

---

## **PERSONALIZED INSTRUCTION FOR LITERACY: Gaining Mastery of Language Arts and Writing Using Student-Centered Adaptive Technology**

*By Dr. Dan Flickinger*

*Senior Research Associate at the Stanford University Center for the Study of Language and Information (CSLI), and Senior Research and Language Consultant for Redbird Language Arts & Writing*

**L**anguage Arts and writing are foundational skills that open wide the doors of learning and prepare students for 21<sup>st</sup> century careers. Mastery of these skills requires sustained effort and practice on the part of the motivated student, and equally sustained instructional feedback for error correction and for individual guidance at each step. Providing this level of detailed and consistent instructional response to each student's efforts is challenging for teachers in classrooms of students learning at differing rates. Recognizing an opportunity for computers to immediately provide much of the necessary feedback and step-by-step instruction in language arts, Dr. Patrick Suppes and his research team at Stanford launched an ambitious project in 2003 to develop an online language arts curriculum for elementary school students, building on decades of ground-breaking work in adaptive learning technology.

Using a similar design methodology to that used in the mathematics curriculum produced in Stanford's Education Program for Gifted Youth (EPGY), a team of educators, linguists, and software developers collaborated for several years to construct a fully automated course that paired short instruction with sets of exercises and sophisticated answer evaluation. This course, for students in grades 2-7, covered the core concepts in language arts, with a strong emphasis on teaching and practicing writing skills. Built on the same underlying technology platform as that used in the mathematics curriculum, the implementation of the language arts courses uses an algorithm developed at EPGY to adapt the presentation of each concept and exercise to the individual student's rate of progress in mastering the material. To evaluate the efficacy of this method on a large scale over time, the Stanford EPGY team worked with school districts in several states to implement the course in a diverse set of classrooms, again following the research methodology developed for evaluation of the mathematics curriculum.

Recognizing that EPGY now stood at a critical crossroads of potential outreach and opportunity for technological enhancement, Stanford identified Redbird Advanced Learning to develop the next generation of adaptive learning courses. Building upon EPGY's distinguished legacy, this development continued to be guided by Stanford research in collaboration with Redbird Advanced Learning, part of Rocket Group, a company with a rich history in diverse educational offerings aiming to improve the lives of students through technology, research, and innovation. In 2016, Redbird was acquired by McGraw-Hill Education, with both the Mathematics and Language Arts & Writing courses now undergoing further development, in continued collaboration with Stanford.

This paper provides a discussion of the design and contents of the Redbird Language Arts & Writing course, summarizes the multi-year efficacy studies and their strongly positive results, and describes the ongoing research and development to expand the scope and instructional capabilities of the curriculum.

---

## Language Arts with an Emphasis on Writing

**M**astery of writing comes through steady, guided practice with detailed feedback at every step, teaching the fundamentals of grammar, clear expression of meaning, and an engaging style that makes one's writing enjoyable for the reader. The more a student writes given this feedback, the more fluent and confident the resulting composition, as written language becomes a familiar and powerful tool for communication and delight.

Providing this kind of consistent detailed feedback on writing efforts is an important but difficult and time-consuming burden for teachers and parents. Yet with the variety and complexity of linguistic expression available in human language, automation of this task by computers has been a daunting challenge for researchers in linguistics and education. Through several decades of research and experimentation in both fields at Stanford University, an effective computer-based method has been developed which provides this sentence-by-sentence analysis and feedback in the context of a full-featured course in English Language Arts.

The Redbird Language Arts & Writing (LA&W) course content spans Grades 2-7, consisting of over 800 concise lectures and 16,000 exercises, organized in eight to ten units per grade, where each unit's lessons include instruction in five broad areas:

- **Parts of Speech.**  
Provides practice identifying and using the grammatical categories of English words in context.
- **Sentence Structure.**  
Provides practice analyzing and identifying properties of English syntax.
- **Sentence Composition.**  
Provides students with immediate and detailed feedback on sentences they compose.
- **Paragraphs.**  
Teaches and gives practice in the writing of basic types of paragraphs: narrative, opinion, information, and persuasion.
- **Reading.**  
Introduces students to grade-appropriate short passages drawn from representative texts in prose and poetry, fiction and non-fiction, print and web, classic and contemporary, historical and scientific.

