Technology Applications for Children with Early Literacy Difficulties: A Framework for Review and Evaluation of Popular Beginning Reading Computer Software

M. J. Bishop
Asst. Prof., Educational Technology
Lehigh University
mj.bishop@lehigh.edu

Lana L. Edwards
Asst. Prof., Special Education
Lehigh University
lle3@lehigh.edu

With the growing prominence of computer-based instructional technologies, it seems appropriate, if not necessary to use empirically supported criteria, to evaluate software applications for children with early literacy difficulties. The purpose of this study was to develop a research-based evaluation framework to evaluate beginning reading software programs that might be used by teachers and parents for children who struggle with beginning reading skills. As part of the instrument's development, the evaluation framework was used to review identified popular beginning reading software. A total of twenty-one software programs were reviewed and evaluated according to criteria addressing the software's interface, instructional design, and beginning reading content. While results should be interpreted cautiously due to continued technical validation and development of the software review instrument, initial findings suggest that the sample of programs generally did not meet evaluation standards for interface, instructional design, and content. An inverse relationship was also observed on scores for interface design and beginning reading content. In other words, programs that were rated more highly in interface design tended to be rated lower on beginning reading content. Next steps for research and implications for practice based on the results from the review are also discussed.

Many children in America struggle with foundational literacy skills. According to the most recent National Assessment of Educational Progress, 40% of U.S. fourth-grade students read below a "basic level" and have "little or no mastery of the knowledge of skills necessary to perform work at each grade level" (National Center for Educational Statistics, 2001). One of the most compelling findings from recent reading research is that children who get off to a poor start in reading rarely catch up. (Torgesen, 1998). Felton and Pepper (1995), for example, reported longitudinal results indicating that most students with poor word recognition skills in the third grade failed to make significant gains in their basic reading skills by the end of the eighth grade without intensive instructional intervention.

Converging evidence suggests that students with reading difficulties require instruction that is quantitatively and qualitatively superior to instruction received by students who are already facile with the reading code (Adams, 1990; Foorman, Francis, Fletcher, Schatschnedier, & Mehta, 1998). Carefully-designed instruction is required to help students at-risk of reading failure attain grade-level expectations and have access to the general education curriculum. Foorman and Torgesen (2001) concluded that reading instruction for children at-risk of reading difficulties should include phonemically explicit interventions, intensity, and strategic cognitive supports. In other words, instruction must include activities that focus on a child's ability to hear sounds in spoken words (such as rhyming or phonemic segmentation); should be delivered with some focused intensity in the context of small group, individual, or peer-assisted interventions; and should progress systematically from teacher-directed to student-directed learning.

Computer-assisted instruction includes many of the critical instructional features found to be effective for students with reading difficulties (Hall, Hughes, & Filbert, 2000; Rieth & Semmel, 1979; Woodward, 1984). For example, computer-assisted instruction is often explicit with immediate feedback, extensive skills review, and consistent error correction procedures. Recent research reviews suggest generally positive outcomes related to acquisition of reading skills when computer-assisted instruction is incorporated in interventions for students with reading difficulties (Blok, Oostdam, Otter, & Overmaat, 2002; Hall, Hughes, & Filbert , 2000; MacArthur, Ferretti, Okolo, & Cavalier, 2001).

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

Despite the promise of computer-assisted instruction's effectiveness, there are a relatively limited number of high-quality, empirical studies investigating the instructional impact on student beginning reading skills (Blok, Oostdam, Otter, & Overmaat, 2002; National Institute of Child Health and Human Development, 2000). In fact, Blok, Oostdam, Otter, and Overmaat's (2002) recent meta-analysis of computer-assisted instruction's use in beginning reading interventions suggested caution in the interpretation of effectiveness due to the poor quality of many studies. In addition, it seems most software used in previous intervention studies is not readily accessible to teachers and parents--often initially developed by researchers, somewhat difficult to acquire due to its unpublished status, as well as technologically outdated (Coley, Crader, & Engel, 1997).

Finally, the relationship between interface design and the instructional presentation of content remains unexplored in research on computer-assisted instruction in beginning reading. Most studies focus on the instructional design and quality of beginning reading content without also exploring the program's interface design (see, for example, Blok, Oostdam, Otter, & Overmaat, 2002). Even if a program addresses beginning reading skills, how the software's interface responds to and supports content learning is also critical (for example, how the computer responds when children respond incorrectly, the attractiveness or clutter of a screen display, how easy it is to repeat and review directions for an activity, and the like) (Alessi & Trollip, 1991; Brown, 1988; Cates & Bruce, 2000; Fleming, 1987; Hannafin & Rieber, 1989)

While many questions concerning the design and use of software for early readers remain unanswered, interest in computer-assisted instruction for beginning reading skill development is high. In a recent survey of over 1,000 special educators, Burton-Radzely (1998) found that 85% of the sample used computer-based technology in literacy instruction and reported that 91% expected their use of technology to increase in the future. With the growing prominence of computer-based technology in the schools, it seems appropriate, if not necessary, to use an empirically supported criteria to evaluate beginning reading software programs. Evaluating the instructional design and development of beginning reading software is necessary to document the efficacy of popular programs for at-risk learners (Lyon, 1997).

Agencies such as the National Institute of Child Health and Human Development (NICHD) have called for a formal procedure to assess the current status of scientific research-based knowledge relevant to the effectiveness of various approaches to teach children to read (Lyon, 2001). As a result of the NICHD initiative, the need to evaluate the efficacy of commercial reading curricula is nationally recognized. Despite the importance of conducting content analyses on reading curricula and software, there is a paucity of empirical research investigating the quality and adequacy of commonly used, commercially available beginning reading programs for children at-risk of reading disabilities and early reading delays (Smith et al., 2001). To address the need for systematic examinations of basal reading program content, Smith et al. developed a research-based evaluative framework for beginning reading curricula. The evaluation tool emphasized well-established principles of instruction (Kame'enui & Carnine, 1998) and content criteria related to phonological awareness (Smith et al., 1998), a critical skill in beginning reading instruction. While the Smith et al. evaluation framework was developed to review commercial beginning reading programs, no such review criteria exists for the purpose of evaluating the content of beginning reading software.

Therefore, the purpose of our study was to develop a research-based evaluation criteria, such as the criteria used by Smith et al. (2001), for the evaluation of beginning reading software programs. Specifically, we extended the Smith et al. criteria and developed an evaluation framework that addressed a software program's interface, instructional design, and beginning reading content. As part of the instrument's development we also reviewed an identified sample of popular beginning reading computer software programs. Overall, therefore,

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

there were three intended outcomes from this pilot study: (a) to test our evaluation instrument by using it to evaluate current commercially available beginning reading software programs with respect to interface, instructional design, and content requirements for children with and at-risk of reading disabilities, (b) to identify the extent to which the reviewed current beginning reading software programs meet empirical standards for interface, instructional design, and content, and (c) to identify any highly rated software programs for possible use in future intervention research.

