Improving Child Literacy in Africa: Experiments with an Automated Reading Tutor

G. Ayorkor Mills-Tettey, Jack Mostow, M. Bernardine Dias, Tracy Morrison Sweet,
Sarah M. Belousov, M. Frederick Dias, Haijun Gong

***Abstract*—This paper describes a research endeavor aimed at exploring the role that technology can play in improving child literacy in developing communities. An initial pilot study and subsequent four-month-long controlled field study in Ghana investigated the viability and effectiveness of an automated reading tutor in helping urban children enhance their reading skills in English. In addition to quantitative data suggesting that automated tutoring can be useful for some children in this setting, these studies and an additional preliminary pilot study in Zambia yielded useful qualitative observations regarding the feasibility of applying technology solutions to the challenge of enhancing child literacy in developing communities. This paper presents the findings, observations and lessons learned from the field studies.**

***Index Terms*—Developing Regions, Education, Educational Technology, Literacy**

I. Introduction

Literacy is a key part of the global development agenda. It is a complex concept with multiple definitions. An international expert meeting at UNESCO in 2003 proposed the following definition: "Literacy is the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts. Literacy involves a continuum of learning in enabling individuals to achieve their goals, to develop their knowledge and potential, and to participate fully in their community and wider society." [22] The United Nations recognizes literacy as a human right, noting that basic education, of which literacy is the key learning tool, was recognized as a human right over 50 years ago in the Universal Declaration of Human Rights.

The research reported here was supported in part by the discretionary gifts to the TechBridgeWorld research group at Carnegie Mellon University, by the Qatar Foundation for Education, Science, and Community Development, by the National Science Foundation under ITR/IERI Grant No. REC-0326153, by the Institute of Education Sciences, U.S. Department of Education through Grant R305A080628 to Carnegie Mellon University, and by the Heinz Endowments. The opinions expressed are those of the authors and do not necessarily represent the views of any of our sponsors.

G. A. Mills-Tettey is with the Robotics Institute at Carnegie Mellon University, Pittsburgh, PA 15213 (phone: 1-412-268-8645; fax: 1-412-268-6436; email: ayorkor@cmu.edu). J. Mostow, M. B. Dias, S. Belousov, and M. F. Dias are also with the Robotics Institute at Carnegie Mellon University, Pittsburgh, PA 15213 (email: mostow@cs.cmu.edu, mbdias@ri.cmu.edu, sarahtbw@ri.cmu.edu, mfdias@ri.cmu.edu). T. M Sweet and H. Gong are with the Statistics Department at Carnegie Mellon University, Pittsburgh, PA 15213 (email: tsweet@andrew.cmu.edu, haijung@andrew.cmu.edu)

As a contribution to the discourse on applying information and communication technologies (ICTs) to address development challenges, we describe a study exploring the potential role of computing technology in improving English literacy among Ghanaian and Zambian children who attend school in English but have low reading achievement levels.

Low functional literacy among individuals who have completed primary school is not an uncommon problem in developing communities. For example, UNESCO reports that in 2000, more than 1 in 3 adults with a fifth-grade education in Chad and Niger reported that they could not read [21]. In other cases, individuals may finish primary school reading below the expected level. In a representative sample of Ghanaian public schools, reading achievement levels measured by the government-administered Criterion Referenced Test in 2000 indicated that fewer than 10% of the children in grade six were able to read with grade level mastery [9].

Several factors contribute to this problem. For the average child from a rural or low-income urban background in Africa, reading is not part of daily family life, and sometimes parents are not themselves literate. In Ghana and Zambia specifically, most children speak one of a number of local languages at home but attend a school taught in English, the official language for both countries. Typically, under-resourced schools with overcrowded classrooms offer few opportunities for individual attention while developing reading skills.

The project described in this paper is a proof-of-concept study to investigate whether an automated computer-based reading tutor that provides guided reading practice can significantly improve the reading proficiency of children in a developing community, even if they have no prior familiarity with computers. We focus our study on children in Accra, Ghana, and Mongu, Zambia.