### Aligned to Modern Standards for Writing and Language Arts

**T**he LA&W curriculum is designed to align with the requirements of widely adopted standards for English Language Arts with a strong emphasis on writing. The key concepts identified in the standards are introduced at the appropriate grade level, together with sets of exercises that reinforce and elaborate on the lecture material but in addition, the ubiquitous composition exercises in every grade provide the student with steady practice using concepts and skills that have already been mastered. This layered reinforcement is employed in each subject area of the course, and particularly in the Reading exercises, where each lesson builds on and often refers directly to concepts introduced in earlier lessons, for comparison and review.

One central element of the end-of-year exams to evaluate student success in meeting state standards is on-the-spot composition by the student of short essays demonstrating mastery of the relevant concepts and skills across the full range of subject matter. The LA&W curriculum program, with its strong emphasis on practice in writing well-formed sentences and paragraphs, provides students with valuable and sustained practice in developing the necessary skills to be able to compose these written demonstrations of their academic progress with confidence and fluency.

---

Detailed concept-by-concept correlations have been established in the LA&W course for state-level standards for Language Arts, including those for California and Tennessee, where the multi-year efficacy studies were conducted. This correspondence with multiple established standards has led to enhanced depth and breadth in the curriculum, particularly where one standard provided more detailed specifications of certain concepts or targets.

### Pace Personalized to Each Student's Individual Needs

**A**s in the Redbird Mathematics curriculum program, LA&W provides highly individualized instruction through the use of a stochastic course motion engine developed at Stanford, which evaluates a student's level of mastery of a concept based on previous results as each exercise is attempted, and presents the next exercise or concept accordingly. When a particular concept proves more difficult to master, more exercises and more guidance are presented, while concepts that are more readily mastered quickly lead on to more challenging content. With this proven method, each student can proceed as quickly or slowly as necessary in gaining mastery of the concepts taught in the multiple strands of the course, steadily building on what has been learned previously, and constantly practicing and strengthening that mastery through the composition exercises woven into all aspects of the course.

The Language Arts & Writing course is self-paced and designed to detect and address gaps in student preparation. As a student progresses through the course, the software uses the results of prior assessments to individualize the student's sequence of exercises. Thus, students who readily master a concept move on quickly to new material, while students who need more guidance with the concept receive further instruction and practice. In addition, material that students have trouble mastering is reviewed with a higher degree of frequency than material that students learn quickly. Because of this continual assessment and individualization, the courses are ideal for students who have weak foundations, and who may be performing below grade level in some skills, while not in others.

As explained in Suppes et al. (2014), the stochastic or probabilistic aspect of the motion algorithm used to individualize the content of the courses provides an appropriate method of smoothing, both of student progress in a given concept class, or in the selection of which strand is to follow the current one. The stochastic choice of which strand appears next also keeps students from being inappropriately conditioned to expect that one particular subject always follows another. The real world does not have that restricted simplicity of predetermined order in which challenges appear, and neither should student learning. Rather, students should learn to move easily from one kind of concept in language arts to any other, from reading to composition to a lesson on paragraph structure.

In this approach to learning based on stochastic motion, a decision must be made after each exercise about what the student should do next. The structure of the Redbird course in Language Arts & Writing, as in Mathematics, provides for a student to work a sequence of exercises in a given concept class before leaving that class for the next one. Because of the striking individual differences in student learning of cognitive concepts, there is a learning model attached to each concept class, so that a student who is learning to master the particular concept quickly is required to do only a sequence of three or four exercises, the usual mastery criterion of the learning model. If this first short sequence of exercises is answered correctly, the student is moved on to the next concept, whereas if the mastery level is not reached even after working additional exercises in this class, the student is either set to repeat the concept, or to move back to a precursor concept for further review and preparation.

One further benefit of the stochastic motion in determining course behavior is that it permits a natural adaptation of movement in the courses to the individual learning differences of each student.

---

For example, one student may learn parts of speech more readily than sentence structure, or reading more readily than composition, while another student may show the opposite tendency. It is impossible to estimate in advance these subtle student differences, but by continually assigning different amounts of time to each strand of a course, depending upon the student's own work, accommodations of such differences can be made, by computing new probabilities each time a student has finished a concept. This motion demonstrates continuous and precise adjustment for each student that could not be done by a teacher on such an individual basis in the group setting of a standard classroom.