Methodology

The research team consisted of a special education professor who is an experienced developer and evaluator of beginning reading instruction and an educational technology professor with many years of experience in instructional design and software development.

To achieve our purpose, we employed a modified content analysis approach. Content analysis is a technique widely used in education for systematically describing the objective content of instructional materials and the opportunities for learning these materials offer (Berelson, 1952, Tamir, 1985). Generally, content analyses are aimed at answering questions directly relating to the material analyzed to obtain the information needed to identify or solve some educational problem (Borg & Gall, 1983). For example, this technique has been used in previous studies to evaluate quality of learner activities in educational software (Shiratuddin & Landoni, 2002); assess strategies suggested in the literature for enhancing reading through technology (Lang, McCarty, Norman, & Upchurch, 1999); analyze the suitability of middle-school geography textbooks for students with learning problems (Jitendra, Nolet, Gomez, Xin, 1999); and investigate the kinds of technological resources available to support language learning (Garrett, 1991).

The steps in a content analysis typically involve: deciding on the sample to be used; deriving a framework for coding observations; training the coders to consistently apply the coding scheme; and analyzing the data (Ary, Jacobs, & Razavieh, 2002). In cases where a pre-established coding framework does not exist, researchers must derive their own criteria and define categories for content indicators that will measure the variables of interest. When these categories involve inference or evaluation on the part of the rater, as is the case in this study, these coding frameworks can be more difficult to develop (Borg & Gall, 1983).

Deriving the Initial Framework

In preparation for developing our software evaluation framework, we taught each other our individual areas of expertise through three "information" sessions and drafted the portions of the instrument that were exclusively within our individual knowledge domains. We then pulled these pieces together with a jointly-derived section on instructional design to create an initial conceptual framework that arises from our knowledge of accepted interface design criteria (see Cates & Bishop, 2003; Laurel, 1986, 1990; Norman, 1988; Rieber, 1996), instructional design criteria from the empirical literature (see Carnine, Silbert, & Kame'enui, 1997; Gagné & Driscoll, 1988), as well as strategies found within well-established and widely-accepted explicit reading instructional programs (see Good, Simmons, & Smith, 1998; Smith et al., 2001).

The items were divided into four broad categories: background information, interface design, instructional design, and reading content. The background information section gathered data on the program's cost, package contents, supplementary materials, target audience, necessary prerequisite skills, and stated goals and objectives. Interface design criteria included categories that addressed software affordances (things within the program driving the learner to action), quality of interactions, media aesthetics, and motivational aspects. The

instructional design criteria addressed whether the software included explicit instructional sequences, engaged learners in well-designed instructional activities, and monitored learner progress throughout the program. The beginning reading instructional criteria addressed the extent to which the software covered phonological awareness skills (at the sound, sentence, word, syllable, onset-rime, and phoneme levels) and alphabetic understanding (letter-sound correspondences, blending, segmenting, and word reading).

With the broad categories established and the preliminary indicators in place, we struggled a good deal over an appropriate rating scale to use. The trouble was in how to deal with two different types of indicator –those that appeared to involve a binary choice (the presence or absence of direct instruction, for example) versus those that appeared to involve a decision over quality (how well that instruction was designed). In the end, we settled on using a 5-point Likert scale (strongly disagree = 1, disagree = 2, neutral = 3, agree = 4, strongly agree = 5) for rating individual items in a 2-tiered approach (see Table 1). Once the rater had answered the "subset" items (indented in Table 1), those scores would be used as part of the rater's decision about how to rate the broader item (emboldened in Table 1). In cases where the element was absent from the software, the rater was to check "Strongly Disagree" for the larger item, skipping over the sub-items and moving on to the next set.

The program includes well-designed, explicit instructional sequences (or processes) for skills.	
	The program gains learners' attention at the beginning of each instructional cycle.
	The program reminds learners of prerequisite knowledge at the beginning of each instructional cycle.
	The program informs learners of objectives at the beginning of each instructional cycle.
	The program uses a "model-lead-test" framework.
	The program offers multiple examples of a target skill.
	The program supplies adequate opportunities for learners to practice new skills.
	The program presents the same requirements in embedded activities as presented in instructional sequence.
	The program requires learners to demonstrate mastery of newly introduced skills before moving on to new skills.
	The program supplies larger conceptual anchors for retention and retrieval (transfer of knowledge).
	The program moves systematically and cyclically through the phonological awareness skill hierarchy (sounds, sentences, words, syllables, onset-rimes, phonemes) and/or word-reading skill hierarchy (letter-sound, blending, segmenting, word reading).

Table 1. Excerpt of instrument showing relationship between broad item and supporting subitems.

Sample

Before determining which software programs to examine, we operationally defined "instructional software for beginning reading." As a foundation, Martindale, Qian, and Cates' (2001) requirement for classifying computer-based materials was used. According to Martindale, Qian, and Cates,' computer-based materials can be classified as *instructional* if the materials included: (1) intended learning goals, (2) instructional strategies (specifications for selecting and sequencing events and activities within a learning environment), and (3) learning materials and activities that may also include learner assessment and/or feedback. Based on this criteria, "storybook" and "tool-based" computer programs were excluded. "Storybook" computer programs

focus exclusively on reading stories to children as they interact with screen elements (see for example *Grandma and Me*, n.d.) and "tool-based" computer programs only supply learners with the ability to manipulate words and spellings on the computer (see for example *Inspiration*, 2003; and *KidPix*, 2003).

To find commercially available software programs that are commonly used in homes and schools, software for the evaluation was identified through the Association for Supervision and Curriculum Development's "Only the Best" list of top rated software (ASCD, 1999, 2000). This list was selected because ASCD is a well-known organization with a national reputation and because several nationally recognized evaluation sources had also participated in rating and ranking the reviewed software (such as the American Library Association, the California Instructional Technology Clearinghouse, and the Children's Software Revue). The "early childhood" and "language arts" categories were scanned for those titles aimed at pre-kindergarten to 3rd grade (ages 4-8) that appeared from their descriptions to have a focus on teaching beginning reading skills (some mention of phonological awareness, alphabetic understanding, or spelling). A total of 39 titles matching the criteria were identified.

Attempts were made to acquire all 39 titles. Eight of the 39 titles were unobtainable for various reasons. During our review, an additional 10 of the 39 titles were found not to meet our selection criteria (5 ended up being storybooks and 5 were more "general studies" with very little focus on reading skills). A total of 21 software titles were included in the final analysis. Appendix 1 lists the 39 titles from the original sample, divided into those included in the sample and those eliminated (along with our reason for excluding them from the review).

Instrument Evolution

We piloted the initial instrument by together examining and rating four software titles not included in our sample of 21. During these sessions several adjustments were made to reduce the instrument's scope and make it more feasible to use.