As illustrated in Figure 1, we began by employing the Reading Tutor in a preliminary three-week-long pilot study in Accra, Ghana. The pilot study was used to explore technical and operational feasibility and to motivate partnerships and funding for a longer term study. We followed up with a four-month-long controlled study in Accra, in collaboration with the Ghana-India Kofi Annan Centre for Excellence in ICT (AITI-KACE) and with input from Associates for Change, an educational research firm in Accra. The controlled study

aimed at quantitatively measuring the educational effectiveness of the tutor with children in Accra. Observations and lessons from these experiences fed into another pilot study in Mongu, Zambia, with the plan of following up with an additional controlled study in Zambia in the future.

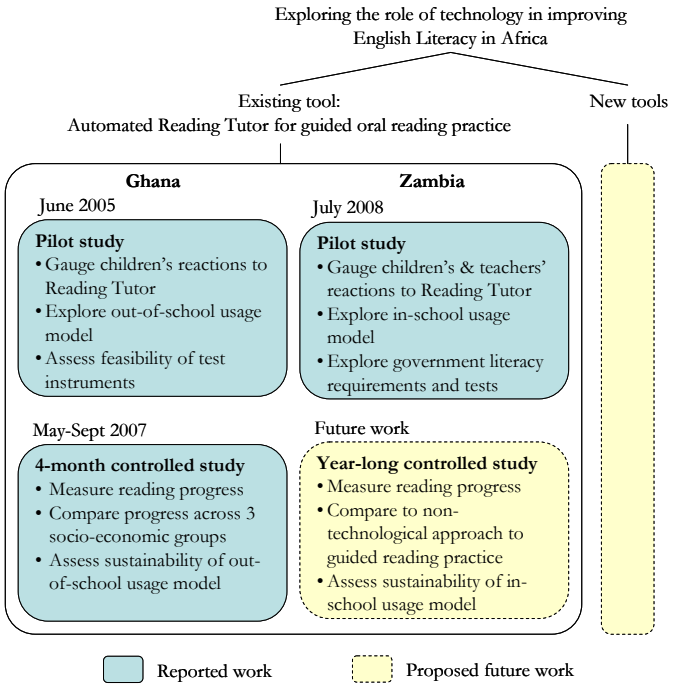

Figure 1 - Project outline

Section II describes the skills that are involved in reading and introduces the automated tutor we used in this project. Sections III, IV, and V describe the pilot field study in Ghana, the 4-month-long controlled study in Ghana, and the pilot study in Zambia respectively. Section VI distills practical lessons learned in the implementation of these field studies and Section VII concludes with a discussion of future work.

II. Background

As a cognitive proficiency, reading involves several component skills such as phonemic awareness, decoding, fluency, vocabulary, and comprehension [14]. Phonemic awareness is the ability to perceive individual sounds or *phonemes* in words. Building on this awareness, a child learns the alphabetic principle that spelling generally maps systematically to pronunciation. The child also learns specific letter-sound correspondences and the ability to correctly pronounce written words (decoding). Fluency is the ability to read text accurately, quickly, and expressively and is an important foundation for comprehension. A rich vocabulary is essential for comprehension and effective communication. Children develop these skills through a variety of experiences, including skilled instruction. Research has shown that regular guided oral reading plays an important role in developing reading skills, particularly fluency and comprehension [14].

Such guided oral reading may happen in small groups in a classroom setting, or with parents at home. Technology may also assist in this process: Scientific Learning's *Reading Assistant* [1] and Project LISTEN's *Reading Tutor* [11] are examples of computer-based tools that use automated speech recognition to provide a guided reading experience for the user. Project LISTEN's Reading Tutor has demonstrated success in improving the reading ability both of children whose first language is English [11][12][13] and of children learning English as a second language (ESL) in the United States [16] and in Canada [18].