## Immediate Evaluation and Feedback in Composition Exercises

A distinctive feature of the course is the inclusion, throughout the grades, of exercises prompting students to compose sentences and paragraphs which are automatically evaluated for immediate and error-specific feedback, both syntactically and semantically. This detection of errors in the meaning of sentences and their interaction with errors of grammar and word choice presents a non-trivial research challenge and is one of the ongoing topics of investigation at Stanford's Center for the Study of Language and Information (CSLI). Further discussion of this research follows in this paper.

The automatic evaluation of student answers provided in the LA&W curriculum makes use of an efficient parser and a broad-coverage, precise grammar (Flickinger 2002, 2011) which has been extended to accommodate frequently occurring types of grammatical errors. Novel errors are analyzed using this extended set of grammar rules, and more than a hundred common error types can be detected in student compositions, triggering sentence-specific advice on what to correct, and how. Frequently occurring well-formed and ill-formed answers are manually validated offline and cached to maximize the accuracy of the automatic online evaluation. Presented with the error diagnosis, the student is given the opportunity to correct the sentence and submit it again for evaluation. Then, after giving the student feedback on this second attempt, the system moves on to the next exercise.

Computational linguists who develop grammar implementations often begin with the motivation to encode their hypotheses about the particular structures and the general principles which illuminate the analysis of any one language. Often these hypotheses are tested either on naturally occurring text corpora, or on systematically constructed test suites illustrating the range of linguistic phenomena under study, including both well-formed and ill-formed example sentences. Since many applications that make use of grammar implementations emphasize robustness of analysis over precision, it is rewarding for the grammarian to encounter an application where precision is demanded. The grammar-checking function with this sentence composition task is just such an application.

In order to provide accurate and detailed automatic analysis of students' errors in composing sentences, a linguistically informed grammar (and parser) can be useful, and perhaps even essential as the complexity of the sentences increases. Of course, a grammar designed to analyze only well-formed utterances needs to be augmented for this kind of application to include rules or mechanisms that also accommodate certain ill-formed inputs, since students' errors are exactly what present opportunities for learning. Within the Redbird LA&W course, we use a modified version of the English Resource Grammar (ERG), a highly accurate, broad-coverage implementation of a lexicon and rule set within the theoretical framework of Head-driven Phrase Structure Grammar (HPSG: Pollard and Sag 1994). Adaptation of this HPSG grammar has involved both extensions via so-called mal-rules (Schneider and McCoy 1998, Bender et al. 2004), as well as reductions in the grammar's coverage via masking to avoid unwanted ambiguity, given the restricted vocabulary made available to the students for the single-sentence exercises. While the present implementation has focused primarily on judging syntactic well-formedness, support is for identifying semantic errors, by testing the equivalence of the semantic representation that the grammar assigns to the student's sentence with that of a set of correct answers supplied for that

exercise. To accommodate systematic mismatches in this equivalence due to errors in the meaning of a student's sentence, the program employs both a paraphrase mechanism analogous to semantic transfer approaches to machine translation (Lønning et al. 2004), and a statistical semantic graph-matching algorithm.

The goal is to provide accurate and detailed instruction to students in response to each sentence that they write while taking these online courses. Fine-grained knowledge of English encoded in the ERG together with its mal-rule extensions, enables the program to identify characteristic properties of student sentences which are associated with error types, anchored in the type hierarchies of the lexicon and the syntactic constructions defined in the ERG. One of the central tensions confronted in this teaching application is in balancing the need for a consistent view of grammaticality with a more nuanced approach to good writing style. Since the ERG, like many linguistically-motivated grammar implementations, is designed to be descriptively comprehensive, adapting it to the service of prescriptive instruction is not a straightforward task, though an interesting one. A second continuing challenge is to maintain a good balance between the flexibility of the grammar and high accuracy in disambiguating to choose the one intended analysis for each student sentence, to make the right diagnosis of errors as the basis for explanations to the students.

The Redbird Language Arts curriculum and its precursor EPGY have been in regular use in elementary public schools in several states for several years now, including the Memphis public school system in which an efficacy study was conducted. During the two years of that study, from 2009-2011, weekly usage of the LA&W course ranged between 5,000 and 20,000 students in grades 2-6. The table here summarizes the number of sentences attempted by the students, with most of the usage occurring from January 2010 through May 2011, and with the answers divided into those evaluated as correct or incorrect, and further separated according to answers given on the first attempt and on the second attempt, when required.

