analysis. Thus, Students C, J, and R scored 100 percent PND when using WordQ demonstrating 100 percent improvement in the total number of words as compared to the word processor. The PND scores for Student B were 100 percent for each of the programs. Student D’s PND scores reached 100 percent PND with the WriteAssist program only. Given changes in the total number of words with the other word prediction software for the aforementioned participants and Student M’s modest writing gains with all three programs, his results were considered unreliable and were not interpreted.

**Proportion of Words Spelled Correctly With Different Word Prediction Programs**

Visual inspection of the data points for the proportion of words spelled correctly across the three different programs suggested the following results. All students considerably improved their spelling when using word prediction software compared to word processing regardless of the program. Students showed similar gains in the mean values with each of the three word prediction programs. Student B improved from an average of 61 percent (SD = 4) spelling accuracy to 92.3 percent (SD = 4.8) with word prediction. Student C’s spelling accuracy increased from 61 percent (SD = 7) in the baseline to 97.5 percent (SD = 2.2) using software. Student M went from 75.3 percent (SD = 1.2) in the baseline to 97.3 percent (SD = 4.2) with the programs. Student D increased from an average of 59 percent (SD = 2) in the baseline to 99 percent (SD = 2.5) in the treatment condition. Student J improved from 47 percent (SD = 1.5) using word processing in the baseline condition to 91.9 percent (SD = 5.3) using word prediction. Student’s R gains in the spelling accuracy measure went from 44.7 percent (SD = 1.5) in the baseline to the averaged 95.1 percent (SD = 5.8) with the treatment. Individual gains in spelling accuracy with each of the three word prediction programs can be found in Table 2. All students accelerated immediately demonstrating 100 percent PND across all word prediction conditions as compared to word processing.

**Composition Rate With Different Word Prediction Programs**

Students increased their composition rates to differing degrees and with different results depending upon the word prediction program. A few consistent and evident through the visual analysis changes are as follows. Graphic representation of data points in composition rate for Student B revealed an improvement from 1.1 (SD = 0.4) wpm in the baseline to 3.6 (SD = 0.4) wpm with Co:Writer, to 2.9 wpm (SD = 0.7) with WordQ, and to 2.9 wpm (SD = 0.4) with WriteAssist. One hundred percent PND score supported this conclusion. Graphic representations for Students M and D suggested a slightly higher composition rate with the WriteAssist program. The mean lines increased from 1.04 wpm (SD = 0.1) to 2.2 wpm (SD = 0.4) for Student M and from 0.7 wpm (SD = 0.2) to 2.1 wpm (SD = 0.6) for Student D. Visual analysis for Students C, J, and R indicated a higher composition rate when writing with WordQ. Student C went from 1.5 wpm (SD = 0.7) to 2.5 (SD = 0.8); Student J from 1.4 wpm (SD = 0.1) to 3.6 wpm (SD = 0.9); Student R from 0.8 wpm (SD = 0.1) to 1.6 wpm (SD = 0.2). PND scores (80–100 percent) corroborated the effectiveness of WordQ over other programs for these students. Additionally, for student R, CoWriter was associated with an increase from 0.8 wpm (SD = 0.1) to 1.4 wpm (SD = 0.9) and 100 percent PND score in composition rate.

The performance of Students M and D on the composition rate measure while using WordQ and CoWriter programs as well as the rate of Student J with CoWriter and WriteAssist and Student R with WriteAssist is characterized by minor and inconsistent changes (see Figure 3). Given such methodological issues as limited typing skills of the study participants discussed later, these results should only be viewed as inconclusive. In addition, no changes in composition rate were observed for Student C while using WriteAssist and CoWriter programs.

**Social Validity**

Overall, all the students enjoyed the word prediction programs and found them beneficial. They indicated that writing was much easier when they used word prediction. Student M noted that he did not have to write the whole word and the program would finish it for him. Another student mentioned that word prediction made him type words faster. One more example of the advantage of word prediction as reported by Student B was that it “helped find words and see if they were correct or not in order to use them.” In addition, Student C reported that word prediction made him think faster.