The project described in this paper used Project LISTEN's Reading Tutor, developed at Carnegie Mellon University. The Reading Tutor displays stories on a screen and "listens" to a child read aloud. By using speech recognition to analyze the child's reading, the Reading Tutor is able to give graphical and spoken feedback. It gives help when it detects a long pause, a severely misread word or a skipped word, and also when the reader clicks for help. The tutor may speak the whole word out loud or decompose the word into syllables or phonemes and speak out each part while highlighting it. It may also give a "rhymes with" hint, or read the sentence by playing a fluent human narration to model expressive reading. The Reading Tutor includes a wide variety of stories at different reading levels. It takes turns with the child in selecting a story to read. It monitors the child's reading progress and selects stories at an appropriate level for the child. For readers at early stages of development, the Reading Tutor also includes word-building exercises to develop knowledge of spelling-to-sound correspondences. Videos of the Reading Tutor in use may be found at Project LISTEN's website [17].

III. Ghana Pilot Study

A. Goal

The 2005 Ghana pilot study aimed to evaluate the practicality of a technological approach to guided reading practice in Accra, and investigate the feasibility of conducting a longer term controlled study.

Specifically, we wished to answer the following questions:

1. *Partners and logistics:* How feasible is it to engage partners and arrange the logistics for such a study?
2. *Learning to use the Reading Tutor:* How quickly do children with no prior computer experience learn to operate the Reading Tutor, and what instruction do they need in order to do so?
3. *Speech*: Does the speech recognition software perform acceptably with Ghanaian accents, and can the students understand the narrated speech?
4. *Tutor content:* Do the children find the reading material in the tutor engaging?
5. *Usage sessions:* What is an effective length for a tutoring session?
6. *Reading measures:* Which test instruments can be used easily and effectively to assess reading proficiency in the Ghanaian setting?

B. Participants and Methodology

We chose to focus on the needs of children from low-income families attending public school because they have a high risk of low achievement in reading. In consultation with Associates for Change, we targeted children in grades two through four since this is a key period for developing reading skills after the initial adjustment to the primary school environment. We restricted our work to an urban environment where computers are more readily available, but where significant literacy challenges still exist. We employed an out-of-school usage model in which practice with the Reading Tutor was supplemental to regular school activities.

The pilot study involved qualitative observations of children as they used the Reading Tutor. The study was conducted by the first author who is a native of Ghana. She was assisted by a local volunteer. Two groups of children participated in this study. One group comprised twelve children in grades two through four from an under-resourced public school. They used the Reading Tutor at an internet café near their school for 20-30 minutes each day over a three-week period. The other group comprised six children from a mixed low- and middle-income neighborhood. They used the tutor on laptops in the home of the researcher, for 20-30 minutes each day, three days a week over the same period.

C. Results and Observations

Partners and Logistics: The school, internet café, and the parents of all participating children agreed without hesitation to participate in the project, and logistics were arranged very quickly. The internet café donated time on four desktop computers; thus, four children could use the Reading Tutor at a time. We engaged the remaining children in other activities while they waited their turn. This worked fairly well for the pilot project, but would not be a feasible approach for a longer study involving a larger number of children.

Learning to use the Reading Tutor: Although one child had previously played a computer game, none of the other children had used a computer before. The children were given an initial 10-15 minute hands-on lesson introducing them to the computer and showing them how to use the mouse, keyboard and software. The children had trouble understanding the Reading Tutor's built-in automated tutorials on how to operate it and use the keyboard, perhaps due to the unfamiliar narration accent or the use of words such as "roster" that are common in American but not Ghanaian English. This difference in language use is an example of the need to localize tools for a given setting. We explained the tutorials one-on-one to the children during their first two sessions. Some of the questions the children asked during the first few sessions were on keyboard use and on navigating the tutor. By their second or third session, most of the children were able to operate the Reading Tutor without help.