**Table 1. Language Arts sentence composition exercises for Grades 2-6 in 2010-2011.**

	Correct			Incorrect			All
Total	1 <sup>st</sup> attempt	+	2 <sup>nd</sup> attempt	1 <sup>st</sup> attempt	+	2 <sup>nd</sup> attempt	
	1,446,253		190,729	709,489		509,995	
	1,636,982			1,219,484			2,856,466
	57.3%			42.7%			
Distinct	12,137			180,795			192,932
Unique	2,323			79,580			81,903
	19.1%			44.0%			

Note: The "Distinct" row gives the total number of different (non-duplicate) answers presented by students, while "Unique" gives the number of answers (word sequences) that appear only once in the whole set of student answers.

This table shows that students produced a correct answer on their first attempt in one of every two exercises (1,446,253, or 50.6% of the total), and improved to 57.3% correct after the second attempt. Interestingly, the 1.6 million correct answers consist of only 12,137 distinct answers (where duplicate answers are merged), while the 1.2 million incorrect answers consist of more than 180,000 distinct answers, showing much more variation than correct answers do. Similarly, while there were only 2,323 correct answers which appeared just once each in all of the corpus, students produced a much larger proportion of unique incorrect answers, more than 40% of the total number of distinct incorrect answers.

---

Of the nearly three million individual answers, many are repeated with high frequency, but even with the strong exercise-specific limits on words to choose from, many of the answers are unique or rarely repeated. The extent of variation is markedly different in correct and incorrect answers, with 6,588 distinct correct answers (roughly half) appearing fewer than 10 times, contrasted with 166,505 distinct incorrect answers with this low frequency (over 90%). This great variation in the range of distinct incorrect answers underscores the importance of using a linguistically sophisticated parser which automatically evaluates and diagnoses grammatical errors in these incorrect answers.

The length of the students' answers is normally under 10 words, with an overall average length of 5.45 words per distinct correct answer, and 5.70 words per incorrect answer. The exercises are designed to need no more than 20 words per answer, and normally no more than 10 words, with an average length of 6.77 words for the expected correct answers.

As students progress through the five grades offered at that time, the exercise-specific vocabulary lists grow in complexity, as do the sentences that the students are asked to compose. By Grade 6, these sentences include subordinate clauses, relative clauses, and conjoined noun phrases, verb phrases, and clauses. These recursive syntactic devices quickly lead to the wide variation summarized above for the particular sentences that students compose, and accurate analysis of errors in these sentences depends ever more crucially on the linguistically sophisticated grammar implementation used for these courses.

## Overview of the Curriculum Instructional Subject Areas

### The Language Subject Areas

Concepts addressed in the Parts of Speech subject area include the basic English word classes both for open vocabulary (nouns, verbs, adjectives, and adverbs) and for basic functional vocabulary (articles, helping verbs, prepositions, pronouns, demonstratives, conjunctions, etc.). Students get practice working with spelling, inflectional rules and their exceptions (irregular plurals for nouns, tensed forms for verbs, and comparative adjective forms), and learn to use productive word-formation rules such as adding the "-ly" suffix to form adverbs from adjectives. Lectures and exercises also teach the use of contractions and abbreviations, provide practice with often-confused word pairs such as "lie" and "lay", and present relations among words including synonymy and antonymy.

The Sentence Structure subject area addresses the syntactic rules governing the construction of well-formed phrases and sentences in Standard English, including types of sentences (statements, questions, and commands), phrase types, and how they are combined to form larger phrases and clauses. Lectures and exercises also teach simple and complex predicates, modification both in nominal phrases and in verb phrases, the proper use of punctuation and capitalization, and the construction and use of conjoined phrases.

In the Paragraph subject area, lectures and exercises address the basic types of paragraphs, their internal structure (lead sentence, supporting sentences, and concluding sentence), and concepts to guide the student both in understanding and in composing well-constructed paragraphs, including choice of a main idea, the use of detail, adoption of a point of view, and choice of style appropriate for the intended reader. Here, as in Sentence Composition, students are regularly presented with exercises that give them carefully guided practice in composing paragraphs of several varieties, complete with analysis and grammaticality feedback for every sentence they write. This scaffolded practice in writing whole paragraphs online, as discussed above, provides important preparation for students, most of whom will be required to produce paragraph-length answers in state-wide end-of-year school examinations.