We then each randomly chose three titles from our sample and examined them independently using our instrument. Individually, we carefully examined each software package and recorded our impressions in the appropriate cells of the instrument. After we completed those evaluations, we "swapped" titles so that we had both independently examined the initial set of six. We met to review each title, check for agreement between us, and negotiate differences until we had 100% consensus on every item for each of the six packages. We then went on to analyze another set of four or five titles using our new, shared understandings, and repeated the process until we had finished reviewing all 21 titles.

As we encountered unanticipated positive and negative software traits, we discovered that items in the instrument needed to be added or altered in certain ways (reworded, made more inclusive, required more specific examples). In addition, we found that there were redundancies that called for eliminating certain items from the initial instrument. There were also items and categories of items that were reordered to make the instrument easier to use. In this way, the instrument and its underlying operational definitions were revised several times throughout this process. Consequently, we decided to re-review all of the titles to test our understandings of specific items and assure that we had formulated exhaustive and mutually exclusive coding categories. Additionally, to test the instrument's "usability" and run a very preliminary test of reliability, we also employed and trained two students from the graduate program in Special Education. The graduate students rated six randomly selected titles from our sample (20% of the sample). The mean percent of agreement between these two raters was 93% (.89 to .96).

Additional modifications to the software evaluation instrument were made after our data analyses (see below). The final version of the instrument available in Appendix 2. While that final version retains the four broad categories from our initial instrument (background, interface design, instructional design, content), the indicators and sub-indicators within each category have been refined to address item redundancy, specificity, and clarity.

Analysis and Findings

We found that a significant factor in the development of our instrument was the opportunity to review software and analyze preliminary data. Only then was it possible for us to understand how the instrument must be designed in order for the data generated to be valid, reliable, and meaningful.

Because we did not meet in the final round of reviews to obtain consensus on the items, we began our data analysis by averaging our Likert-scale responses across all items. One difficulty that became clear was that while the Instructional Design and Content sections of the instrument relied upon the item/sub-item approach discussed above, the Interface Design section did not. This meant that the Interface Design section had many more items that were not being "synthesized" into a broader appraisal. Consequently, we made one final adjustment to the instrument and added larger, synthesizing items for the learner support, affordances, aesthetics, and motivation subsections within the Interface Design section (Appendix 2). For purposes of this data analyses we calculated the means of the sub-items within each subsection and used those means to represent the larger item score.

Overall, the instrument included 17 larger items with scores from 1 (strongly disagree) to 5 (strongly agree) to examine across the Interface Design, Instructional Design, and Content categories. Means for each main category were calculated using the larger items scores within each section. Total scores for each software package were then calculated by adding the 3 scores of those main categories. The largest possible total score was 15. In our sample, M=6.47 (max = 8.7, range = 8.67 to 4.48, SD = 0.99). Table 2 shows the breakdown of mean and the maximum score given to an item per category and subcategory.

	Mean	Max	Total Possible	SD
TOTAL SCORE	***6.47***	***8.7***	***15***	***0.99***
Interface Design	**3.02**	**4.00**	**5.00**	**0.66**
Learner Support	3.02	4.00	5.00	0.67
Affordances	3.20	4.61	5.00	0.84
Aesthetics	3.48	4.80	5.00	1.04
Motivation	2.37	3.69	5.00	0.83
Instructional Design	**2.04**	**3.70**	**5.00**	**0.67**
Instructional sequences	1.95	4.00	5.00	0.77
Well-designed activities	2.00	4.00	5.00	0.88

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

	Mean	Max	Total Possible	SD
Tracks learner progress	2.17	4.00	5.00	0.95
Content	**1.41**	**1.90**	**5.00**	**0.26**
Phonological Awareness	1.20	2.80	5.00	0.45
Sound Level	1.29	5.00	5.00	0.96
Sentence Level	1.00	1.00	5.00	0.00
Word Level	1.48	4.00	5.00	1.03
Syllable Level	1.14	4.00	5.00	0.65
Onset-Rime level	1.10	3.00	5.00	0.44
Phoneme Level	1.19	4.00	5.00	0.68
Alphabetic Understanding	1.16	2.75	5.00	0.56
Letter-Sound Correspondence	2.10	4.00	5.00	0.94
Blending	1.29	3.00	5.00	0.64
Segmenting	1.00	1.00	5.00	0.00
Word Reading	2.07	4.00	5.00	1.19

Table 2. Mean, max, total possible points, and standard deviation per category and subcategory.

Overall, our software sample did not rate highly on instructional design items (Mean = 2.04, Max. Given = 3.7, Total Possible Points = 5). Only 24% (5 out of 21 programs) of the sample rated above "disagree" for explicit instruction. Similarly, only 24% of the sample included content in phonological awareness. Only 2 programs (10% of the sample) included activities at the sound level (*Thinkin' Things, Kid Phonics 1*). None of the programs addressed phonological skills at the sentence level, 19% (4 out of 21) of the sample included activities at the word level (*Dr. Seuss Preschool and Kindergarten Reading, Richard Scary's Busytown, Kid Phonics I*), and one program addressed phonological skills at both the syllable level and onset-rime levels (*Lexia Early Reading*). Finally, only 2 programs (10% of the sample) included activities at the phoneme level (*Kid Phonics I, Lexia Early Reading*).

Alphabetic understanding was equally underrepresented with only 28% (6 out of 21 programs) of the sample rating above "disagree" for general alphabetic understanding. Within the alphabetic understanding category 28% (6 out of 21) rated above "disagree" for letter-sound correspondences, 9.5% (2 out of 21) rated above "disagree" for decoding (i.e., blending), 0% (0 out of 21) rated above "disagree" for segmenting skills for spelling, and 33% (7 out of 21) rated above "disagree" for word reading. Programs including alphabetic understanding content were *Simon Sounds It Out, Letter Sounds, Reader Rabbit Interactive Reading Journey, Reader Rabbit Kindergarten,* and *Read, Write, and Type*.

Because of the preliminary nature of our findings, we are reluctant to report the overall rankings of our sample. However, Table 3 does list the software packages that might be described as "exemplary" and "non-exemplary" in each category based on our review.

	Exemplary	Non-exemplary
TOTAL SCORE	• Jump-Start 1st Grade • Lexia Early Reading • Dr. Seuss Preschool & Kindergarten Reading	• Reading Who? Reading You! • Kid Phonics 1 • Bailey's Bookhouse
Interface Design	• Dr. Seuss Preschool & Kindergarten Reading • Jump-Start 1st Grade • Arthur's Reading Games	• Reading Who? Reading You? • Kid Phonics 1 • Lexia Early Reading
Instructional Design	• Jump-Start 1st Grade • Lexia Early Reading • Simon Sounds it Out	• Kid Phonics 1 • Bailey's Bookhouse • Alphabet Express
Content	• Lexia Early Reading • Simon Sounds it Out • Letter Sounds	• Alphabet Express • Reading Blaster-Spelling Blaster • Bailey's Bookhouse • Jump-Start 1st Grade • Kid Phonics 2

Table 3. Examples and non-examples within each category.