In regard to which program students found the most helpful and enjoyable, four out of six students preferred WordQ to other programs, while the other two students liked WriteAssist and Co:Writer the most. Students who chose WordQ referred to it as the fastest and having “better words.” One student mentioned that “it was like telepathic,” because “the words came up” as you were just thinking about them. Another student enjoyed that the program read the sentences “exactly as you read them.” Student C refused to use other
programs because they did not have as many voices as WordQ did. As for those students who did not prefer WordQ, they noted that constant speech-feedback was annoying and that the window although smaller “moved around to places where they did not want it.” Both of those problems were eliminated within program options as students expressed their opinions. The main problem with the WriteAssist program as reported by students was the big window that “did not move and covered the words.” WriteAssist was found to have less word choices. In addition, the vocabulary was not appropriate for students since it offered “bad words” as reported by students and teachers. Co:Writer had more “technical glitches” within the software in comparison to other programs. Since the time of the study several glitches were addressed and solutions offered on the manufacturer’s website.

The teachers supported students’ opinions and preferred WordQ to the other word prediction programs. They commented it was the easiest to use while offering a large choice of features. WordQ’s four-button toolbar was “very straight forward” and simple to handle. Teachers noticed vocabulary issues with WriteAssist as it offered curse words. One teacher described it as the “most primitive of all.” In addition, one teacher with vision impairments noted that the WriteAssist program “would be hard to use with a screen enlarging program,” as the prediction window stays in the locked position on the screen as opposed to WordQ and the word window of Co:Writer. In turn, the Co:Writer program had “cleaner language” and had more features for students. However, it was also more difficult to use for younger students in this study and presented more technological difficulties.

**DISCUSSION AND PRACTICAL IMPLICATIONS**

The primary goal of this study was to explore the effects of various word prediction programs on students’ journal writing as compared to word processing. Consistent with previous research (Handley-More, 2003; MacArthur, 1998, 1999; Williams, 2002) the results of this study demonstrated the relative effectiveness of word prediction on various aspects of the writing process for some students with writing difficulties as compared to word processing alone. Word prediction regardless of the software was effective in improving written spelling accuracy as measured by the proportion of words spelled correctly for all the participants. Writing performance increased in the total number of words for one student with each of the three programs and for the other five students with at least one of the programs. Composition rate noticeably improved for one student with each program, for four students with at least one of the programs, and did not visibly change for one of the participants.

The latter two measures need to be discussed considering one major limitation of this study. The proficiency in typing skills is essential to students’ success in word processing. Otherwise, the computer can be a frustrating and inefficient tool to use (MacArthur & Graham; 1987). While all parents reported students’ familiarity with the computer and word processor, the participants in this study were characterized by limited typing skills ranging from 6 to 10 words per minute. The existing research reports that students, particularly boys in third to sixth grades are able to handwrite 43–78 characters per minute, which is approximately eight to 15 words per minute (Chwirka, Gurney, & Burtner, 2002; Graham, 1999; Graham, Berninger, Weintraub, & Shafer, 1998; Rosenblum, Weiss, & Parush, 2003). If the participants in this study were able to type 6–10 words per minute, they were approaching the production consistent with handwriting speed at their grade level. However, the students’ typing skills were determined based on the tests within the typing software, which overestimates the speed. Thus, possible weak keyboarding skills should be recognized when discussing the total number of words and composition rate measures.

Overall, the improvements in the amount and the writing speed were less obvious than gains in spelling. Due to limited typing skills and the nature of the changing conditions single-subject research design, these results should be interpreted with great caution (Kennedy, 2005). The absence of a return to baseline prior to starting each new program alters the conclusions about functional relationships between word processing and each word prediction condition. However, this study presents unique information comparing different word prediction programs. External validity of this study is enhanced through replication across different participants and random assignment of students to a different order of program implementation (Horner et al., 2005).

Overall, regardless of the order in which it was introduced, three and four students performed relatively better with WordQ on the total number of words and composition rate, respectively. Student D demonstrated a greater number of words and faster composition rate with WriteAssist. Student B’s performance improved similarly with all three word prediction programs, so the recommendation of one of them depended solely on his preferences. As far as the proportion of words spelled correctly, written spelling accuracy improved similarly for all of the students while using each of the three word prediction programs, thus demonstrating overall effectiveness of word prediction over word processing.

Social validity of the goals, procedures, and effects were examined through student and teacher interviews. Both students and teachers enjoyed using word prediction programs and found them helpful for the writing process. All students benefited from word prediction features that supported their writing difficulties. For example, a student with fine motor/handwriting difficulties mentioned the ability to type faster with word prediction. Furthermore, a student who experienced difficulties with putting his ideas on paper reported that word prediction made him “think faster.”