## The Reading Subject Area

The Reading lessons of the course are designed to align closely with widely used ELA standards, while being integrated smoothly into the existing course with its equally strong emphasis on writing. In the Reading lessons, students are introduced to grade-appropriate short passages drawn from representative texts in prose and poetry, fiction and non-fiction, print and web, classic and contemporary, and historical and scientific.

Students are taught each concept or skill through a set of guided exercises centered on a text passage for each unit. While most instruction in this strand is on-screen, audio for the reading passages is included in order to smooth the path for new readers. Some sets of exercises are constructed around images to teach skills in interpreting information in multiple modalities.

### Efficacy of the Method

Stanford University researchers investigated the effectiveness of this composition-centered approach to teaching language arts, in a large-scale, two-year study for 2009-2011 of several thousand elementary public school students, with the results reported in a published refereed journal article (Suppes et al. 2014). In this study in the Memphis public schools, conducted in parallel with one of similar design for measuring efficacy in mathematics education, we recorded year-over-year student performance on annual Tennessee state test scores in Language Arts, and measured the correlation of these scores with the amount of positive work in the online courses during each of the two years. The study found that more positive work in the course correlated strongly and increased monotonically with improved state test scores. That is, the more effort a student invested in doing the exercises in the LA&W course during the school year, the greater the observed increase was in the score on the state exam compared to the previous year.

These test score gains, closely correlated with positive work in the course, are especially encouraging given that most of the students in the study were in Title 1 schools, where state test scores typically decrease from year to year. The strong statistical significance of these results was established by computing mean effect size as a function of the amount of positive work, comparing state test score changes for pairwise groupings of the students who used the LA&W courses, with the groupings determined by the amount of positive work completed.

An overview of some of the results of the Memphis study is shown in Table 2, with the four columns showing the test score ranges between 650 and 800 for the Tennessee annual state exam (TCAP), and the three rows showing the three broad groupings of students by amount of positive work (diff) done in the LA&W course during the two year period from 2009-11. In each cell of the table, the lower right number records how many students achieved a particular test score range and did some amount of work in the LA&W course. The upper left number in each cell shows the average change in state test score over the two-year period for that group of students.

**Table 2. Memphis Language Arts 2009-11**

500		18.6 90	6.7 94	
0	18.2 342	4.2 2006	-8.9 924	-31.8 62
-500	4.6 419	-14.5 855	-32.4 144	
<i>Diff</i> (09-11)	650	700	750	800
	TCAP 2009			

For example, the group of 342 students who completed up to 500 exercises of positive work (the measure of the number of correctly answered exercises minus the number of incorrectly answered ones) showed an average gain of 18.2 points on the state exam compared to their score in 2009, where the average state test score for this group of students in 2009 was between 650 and 700.

The numbers recorded in Table 2 also show that students who completed more positive work in the LA&W course achieved improved test scores, even when their initial (2009) test scores were higher. As noted above, a certain amount of positive work in the LA&W course is seen to be more immediately beneficial where the student's initial test scores are lower, but it is noteworthy that even the better prepared students benefit from positive work in the course to a degree that increases as the amount of work increases.

These efficacy results from a carefully designed, large-scale, multi-year study provide strong support for the benefits to students of sustained practice doing a wide variety of exercises with individualized instruction and feedback to develop mastery of the skills and concepts in English Language Arts. These benefits accrue especially quickly to students who enter the LA&W course with the least preparation, but mastery gains are also enjoyed by better-prepared students who are encouraged and motivated to invest more time in doing the coursework, at their own pace.

## Research and Development at Stanford

**M**cGraw-Hill Education continues to participate in and support a vigorous research program at Stanford in several departments, including work on enhancements to the Redbird Language Arts & Writing course offerings. Current research in this program at CSLI is focused in three areas, all aimed at improving the precision and breadth of error analysis in student writing. This work builds on the student composition data accumulated over the past nine years of the use of our grammar-based evaluation method, now comprising some ten million individual sentences and several hundred thousand paragraphs.