We also examined potential correlations among various indicators. Pearson product-moment correlation coefficients were calculated for the software programs total scores, price, interface design subscores, instructional design subscores, and content subscores. We found there was no apparent relationship between the software program's cost and its total score on the evaluation (r = 0.18) (see Figure 1).

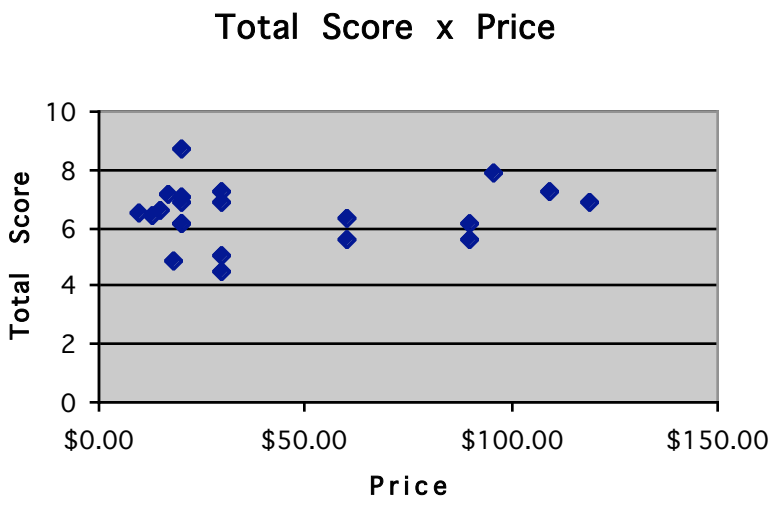

Figure 1. Total Score compared to Price of software program.

Similarly, we found no apparent association between interface and instructional design subscores (r = 0.06). However, we did find a strong positive correlation (r = 0.65) between instructional design and content subscores (see Figure 2) as well as a strong negative correlation (r = -0.55) between interface design and content subscores (see Figure 3).

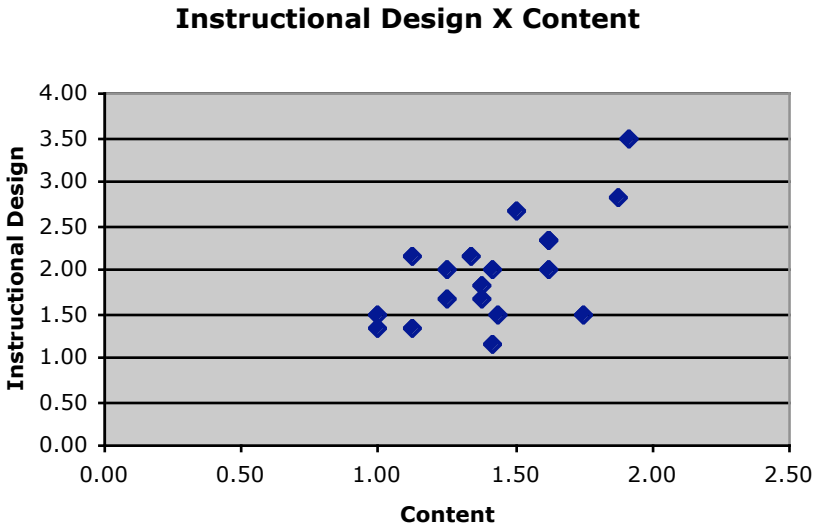

Figure 2. Instructional Design compared to Content.

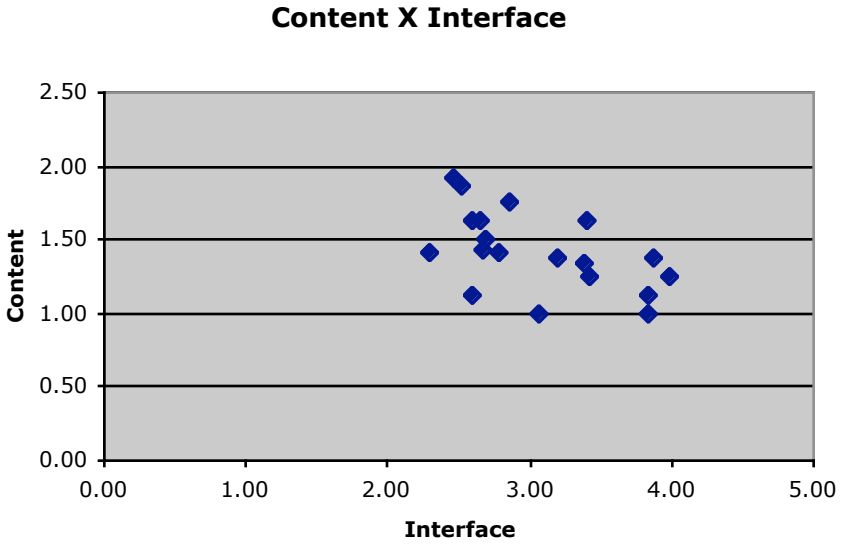

Figure 3. Content compared to Interface.

Discussion

While results must be interpreted cautiously due to continued development and technical validation of our software evaluation instrument, the findings from our review support existing research on computer-assisted instruction in beginning reading interventions. Like the most recent meta-analysis of software's effectiveness in beginning reading instruction (Blok, Oostdam, Otter, & Overmaat, 2002) and the National Reading Panel's (2000) discussion on software use in intervention research, our findings also suggest that the quality of

instructional software for beginning reading instruction "needs a great deal of additional exploration" (National Institute of Child Health and Human Development, 2000, ch. 6, p. 2). Even though we would not expect all software programs to receive perfect scores on our evaluation instrument, not all programs need to include all aspects of beginning reading content, there was still a considerable difference between scores software programs received on our evaluation instrument and the total possible score programs could receive when reviewed. Unfortunately, as a whole, our reviewed sample did not meet research-based criteria for interface, instructional design, and beginning reading content required for at-risk learners.

Our preliminary findings also suggest that you don't necessarily get what you pay for when considering the software in the reviewed sample. In other words, price didn't appear to be associated with final scores on the evaluation instruction. This finding has implications for teachers and parents, as well as suggests the importance of using a validated software evaluation tool to assess program content and overall quality. If purchasing the most expensive software program doesn't ensure quality content and instructional design, then using a software evaluation instrument to evaluate a program's design features and beginning reading content may have practical utility for teachers and parents making software purchasing decisions.

Our software review also revealed that there was a relation between content and instructional design. In other words, programs that were rated more highly on beginning reading content also tended to be rated more highly on instructional design. For example, if a program had strong alphabetic content, then activities with letter sound correspondences would include an explicit instructional model, consistent error corrections, and scaffold levels of difficulty. In other words, our preliminary results suggest a somewhat intrinsic link between the quality of the beginning reading content and how well it is taught.