Speech: The speech recognition capability appeared to work adequately with the children's accents: it accepted a great majority of words that were read correctly. Also, based on the graphical feedback given by the tutor, the children quickly learned to recognize when it did not "hear" them correctly and to repeat themselves when necessary. After the initial difficulty with the built-in tutorials, the children appeared to understand the prompts given by the Reading Tutor.

Tutor content: Overall, the children were very enthusiastic about using the Reading Tutor. Most of them appeared to enjoy the world-building exercises and the stories in the tutor. However, we noticed that a couple of children who were older than the norm (11 or 12 years old) but had a kindergarten or first grade reading level were not fully engaged by the content of the simple stories, such as "Sam sat on a mat," available for their reading level. There was a discrepancy between these children's maturity and their reading ability. Furthermore, we believe it would be good to incorporate more local content, such as Ghanaian folk tales, into the tutor.

Usage sessions: We noticed that the better readers could use the tutor for more than half an hour at a time without getting bored, whereas the more challenged readers would tire after about twenty minutes, but would still look forward to their next turn. We thus decided to limit the length of usage sessions in the controlled study to 20-30 minutes.

Reading measures: Although the pilot study did not aim to quantitatively measure reading progress, we identified and tested reading measures that we would use in the controlled study. The instruments we chose to use to assess reading ability were an oral reading fluency test [5] and the Test of Written Spelling (TWS) [8], both of which are hand-scored tests that have been frequently used in previous studies involving the Reading Tutor, and are psychometrically reliable, fast and easy to administer. The fluency test is timed reading exercise in which the child is given a grade-appropriate story to read. It is scored as the number of words read correctly in one minute. Fluency is essential to comprehension and is a sensitive measure of growth in proficiency [19]. The TWS is a dictation exercise in which the child writes down words read aloud by the tester in order of increasing difficulty. The test ends once the child has incorrectly spelled five words in a row, at which point it is assumed that the difficulty of the remaining words exceeds the child's proficiency. This allows the same test to be used for multiple grade levels. TWS is scored as the number of words spelled correctly. Spelling is like decoding in that it tests print-to-sound mappings but it is easier to assess reliably.

We tried out the fluency test and the TWS with the pilot study participants, and they did not have trouble with the format of either test. To combat their nervousness about being examined, we avoided using the term "test." We instead explained that these exercises were to help figure out how the Reading Tutor could assist them and we reassured the children that they would not be given a grade. We modified the passages for the fluency test to use Ghanaian names rather than American names, to make them more recognizable to the students. The children were familiar with the dictation format of the TWS, but again were unfamiliar with some words such as "tardy" which are not regularly used in Ghanaian English.

Figure 2 - A child in Accra reading a story using the Reading Tutor

IV. Ghana Controlled Study

A. Goal

Following on the successful pilot study, the goal of the controlled study in Ghana was to quantitatively measure the efficacy of the Reading Tutor in helping children improve their reading skills. Specifically, we wished to determine:

1. Does regular use of the Reading Tutor improve oral reading fluency and spelling?
2. Do treatment effects depend on other factors such as school/socio-economic background, gender, or grade level?

In addition, we wanted to learn about the operational sustainability of an out-of-school usage model, taking into consideration installation and maintenance of the software, training of staff responsible for the day-to-day running of the project, transportation of the children between the school and the project site, and other logistics. The controlled study was deployed by project staff at AITI-KACE, with remote training and support by our team in Pittsburgh.

B. Study Design / Methodology

The controlled study involved eighty-nine children from three schools, representing three socio-economic backgrounds. The participating schools, recruited by Associates for Change, were S1: a private school in a middle-income community, S2: a public school in a low-income community, and S3: an informal educational program for highly disadvantaged children who have never attended formal school. The study involved children in grades two through four of S1 and S2, and in the "Intermediate" and "Advanced" levels (roughly corresponding to grades two and three respectively) of S3.