A majority of students and teachers preferred WordQ to the other two word prediction programs due to its extended and appropriate vocabulary and ease of use, despite the fact that WordQ’s intermediate vocabulary level had the least number of words as compared to other two programs. It was interesting that several students referred to WordQ’s “telepathic” abilities, recognizing its contextual prediction. Although Co:Writer and WriteAssist also have similar features, WordQ predictions were identified as more precise by both students and teachers. One student who preferred WriteAssist performed better with that program on all dependent variables. The student who chose Co:Writer demonstrated equally significant improvements with all the programs, so following his preferences, Co:Writer was recommended as a word prediction program for him.
Based on the results of this study, Co:Writer was reported as being more difficult to use due to the extensive number of features. This suggests that Co:Writer may be a better choice for older students who could use and benefit from those features. WriteAssist was the only program that offered users up to 30 choices in the prediction list. This program was also reported as having a more grown-up-oriented vocabulary, suggesting its greater value for adults, including those with physical disabilities. It is also important to note that a student with ADD considered the prediction window moving with a cursor annoying and distracting. So, it may be important to consider turning off such features in WordQ and Co:Writer with “word window” prediction. In addition, for students with autism spectrum tendencies a program with a larger selection of voices (WordQ) could be preferable.

When interpreting the results of this study, it is important to remember that the presented word prediction programs offer a number of extended features that can further enhance students’ writing (e.g., eWordBank in Co:Writer and Topical Dictionaries in Co:Writer and WordQ). These features were not used in this study. This was the first attempt to investigate the rate of production and student preferences of three word prediction programs while controlling for any additional, extraneous characteristics. Thus, it is impossible to fully compare these three programs based on this study alone. More comprehensive evaluation of extended features conducted during a longer period of time is necessary in order to more thoroughly investigate each program’s effectiveness.

**Limitations and Future Research**

Unfortunately, interpretation of the findings from this study is clouded by a number of methodological issues. First of all, the keyboarding skills of participating students were quite weak at the beginning of the study. Participants were not retested on these skills at the end of the study, so it is impossible to say how their typing speed prior to and following the study affected the results. Second, the short duration of the writing camp guided the number of journal writing sessions randomly assigned for each word prediction program. Thus, time also influenced the choice of the research design that excluded return to baseline, preventing the establishment of a stronger functional relationship between different word prediction programs and improvements in students’ journal writing. Furthermore, the length of the summer camp hindered the researchers’ ability to test maintenance and level of continuous improvement of students writing with a particular word prediction program. Finally, the research setting at CompuWrite summer camp was different from a general education classroom where a majority of participants received special education services. Thus, generalization of the writing improvement with word prediction program was not assessed in a realistic school environment.

At the same time, this pilot study provided several areas for future research. First, the study could be replicated allowing longer periods of time for each word prediction program and employing different single-subject research designs (e.g., multiple baseline across participants) to establish stronger functional relationships between the writing improvement and each word prediction program. Second, it is important to explore the true impact of word prediction on students’ writing. Third, it would be interesting to examine the effectiveness of different word prediction programs on other more meaningful writing activities that require editing in more natural school settings. Such research may also focus on more extensive features of word prediction programs including prediction based on topic and genre. Last, a majority of research studies examining the effectiveness of word prediction are single-subject research studies (Siko et al., 2005). This may be explained by the specificity of word prediction programs that makes them appropriate for students with very specific abilities and needs. Such studies, including the present one, discourage interpretation and generalization of findings to a larger population. Group design experimental studies with a large number of participants would provide generalizable information on word prediction effectiveness.

**REFERENCES**


### About the Authors

**Anna S. Evmenova** is an assistant professor of special education at George Mason University, Virginia. Her current research interests focus on assistive and instructional technology tools for providing academic content-based instruction to students with disabilities.

**Heidi J. Graff, Ph.D.** is the director of the Mason LIFE program at George Mason University. Her interests are post-secondary education for students with intellectual and developmental disabilities as well as technology tools for improved instruction in the general education setting.

**Marc K. Jerome** is an assistant professor, Department of Special Education, George Mason University. She earned her Ph.D. in Special Education in 2008 at George Mason University. Her current research interests include evaluating the effectiveness of assistive technology tools on student learning and distance education in special education teacher preparation.

**Michael M. Behrmann** is a professor of special education and the director of the Helen A. Kellar Institute for Human DisAbilities at George Mason University. He is a leader/innovator in special education and technology and his research interests include assistive technology, instructional design, universal design for learning and distance education.