Unlike the relationship between content and instructional design, however, findings suggested a negative relation between beginning reading content and interface design. Programs that tended to have stronger content, also tended to have a weaker interface. Perhaps not completely an unexpected result, the sample generally supported the stereotypical impression that instructional software usually had a somewhat boring and cumbersome interface design. For example, programs with strong beginning content were more likely to have uninteresting graphics, poor quality audio features, and disconnected activities. Programs rated more highly on interface design, with engaging and motivating activities, immersive and thematic environments, and easy to navigate directions, were more likely to have poor beginning reading content. These initial findings suggest the importance of further research in the relation between interface design and content when reviewing software programs and evaluating their effectiveness in intervention research.

Our review also revealed that most of the beginning reading software in our sample did not adequately address phonological awareness skills. While some software programs included sound matching and rhyming activities, only 2 programs (10% of the sample) included activities at the phoneme level. Our results are similar to findings by the Smith et al. (2001) review of commercial basal reading programs. The Smith et al. study found that very few programs adequately addressed phonemic segmentation skills across the reviewed curricula. Given the importance of phonological awareness in beginning reading instruction, sparse attention to phonological awareness in both reviews is a serious concern (Smith, Simmons, & Kame'enui, 1998). Finally, while we found that most programs did address skills related to alphabetic understanding, improvement in instructional quality was still needed. For example, most of the programs included activities directed at recognizing and identifying letter names and reading and spelling words. Despite the inclusion of alphabetic and word reading activities across the reviewed sample, the instructional presentation of content was often not explicit enough for children with reading difficulties.

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

Limitations and Future Directions for Research

While our findings are suggestive of possible emerging trends in the interface, instructional design, and beginning reading content of a sample of early reading software (e.g., inverse relationship between interface and beginning reading content, inadequate attention to phonological awareness), the primary focus of our study was the development of an evaluation instrument. Therefore, the results of our software evaluation must be viewed as preliminary. For example, even though we were able to achieve high levels of reliability by asking two reviewers to independently rate 20% of our sample, more stringent reliability with additional reviewers is required. The software evaluation instrument also requires further validation. For example, though components of our evaluation instrument are based on research-based criteria for interface, instructional design, and beginning reading content, experts in instructional technology and beginning reading instruction could also be asked to review the instrument.

In addition, future research will require empirical and quantitative support for the interface, instructional design, and beginning reading constructs used in the evaluation framework. To truly assess whether interface, instructional design, and content are contributing factors to overall software quality, a follow-up software review with a larger sample of software, additional raters, and confirmatory factor analyses is required. Future research may also include using a final version of the software evaluation instrument to review a newly identified sample of software and conduct a comparative analysis of the software to determine the highest ranking programs. If any highly ranked programs are identified, the rankings could be validated by use of the software in beginning reading intervention studies. Finally, the researchers hope to derive a set of guidelines for the development of new software for learners who struggle with reading skills.

Conclusions

Similar to Blok, Oostdam, Otter, and Overmaat (2002) and the National Reading Panel (2000) our initial findings suggest caution when considering the overall content of the popular beginning reading software reviewed in this study. We cannot recommend any specific titles until additional research is conducted.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT.

Alessi, S. M. & Trollip, S. R. (1991). Computer-based instruction: Methods and development (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Ary, D., Jacobs, L.C., & Razavieh, A. (2002). *Introduction to research in education* (6th ed.). Stamford, CT: Wadsworth/Thomson Learning.

Berelson, B. (1952). *Content analysis in communication research*. Glencoe, IL: Free Press.

Blok, H., Oostdam, R., Otter, M. E. & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A Review. *Review of Educational Research, 72*(1), 101-130.

Borg, W.R., & Gall, M.D. (1983). *Educational research: An introduction* (4th ed.). New York: Longman.

Brown, C. M. (1988). Human-computer interface design guidelines. Norwood, NJ: Ablex.

Burton-Radzely, L. (1998). *A national perspective on special educator's use of technology to promote literacy: Technical report*. Washington DC: MACRO International and the Council for Exceptional Children.

Cates, W.M., & Bishop, M.J. (in press). Learner as bobsled operator: The physics of learner engagement. *Journal of Educational Technology Systems, 31*(3).

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

Cates, W. M., & Bruce, R. (2000). Conceptualizing learner support space. Educational Technology Research & Development, 48(1), 85-98.

Coley, R., Cradler, J., & Engel, P. K. (1997). *Computers and classrooms: The status of technology in U. S. schools*. Princeton, NJ: Educational Testing Service, Policy Information Center.

Fleming, M. L. (1987). Displays and communication. In R. Gagné (Ed.), Instructional technology: Foundations (pp. 233-260). Hillsdale, NJ: Erlbaum.

Felton, R. H., & Pepper, P. P. (1995). Early identification and intervention of phonological deficits in kindergarten and early elementary children at risk of reading disability. *School Psychology Review, 24*, 405-414.

Foorman, B.R., Francis, D. J., Fletcher, J. M., Schatschnedier, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*, 37-55.

Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children*, Learning Disabilities Research and Practice, 16*(4), 203-212.

Gagné, R. M., & Driscoll, M. P. (1988). Essentials of learning for instruction (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hall, T., Hughes, C., & Filbert, M. (2000). Computer assisted instruction in reading for students with learning disabilities: A research synthesis. *Education and the treatment of children, 23*(2), 173-193.

Hannafin, M. J., & Rieber, L. P. (1989). Psychological foundations of instructional design for emerging computer-based instructional technologies: Part II. Educational Technology Research and Development, 37(2), 102-114.

Inspiration 7.0 (2003). [Computer software]. Portland, OR: Inspiration Software.

Jitendra, A., Nolet, V., Gomez, O., Xin, Y.P. (1999). An Analysis of Four Middle School Geography Textbooks: Meeting the Needs of Students with Learning Problems. ERIC Document Reproduction Service No. ED433271

Just Grandma and Me (n.d.). [Computer software]. Novato, CA: Broderbund/The Learning Company.

Kame'enui, E. J., & Carnine, D. W. (1998). *Effective teaching strategies that accommodate diverse learners*. Columbus, OH: Merrill, Prentice Hall.

KidPix Deluxe 3 (2003). [Computer software]. Novato, CA: Broderbund/The Learning Company.

Krippendorff, K. (1980). *Content analysis, an introduction to its methodology.* Beverly Hills, CA: Sage.

Lang, M., McCarty, C., Norman, L., & Upchurch, N. (XXXX). Improving reading strategies through the use of technology. ERIC Document Reproduction Service No. ED433500

Laurel, B. (1986). Interface as mimesis. In D. A. Norman and S. W. Draper (Eds.), User centered system design: New perspectives on human-computer interface (pp. 67-85). Hillsdale, NJ: Erlbaum.

Laurel, B. (Ed.) (1990). The art of human computer interface design. Reading, MA: Addison-Wesley.

Lyon, G. R. (1997). *Children's Literacy*. Testimony before the Committee on Education and the Workforce. U. S. House of Representatives, Washington, DC.