The children were split randomly across school, grade and gender boundaries into two groups, Tutor-1st and Control-1st, as shown in Table I. We used a two-treatment crossover study design: for the first half of the study, the Tutor-1st group used the Reading Tutor while the Control-1st group had no additional reading intervention, and for the second half, the roles were switched: the Control-1st group used the Reading Tutor while the Tutor-1st group had no additional reading intervention. Each half of the study lasted nine weeks (two months) during which there were daily usage sessions of approximately half an hour per child, although attendance and usage varied considerably. This crossover study design, illustrated in Table II, had the advantage of equity in the sense that all participating children had the opportunity to use the reading tutor. It also avoided having to pair similar children for comparison since each child essentially served as his or her own control. The potential disadvantage, which was that carryover treatment effects from the first half may have affected the second half, was mitigated by having half the children participate first in the control, while the others participated first in the treatment group.

Table I - Study participants by school, grade and gender

School	Grade	Number of Children			
		Tutor-1st		Control-1st	
		Female	Male	Female	Male
S1 (29 children)	Grade 2	3	2	2	3
	Grade 3	5	2	1	3
	Grade 4	1	3	3	1
S2 (30 children)	Grade 2	2	3	3	2
	Grade 3	3	2	1	4
	Grade 4	3	2	3	2
S3 (30 children)	Intermediate	3	4	3	5
	Advanced	3	5	5	2

A battery of three fluency tests and one TWS were administered to all the children at the beginning of the study (*pre-testing*), between the two halves (*mid-way testing*), and at the end (*post-testing*). The testing was conducted at the children's schools by AITI-KACE project staff. Three fluency tests were used to better estimate reading ability. The passages used for these tests correspond roughly to first, second, and third grade reading levels respectively, since a test close to the child's current reading level should give a more sensitive measure of progress. We found that the scores on the three passages were highly correlated, and so for analysis purposes these scores were combined into a single mean fluency score.

Table II – Cross-over Study Design

	Tutor-1st Group	Control-1st Group
	Pre-testing of all children	
First half of cross-over (1st nine weeks)	Reading Tutor	Control: no special intervention
	Mid-way testing of all children	
Second half of cross-over (2nd nine weeks)	Control: no special intervention	Reading tutor
	Post-testing of all children	

The study participants used the Reading Tutor in a computer lab at AITI-KACE because it had equipment and technical capability that their schools did not. The computers used for the project had 2.4GHz Pentium IV processors, 256MB of RAM, and 16GB hard drives.

C. Deviation from Study Design

In a normal crossover study, experimental conditions are held constant over the course of the study. However, due to logistical challenges that delayed the start of the project, the first half of this study took place while the students were attending school, whereas the second half overlapped with the school vacation. Because of the different conditions in place during the two halves of the study, we analyzed the results as two different experiments rather than as a single crossover study. The first experiment measured the effect of the Reading Tutor while the children attended school (and thus attended English class as usual). The second experiment measured the effect of the tutor while the children did not attend school (and so did not attend English class). This modified study design is illustrated in Table III.

Table III – Modified Study Design

	Tutor-1st Group	**Control-1st Group**
	Pre-testing of all children	
Experiment 1 (1st nine weeks)	School + Reading Tutor	Control: school only
	Mid-way testing of all children	
Experiment 2 (2nd nine weeks)	Control: No school	No school + Reading tutor
	Post-testing of all children	

D. Results and Analysis

Figures 3 and 4 illustrate the pre-test scores of children at the three schools, representing their reading proficiency going into the study. It is clear that the S1 children had much higher levels of reading achievement than the S2 children who in turn had higher levels than the S3 children. Pre-testing for our study occurred about two months before the end of the school year. To provide some context for these scores, Table IV compares the fluency pre-test scores with end-of-year norms from schools in the United States [6]. By this standard, the average reading proficiency of the S1 students appears to be at or above the U.S. average whereas that of the S2 and S3 children is lower. Norms from schools in Ghana are not available for comparison.