- **Paraphrase-based detection of semantic errors.**

The software provides evaluation and explanation of student errors in grammar, meaning, and style, with the strongest emphasis on grammar errors. Researchers are now expanding our analysis of semantic errors (where a sentence is syntactically correct but has the wrong meaning), through the

---

addition of a systematic treatment of paraphrase equivalence. They have identified a core inventory of congruence types which enable them to define a mapping from (the meaning of) a student's novel sentence to one of a small set of (nearly) equivalent correct or incorrect sentences for that exercise previously authored by the course developers. This work on formalizing and implementing a robust notion of paraphrase is a substantial and timely research challenge, one which draws on the wide range of expertise available at Stanford in the fields of semantics, logic, grammar engineering, and language processing, including generation of text from meaning. This rule-based method is complemented by a statistical approach that uses a semantic graph-matching algorithm to judge the correctness of meaning of a student sentence by measuring the distance between its associated semantic graph and those associated with the set of pre-supplied correct answers for each exercise.

- **Paragraph composition: Evaluating content and style.**

The program currently analyzes and advises on sentence-specific errors in the paragraphs that students write, without sufficient attention to the kinds of mistakes that involve multiple sentences in a paragraph. Research goals for improved paragraph-level evaluation include the proper use of pronouns and other anaphoric expressions, correct usage of articles including "a/an" and "the" for new and known entities, effective use of variation in sentence structure for better style, good use of dependent clauses (both adverbial and relative clauses), and sequencing of verb tenses. Each of these goals requires substantial linguistic research grounded in the student composition data as well as research and experimentation in effective user interface design to communicate our error analysis to students.

- **Evaluating short essays written with open vocabulary.**

Researchers are working to expand the scope of the Redbird LA&W curriculum beyond the current instruction in vocabulary-bound sentences and paragraphs, to include lectures and exercises teaching the composition of short essays. Research to support automatic evaluation of such essays includes the adaptation of methods for doing logical inference and knowledge representation, with some limits on complexity made possible through our control of the vocabulary available to students for each essay exercise. As with exercises in paragraph composition, the team will initially provide sentence-by-sentence evaluation for grammar, meaning, and style, while gathering writing samples in sufficient volume so as to choose useful opportunities for analysis and explanation at higher levels of text size.

## Summary

The Redbird Language Arts & Writing course is the result of more than a decade of focused design, authoring, and implementation. Ongoing Stanford research will continue to strengthen the distinctive features of MHE's Language Arts offerings, and efficacy studies using candidate enhancements will in-form the research. This interdependence of foundational research and efficacy studies will ensure that the courses continually evolve to address the instructional needs and goals of students, educators, and parents by utilizing research, best practices, and the newest advances in technology and personal-ized learning.

## References

Bender, Emily. M., Dan Flickinger, Stephan Oepen and Annemarie Walsh (2004). "Arboretum: Using a precision grammar for grammar checking in CALL," in Proceedings of the InSTIL/ICALL Symposium 2004, Venice, Italy.

---

Flickinger, Dan (2002). "On building a more efficient grammar by exploiting types," in Stephan Oepen, Dan Flickinger, Jun'ichi Tsujii and Hans Uszkoreit (eds.) Collaborative Language Engineering, Stanford: CSLI Publications, pp. 1-17.

Flickinger, Dan (2011). "Accuracy vs. Robustness in Grammar Engineering," in Emily M. Bender and Jennifer E. Arnold, eds. Language from a Cognitive Perspective: Grammar, Usage, and Processing, pp. 31-50. CSLI Publications, Stanford.

Lønning, Jan Tore, Stephan Oepen, Dorothee Beermann, Lars Hellan, John Carroll, Helge Dyvik, Dan Flickinger, Janne Bondi Johannessen, Paul Meurer, Torbjørn Nordgård, Victoria Rosén, and Erik Velldal (2004). "LOGON. A Norwegian MT Effort," in Proceedings of RASMAT:04, pp. 1-6. Uppsala, Sweden.

Pollard, Carl and Ivan A. Sag (1994). Head-driven Phrase Structure Grammar. The University of Chicago Press and CSLI Publications. Chicago, IL and Stanford, CA.

Schneider, David and Kathleen McCoy (1998). "Recognizing Syntactic Errors in the Writing of Second Language Learners," in Proceedings of Coling-ACL, pp. 1198-1204. Montreal.

Suppes, Patrick, Tie Liang, Elizabeth E. Macken, and Dan Flickinger (2014). "Positive technological and negative pre-test-score effects in a four-year assessment of low socioeconomic status K-8 student learning in computer-based Math and Language Arts courses," in Computers & Education 71.0, pp. 23-32.