Lyon, G. R. (2001, October 19). *White House Early Childhood Cognitive Development Summit: Summary Comments*. Retrieved from http://www.educationnews.org/white_house_early_childhood_cogn.htm

MacArthur, C. A., Ferretti, R. P., Okolo, C. M., & Cavalier, A. R. (2001). Technology applications for students with literacy problems: A critical review. *Elementary School Journal, 101*(3), 273-301.

Martindale, T., Qian, Y., & Cates, W.M., (2001). Categorizing Exemplary Educational Websites. Proceedings of the Annual Meeting of the Association for Educational Communications and Technology. Atlanta, GA.

National Center for Education Statistics. (2001). *National assessment for educational progress (NAEP): Report card for the nation and states*. Washington DC : Author.

National Reading Panel (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and is implications for reading instruction: Reports of the subgroups*. Rockville, MD: NICHD Clearinghouse.

National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction:*

Reports of the subgroups. (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office (www.nichd.nih.gov/publications/nrp/report.htm).

Norman, D. (1986). Cognitive engineering. In Norman, D. & Draper, S. W. (Eds.), User centered system design: New perspectives on human-computer interface (pp. 31-61). Hillsdale, NJ: Erlbaum.

Norman, D. (1988). *The design of everyday things*. New York: Doubleday.

Rieber, L. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. Educational Technology Research and Development, 44(2), 43-58.

Rieth, H.J., & Semmel, M.I. (1979). The use of microcomputer technology to prepare and to enable teachers to meet the educational needs of handicapped children. *Teacher Education and Special Education,2*(2), 56-60.

Shiratuddin, N., & Landoni, M. (2002). Evaluation of content activities in children's educational software. *Evaluation and Program Planning, 25*, 175-182.

Shneiderman, B. (1992). Designing the user interface: Strategies for effective human-computer interaction (2nd ed.). Reading, MA: Addison-Wesley.

Smith, S., Simmons, D. C., Gleason, M. M., Kame'enui, E. J., Baker, S. K., Sprick, M., Gunn, B., Thomas, C. L., & Chard, D. C. (2001). An analysis of phonological awareness instruction in four kindergarten basal reading programs. *Reading and Writing Quarterly, 17*, 25-51.

Smith, S., Simmons, D. C., & Kame'enui, E. J. (1998). Phonological awareness: Research bases. In D. C. Simmons & E. J. Kame'enui (Eds.), *What reading research teaches us about children with diverse learning needs: Bases and basics* (pp. 61-127). Mahwah, NJ: Lawrence Erlbaum Associates.

Tamir, P. (1985). Content analysis focusing on inquiry. *Journal of Curriculum Studies, 17*, 87-94.

Torgesen, J. K. (1998). Catch them before they fall. *American Educator, 22*(1), 32-39.

Woodward, V.A. (1984). Literacy learning in early childhood. ERIC Document Reproduction Service No. ED262880

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

Appendix 1. Software Sample

	Name	Publisher	Date	Ages	Cost	ASCD Subjects	Reason for exclusion if not reviewed
	A to Zap	Sunburst Communications	1993	PreK-1	$79.95	Early Childhood	Alphabet recognition only.
1	Alphabet Express	School Zone Publishing	1996	PreK-1	$19.99	Early Childhood, Language Arts, Reading & Problem Solving	
2	Arthur's Reading Games	Broderbund Software	1997	PreK-2	$19.95	Early childhood, language arts, reading	
3	Bailey's Bookhouse	Edmark Corporation	1995	PreK-2	$29.95	Early childhood, language arts, special education, ESL	
4	Big Thinkers Kindergarten	Humongous Entertainment	1997	K-1	$12.95	Early childhood, pre-reading, reading, music, mathematics	
	Clue In On Phonics	GAMCO Educational Software	1994	K-4	$59.95	Early childhood, reading	Could not obtain.
	Cornerstone Language Arts Series	Skillsbank	1998	3-8	$387.00	Language arts, grammar, reading, spelling	Storybooks
	Dr. Seuss's ABCs	Living Books/ Broderbund	1995	PreK-1	$49.95	Early childhood, language arts	Storybook
	Dr. Suess's The Cat in the Hat	The Learning Company	1998	PreK-1	$19.95	Early childhood, language arts, reading	Storybook
	Dr. Suess's Green Eggs & Ham	Living Books / Broderbund	1996	PreK-1	$49.95	Early childhood, language arts, reading	Storybook
	Eager to Learn	Sierra Online	1996	PreK-3	$30.00	Early childhood, language arts, mathematics, reading	Could not obtain.
5	Dr. Suess's Preschool & Kindergarten Reading	Broderbund	1998	PreK-K	$29.99	Early childhood, spelling, reading	
	Elmo's Preschool Deluxe (Sesame Street Learning Series)	The Learning Company	1997	PreK-1	$34.95	Early childhood, reading, mathematics, problem solving	Alphabet recognition only.
6	Jump Start First Grade	Knowledge Adventure	1995	K-2	$19.95	Early childhood	
7	Kid Phonics 1	Knowledge Adventure	1995	PreK-2	$17.95	Language arts, phonics, reading	
8	Kid Phonics 2	Knowledge Adventure	1996	1-3	$16.95	Early childhood, language arts	
	Kinder Critters	Micrograms Publishing	1991	K	$12.95	Early childhood, problem solving, pre-reading	Could not obtain.
	Learn About Animals	Sunburst Communications	1992	K-2	$79.95	Early childhood, science, special education	Very little on reading.
	Learning to Read on the Prominade: Kindergarten	Sunburst Communications	1998	K-1	$79.95	Language arts, reading, phonics	Could not obtain.
9	Letter Sounds	Tenth Planet / Sunburst Communications	1998	PreK-1	$89.95	Language arts, reading, phonics	
10	Lexia Early Reading	Lexia Learning Systems	1994	1-12	$96.00	Language arts, reading	
11	Lexia Phonics-based Reading	Lexia Learning Systems	1994	1-12	$119.00	Language arts, reading	
	Mixed-Up Mother Goose Deluxe	Sierra Online	1995	PreK-3	$19.95	Early childhood, problem solving, language arts	Could not obtain.
	Putt-Putt Joins the Parade	Humongous Entertainment	1993	PreK-2	$14.95	Early childhood, problem solving	Could not obtain.
	Little Monster At School	Living Books / Broderbund	1994	K-2	$39.95	Early childhood, language arts	Storybook
12	Read, Write, and Type	The Learning Company	1995	1-3	$9.95	Language arts, tools	