In analyzing each experiment, we focus on gains in fluency and spelling test scores. For experiment 1, the gain is the difference between pre- and mid-test scores; for experiment 2, the gain is the difference between the mid- and post-test scores. A positive gain indicates improvement in the child's reading proficiency. In each experiment, we use a standard statistical *t-test* to compare the mean gains of the treatment and the control group. This test yields a *p*-value indicating how significant the difference is between the means of the two groups. For this analysis we consider *p*-values of less than 0.016 to indicate statistical significance. This value is smaller and thus more conservative than the commonly used threshold of 0.05 because multiple comparisons (due to the different schools) require an adjustment of significance levels [10]. We also compute a measure of effect size, that is, the magnitude of the treatment effect. The measure we use for effect size is Cohen's *d*: the difference between group mean gains divided by the within-group pooled standard deviation [3]. The effect size is computed for each school because the treatment effect depends on the school. An effect size of 0.2 is generally considered small, 0.5 medium, and 0.8 large [4].

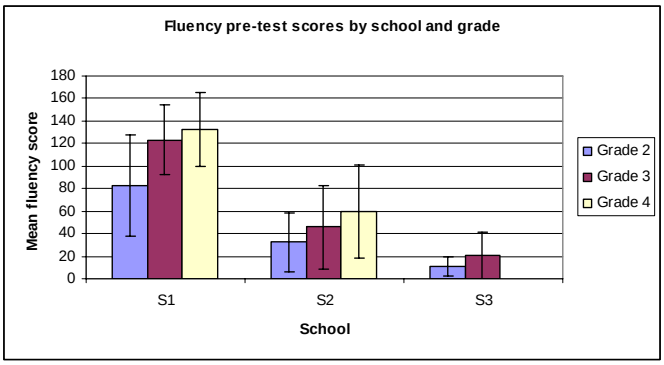

Figure 3 - Fluency pre-test scores by school and grade

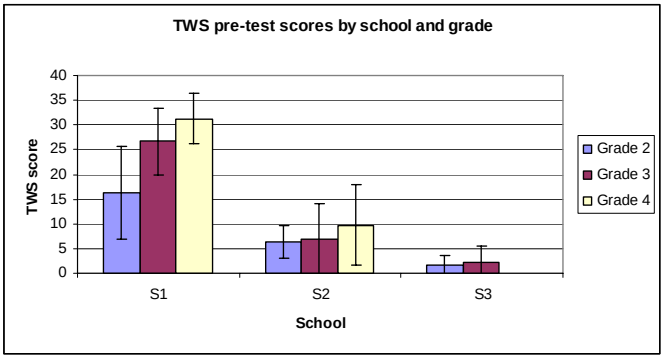

Figure 4 - TWS pre-test scores by school and grade

Table IV - Fluency pre-test scores compared with U.S. norms

Grade / Level	**Mean pre-test fluency (standard deviation)**			**Year-end U.S. Norm [8] (50th percentile)**
	S1	**S2**	**S3**	
2 / Int.	82.4 (44.8)	32.3 (25.9)	11.1 (8.5)	89
3 / Adv	123.3 (31.4)	45.8 (37.4)	20.4 (20.8)	107
4	132.7 (32.7)	59.8 (41.5)	--	123

Experiment 1 results by school: Table V shows the results and *t*-test analysis for experiment 1, when the children were attending school. The reading proficiency of the S2 children who used the Reading Tutor (the "treatment" group) improved significantly more than those who did not (the "control" group), as evidenced by larger gains in both fluency and TWS. The S3 treatment group significantly out-gained the control group in fluency but not in spelling, although there was a positive trend. Finally, there was no significant difference in gains between the S1 treatment and control groups, for either test. Omitted from this analysis are five S1 children who were not present for the midway testing. Due to the small numbers, we did not break down the data to examine gains for each grade level within each school.