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

	Name	Publisher	Date	Ages	Cost	ASCD Subjects	Reason for exclusion if not reviewed
13	Reader Rabbit Reading 2	The Learning Company	1991	1-3	$14.95	Language Arts, Early childhood	
14	Reader Rabbit Interactive Reading Journey	The Learning Company	1994	K-1	$89.95	Early childhood, language arts, reading	
15	Reader Rabbit's Kindergarten	The Learning Company	1997	K-3	$19.99	Early childhood, mathematics, reading, problem solving	
	Ribbit Collection	Little Planet (Vanderbilt University's Learning Technology Center)	1996	K-4	$496.00	Language arts, early childhood	Could not obtain.
16	Reading Blaster: Spelling Blaster	Knowledge Adventure	1998	1-4	$59.95	Language arts, spelling	
17	Reading Who? Reading You!	Sunburst Communications	1996	K-2	$30.00	Language arts, reading, spelling, phonics	
18	Richard Scarry's Busytown	Computer Curriculum Corp.	1993	PreK-1	$19.99	Early childhood	
	Stickybear's Early Leaning Activities	Optimum Resource	1995	PreK-1	$59.95	Early childhood, language arts, mathematics, foreign languages, ESL	Could not obtain.
19	Simon Sounds It Out	Don Johnston, Inc	1997	K-3	$109.00	Language arts, reading, spelling, phonics.	
20	Thinkin' Things Collection	Edmark Corporation	1993	PreK-4	$29.95	Early childhood, problem solving, special education	
21	Vowels: Short and Long	Tenth Planet/Sunburst Communications	1999	K-2	$89.95	Language arts, reading, phonics, spelling	
	Write: Outloud	Don Johnston, Inc.	1994	K-12	$99.00	Early childhood, language arts, writing	Tool software.
	Zurk's Learning Safari	Soleil Software	1995	PreK-1	$59.95	Early childhood	Very little on reading.

Appendix 2: EARLY READING SOFTWARE EVALUATION FORM

Background Information

Software name and version number:							
Publisher:							
Date:		Cost:	$	Operating System (Mac/Win/both):		Evaluated under:	Mac? Win?
Package contents:							
Supplementary materials:							
Required materials:							
Survives "pound-the-keyboard" test:	Yes No						

	Description	SD	D	N	A	SA
Stated goal and objectives are clearly defined as:						
Stated target audience is clearly defined as:						
Stated prerequisite skills are clearly defined as:						

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

Interface Design

LEARNER SUPPORT		SD	D	N	A	SA
1. The program fully supports learners' efforts to use the software.						
	Instructions in manual are specific and to the point.					
	Instructions can be reviewed on the screen, if necessary.					
	Instructions within program (text, voice, and the like) are helpful.					
	The interface responds with prompt and informative invalid action messages when appropriate.					
	Most learners should be able to use the program with little help from adults and without being able to read.					

AFFORDANCES		SD	D	N	A	SA
2. The program supplies strong clues about the operation of elements and functions.						
	Learners don't have to search for commonly used functions. Program functions (like navigation and help) are placed in equivalent (if not identical) locations on screens.					
	The interface anticipates learner needs (prompts learner to action, makes suggestions, interprets aimlessness and remediates).					
	The interface takes advantage of what users already know (represents functions with familiar graphical icons, uses a locational setting to represent program components).					
	Things on the screen are what they appear to be and function as expected.					
	Inputs are discreet for various functions (rollover to see or hear then click to choose rather than click for both).					
	The interface requires as few mouse clicks or keystrokes as possible.					
	It's easy to get in or out of any activity at any point.					
	It's easy to get out of the program at any point.					
	The interface asks for confirmation only when the action results in serious damage to the learning experience.					

AESTHETICS		SD	D	N	A	SA
3. The program's interface is aesthetically pleasing.						
	The interface reduces cognitive and visual load (screen display is uncluttered).					
	Screens are visually attractive and are likely to be pleasing to the user (not boring or tiresome).					

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

AESTHETICS		SD	D	N	A	SA
	Media used is high quality.					
	Visuals (text, graphics, animations, video, color, and layout) are designed in ways that enhance the experience.					
	Uses media to create themes/metaphors that relate to the content and help create meaning.					

MOTIVATION		SD	D	N	A	SA
4. The program motivates learners to action.						
	Learner is primarily the driving force behind what happens on screen (doesn't watch passively as program does things, title sequence is brief or can be bypassed, learner controls pace).					
	The program stimulates curiosity (random generation techniques employed, exploration rewarded with surprises).					
	Activities require learner to be actively engaged (both mentally and physically).					
	Learner is able to control/modify environment in appropriate ways (adjust font sizes, change colors, pick background music, etc.).					
	The learner's ideas can be incorporated into the program (controls avatar/character that acts within the environment, makes decisions that alter the environment in substantive ways).					
	The program encourages play (competition, rewards, consequences).					
	Learners can easily access different levels of program content (challenge).					
	The program is fun to use.					

Instructional Design

INSTRUCTIONAL DESIGN:		SD	D	N	A	SA
5. The program includes well-designed, explicit instructional sequences (or processes) for skills.						
	The program gains learners' attention at the beginning of each instructional cycle.					
	The program reminds learners of prerequisite knowledge at the beginning of each instructional cycle.					
	The program informs learners of objectives at the beginning of each instructional cycle.					
	The program uses a "model-lead-test" framework.					
	The program offers multiple examples of a target skill.					
	The program supplies adequate opportunities for learners to practice new skills.					
	The program presents the same requirements in embedded activities as presented in instructional sequence.					

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

INSTRUCTIONAL DESIGN:		SD	D	N	A	SA
	The program requires learners to demonstrate mastery of newly introduced skills before moving on to new skills.					
	The program supplies larger conceptual anchors for retention and retrieval (transfer of knowledge).					
	The program moves systematically and cyclically through the phonological awareness skill hierarchy (sounds, sentences, words, syllables, onset-rimes, phonemes) and/or word-reading skill hierarchy (letter-sound, blending, segmenting, word reading).					
6. The program includes effective embedded or drill-and-practice activities (if part of an instructional sequence they're "embedded," of not, they're drill-and-practice).						
	The program uses informative and instantaneous feedback messages to support content learning.					
	The program limits number of tries to avoid frustrating learners.					
	The program makes content support available precisely when the learner needs it.					
7. The program monitors and helps direct learner progress.						
	The program saves learners' work.					
	The program graphs or charts learner performance in an easily interpreted way.					
	The program helps determine and/or identify goals and monitor progress.					
	The program supplies learners' fluency rates (such as letters/sounds identified per minute).					
	The program branches automatically to accommodate learner needs (such as going back to earlier material if learner is having difficulty with newer material).					

PHONOLOGICAL AWARENESS:		SD	D	N	A	SA
8. SOUND LEVEL -- The program develops learners' skill at identifying and comparing environmental sounds.						
	The program includes activities that ask learners to match similar environmental sounds.					
	The program includes activities that ask learners to put previously played environmental sounds in the order in which they were heard.					
	The program uses concrete representations (interactive visuals) for manipulating (matching, comparing, sequencing, isolating) whole sounds.					
9. SENTENCE LEVEL -- The program develops learners' skill at identifying individual words within a spoken sentence.						
	The program includes auditory activities that require students to isolate words within a sentence (such as clapping for individual words within a sentence).					
	The program includes auditory activities that require the learner to supply the order of words spoken in a sentence.					

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

PHONOLOGICAL AWARENESS:		SD	D	N	A	SA
	The program uses concrete representations (interactive visuals) for manipulating (matching, comparing, sequencing, isolating) whole words within a sentence.					
10. WORD LEVEL -- The program develops learners' skill at listening to differences and similarities among spoken whole words.						
	The program provides sufficient practice at auditorially matching whole words.					
	The program provides sufficient practice auditorially identifying words that rhyme.					
	The program provides sufficient practice auditorially producing words that rhyme.					
	The program develops rhyming skills auditorially before text of words appears (match pictures of nose and toes rather than the words themselves).					
	The program uses concrete representations (interactive visuals) for manipulating (matching, comparing, isolating) whole words.					
11. SYLLABLE LEVEL -- The program develops learners' skill at auditorially detecting and identifying the syllables in a word (such as /foot/ /ball/; /hot/ /dog/).						
	The program provides sufficient practice at auditorially matching syllables (e.g., "Which word has the syllable 'ball?', baseball or hotdog?")					
	The program includes auditory activities that require students to isolate syllables within a word (such as clapping for individual syllables within a word).					
	The program includes auditory activities that require the learner to supply the order of syllables spoken in a word.					
	The program uses concrete representations (interactive visuals) for manipulating (matching, comparing, sequencing, isolating) syllables in the word.					
	The program develops learners' skills at blending syllables (What word is this? /foot/ /ball/ "football").					
	The program develops learners' skills at segmenting syllables (Identify the two sounds parts in the word "football").					
	The program introduces blending (What word is this? /foot/ /ball/ "football") before segmenting (Identify the two sounds parts in the word "football").					
12. ONSET-RIME LEVEL—The program develops learners' skill at auditorially breaking words into two parts: 1) the sounds before the first vowel and 2) the vowel and the rest of the sounds (m|at; cl|ock; f|ox)						
	The program includes auditory activities that require students to isolate the onset-rime within a word (such as clapping for individual syllables within a word).					
	The program includes auditory activities that require the learner to supply the order of the onset-rime in a spoken word.					
	The program uses concrete representations (interactive visuals) for manipulating (matching, comparing, sequencing, isolating) the onset-rime in the word.					
	The program develops learners' skills at blending onset-rimes (What word is this? /mmm/ /aaat/ "mat").					
	The program develops learners' skills at segmenting onset-rimes (Identify the two sounds parts in the word "mat").					

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

PHONOLOGICAL AWARENESS:		SD	D	N	A	SA
	The program introduces blending (What word is this? /mmm/ /aaat/ "mat") before segmenting (Identify the two sounds parts in the word "mat").					
13. PHONEME LEVEL – The program develops learners' skill at auditorially breaking words into individual sounds (/mmm/ /aaa/ /t/; /c/ /lll/ /ooo/ /ck/; /fff/ /ooo/ /k/ /s/)						
	The program provides sufficient practice at auditorially matching phonemes (e.g., "which words have the same beginning sound: bet, bag, or mop?" or which words have the same ending sound: mop, get, tap?")					
	The program first targets the beginning phoneme, then the last phoneme, then the middle phoneme in a word.					
	The program identifies and separates a single phoneme from a larger context (e.g., "the first sound in run is /r/")					
	The program uses concrete representations (interactive visuals) for manipulating (matching, comparing, sequencing, isolating) phonemes in the word.					
	The program develops learners' skills at blending phonemes (What word is this? /mmm/ /aaa/ /t/ "mat")					
	The program develops learners' skills at segmenting phonemes (Identify the sounds in the word "mat").					
	The program introduces blending (What word is this? /mmm/ /aaa/ /t/ "mat") before segmenting (Identify the sounds in the word "mat").					
	The program introduces "continuous" phonemes before "stop" phonemes.					

ALPHABETIC UNDERSTANDING:		SD	D	N	A	SA
14. LETTER-SOUND CORRESPONDENCE: The program develops learners' skill at identifying letter-sound correspondences (e.g., What is the sound of this letter?).						
	The program introduces first sounds in words.					
	The program introduces last sounds in words.					
	The program introduces middle sounds in words.					
	The program introduces first, then last, then middle sounds in words in this sequence.					
	The program introduces continuous sounds first (such as /mmm/ /lll/ /rrr/ and all vowels –anything you can stretch out without distorting the sound).					
	The program introduces regular sounds first --such as all the short vowel sounds and commonly used consonant sounds (such as the hard "c" in cat rather than the soft "c" in city).					
	The program auditorially introduces letter sounds before letter names.					
	The program separates visually or auditorily similar letters (p/b/d, e/i, m/n).					
	The program introduces more useful letters (e.g., m, s, a, t) before less useful letters.					
	The program introduces lower case before upper case (some exceptions–s, u , w , r, t, f).					

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

ALPHABETIC UNDERSTANDING:		SD	D	N	A	SA
	The program requires learners to manipulate the letter-sound correspondences in words (e.g., What word would you have if you change the /nnnn/ in “nap” to /llll/.”).					
15. BLENDING (DECODING, SOUNDING OUT): The program develops learners’ skill at blending letter-sound correspondences within a word (e.g., “Blend the sounds of these letters to make a word,” /mmm-aaa-nnn/). If program requires to use blending skills but does not explicitly teach blending, then disagree.						
	The program introduces blending before whole-word reading.					
	The program includes pseudowords when teaching blending.					
	The program introduces blending before segmenting.					
16. SEGMENTING: The program develops learners’ skill at segmenting letter-sound correspondences within a word (e.g., What sounds are in the word on the screen? /mmm/ /aaa/ /nnn/). If program requires to use segmenting skills but does not explicitly teach segmenting, then disagree.						
	The program incorporates segmenting into spelling activities.					
	The program uses concrete representations (interactive visuals) for manipulating (matching, comparing, sequencing, isolating) letters in the word.					
	The program includes pseudowords when teaching segmenting.					
17. WORD READING: The program develops learners’ word reading skills.						
	The program introduces regular word reading (words that include common sounds such as “cat,” “map, “let”) before irregular word reading (words that include uncommon sounds such as “are,” “city”)					
	The program introduces regular words before sight words (such as “and,” “the,” “is”).					
	The program uses an appropriate sequence of word types (e.g., VC and CVC that begin with continuous sounds, VCC and CVCC that begin with continuous sounds, CVC that begin with stop sounds, CVCC that begin with stop sounds, etc.).					
	The program introduces new words in isolation vs. in context (within a story).					
	The program includes pseudowords when practicing word reading.					
	The program ensures that the reading passages or words within the activities are at the learners’ skill level (readability).					
	The program provides overlap in words (i.e., words are embedded across multiple texts).					

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003

SUMMATIVE EVALUATION

	Superior	Very Good	Good	Fair	Poor	COMMENTS
Content:						
Instructional Design:						
Interface Design:						
Stated goal and objectives were met:						
Stated target audience was served:						
Stated prerequisite skills were accurate:						

Notable Features:

Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April 2003