Table V - Comparison of treatment and control gains for experiment 1

Test	School	Treatment (Tutor-1st group, N=41)	Control (Control-1st group, N=42)	*t*-test comparison of treatment & control gains	
		Mean Gain (SD)		p-value	Effect size
Fluency	S1	8.6 (20.5)	21.5 (11.7)	0.0833	-0.775
(*# words read correctly per min*)	**S2**	**61.4 (22.9)**	**23.4 (21.0)**	**<0.0001**	**1.731**
	S3	**13.4 (9.3)**	**5.1 (4.4)**	**0.0054**	**1.144**
TWS	S1	-1.9 (4.9)	-0.2 (2.6)	0.2997	-0.457
(*# words spelt correctly*)	**S2**	**3.9 (2.7)**	**0.3 (3.4)**	**0.0039**	**1.153**
	S3	2.8 (4.0)	0.6 (1.7)	0.0648	0.719

Experiment 2 results by school: Table VI shows the results for experiment 2, when the children were not attending school. Interestingly, there were negative fluency gains for the S2 children over the vacation, and these were more dramatic for the control group than for the treatment group. We discuss this observation in the *Discussion* subsection that follows. For the S2 children, the difference in TWS test score gains between the treatment and the control group was not statistically significant. Finally, there was no significant difference between the treatment and the control group in the other two schools for either mean fluency or TWS. Omitted from the analysis are two S1, two S2 and ten S3 children who were absent for testing.

Table VI - Comparison of treatment and control gains for experiment 2

Test	School	Treatment (Control-1st group, N=33)	Control (Tutor-1st group, N=37)	*t*-test comparison of treatment & control gains	
		Mean Gain (SD)		p-value	Effect size
Fluency	S1	0.3 (13.4)	-3.8 (22.6)	0.613	0.222
(*# words read correctly per min*)	**S2**	**-9.1 (17.1)**	**-39.8 (16.8)**	**<0.0001**	**1.816**
	S3	4.8 (6.0)	17.8 (27.0)	0.1321	-0.665
TWS	S1	2.0 (4.1)	2.9 (5.1)	0.6431	-0.197
(*# words spelt correctly*)	**S2**	0.4 (2.0)	0.6 (3.0)	0.858	-0.068
	S3	0.8 (2.4)	-0.3 (2.2)	0.3198	0.474

Effect of Gender: In addition to exploring how test score gains varied by school, we examined the data for evidence that test score gains were affected by gender. We focused this analysis on experiment 1 since that data was more expressive. A *t*-test comparison of the mean gains of females to that of males in each experiment resulted in large *p*-values (> 0.4), indicating that the mean gains of females and those of males were not significantly different from each other. An Analysis of Variance (ANOVA) also confirmed that gender did not have a significant effect on TWS or fluency gains in experiment 1. Details of these analyses are omitted for brevity.

E. Discussion

Effectiveness of the Reading Tutor: The results provide evidence that during the school term, the S2 students (and to a lesser extent the S3 students) who used the tutor gained considerably more than those who did not use the tutor. Thus, the Reading Tutor was helpful for the S2 children. This is a positive outcome since the S2 group most closely represents our target population of children from low-income families attending public school. The proficiency of the S1 children did not appear to be influenced by reading practice with the tutor. A possible explanation for this is that the S1 children might not have had much room to benefit from the tutor since they were fluent readers going into the study.

Loss in proficiency over vacation period: Of the statistically significant results highlighted in Tables VI and VII, the negative gains of the S2 children in experiment 2 stand out. A possible explanation for this might be found in studies in the United States which have documented that reading achievement test scores for children from low-income families deteriorate significantly over the summer vacation (a phenomenon referred to as the "summer reading setback") whereas those for children from middle-income families remain steady or increase slightly [2]. This trend has been attributed to the discrepancy in the reading opportunities and materials available to these two groups over the vacation period. In light of these studies, it is interesting to note that for the S2 children, those who used the Reading Tutor did not deteriorate in reading ability as much as those who did not use it. However, the sessions with the Reading Tutor (totaling, on average, 11.5 hours of reading per child), without English class at school, were not enough to prevent negative gains.

Effect of pre-test scores: A complicating factor in the data analysis is the observed unequal average pre-test scores of the two groups, Tutor-1st and Control-1st, despite supposed random assignment of children to the two groups. Table VII shows that this disparity is statistically significant for the S2 children. Higher pre-test scores can be the cause of greater gains [20], and so to determine whether the greater gains of the Tutor-1st group in experiment 1 were due to the Reading Tutor or the higher pre-test scores, we computed the correlation between pre-test scores and gains. We found no significant positive correlation between pre-test scores and gains: the correlation coefficient of mean fluency gains vs. mean fluency pre-test scores was –0.301 (a small negative correlation), and the correlation coefficient of TWS gains to TWS pre-test scores was 0.005 (no correlation). This suggests that the greater gains of the Tutor-1st group in experiment 1 are indeed attributable to the Reading Tutor rather than to their higher pre-test scores.

Table VII - Comparison of pre-test scores for the Tutor-1st group and the Control-1st group

Test	School	Tutor-1st (N=41) Mean pre-test score	Control-1st (N=42) Mean pre-test score	p-value
Fluency (*# words read correctly per min*)	S1	122.0 (45.9)	99.2 (33.9)	0.137
	S2	**65.8 (33.9)**	**26.2 (26.9)**	**0.0015**
	S3	18.6 (13.2)	12.1 (19.1)	0.308
TWS (*# words spelt correctly*)	S1	24.7 (9.7)	24.0 (9.6)	0.850
	S2	**11.1 (7.0)**	**4.3 (3.7)**	**0.003**
	S3	2.1 (2.2)	1.9 (3.1)	0.840

Tutor usage: Another significant difference between the two experiments is that tutor usage in experiment 2 was much lower than in experiment 1, due to a higher level of absenteeism from the study during the school vacation. As Table VIII shows, this is particularly true for participation of the S3 students whose total usage per student in the second half of the study dropped to under a third of what it was in the first half. This reduced usage may partly explain why there is no statistically significant effect of the Reading Tutor for S3 children in experiment 2. Nineteen children who were absent for testing either at the beginning or end of each experiment are not included in this summary.

Table VIII - Tutor usage in each experiment

	School	Experiment 1	Experiment 2
Number of days of tutor use per student	S1	29.6 (12.0)	22.2 (9.4)
	S2	36.9 (2.4)	24.0 (11.7)
	S3	30.7 (2.8)	12.6 (3.9)
Total time spent reading with tutor per student (hours)	S1	12.4 (5.4)	11.9 (5.2)
	S2	18.6 (3.1)	11.5 (6.7)
	S3	17.3 (2.7)	6.8 (2.5)
Average daily time spent reading with tutor per student (hours)	S1	22.9 (7.1)	32.6 (5.3)
	S2	30.1 (3.7)	26.6 (7.0)
	S3	33.9 (4.6)	31.9 (5.3)

V. Zambia Pilot Study

The work in Zambia complements the prior studies in Ghana by investigating an in-school usage model and testing the tutor in a different English-speaking African country. ProjectEDUCATE, a non-profit organization supporting some schools in Zambia, introduced us to two under-resourced public schools in Mongu, the capital of Zambia's Western Province. The schools had received donated computers and were enthusiastic about the possibility of testing the Reading Tutor. However, the Mongu District Education Board Secretary's office directed us to select only one school for the pilot test. Three of the authors conducted the study in Zambia.

A. Goal

The goals of the Zambian pilot study were to answer the following questions:

1. *Computing infrastructure*: What is the state of the school's computing lab, how is it currently used, and can it feasibly be used for sessions with a computer-based reading tutor?
2. *Training Teachers*: How long does it take to train teachers to guide students in the use of the tutor?
3. *Children's response*: How do the children respond to the Reading Tutor?
4. *Test instruments and Reading Tutor content*: What test instruments and Reading Tutor content are appropriate for the Zambian setting?
5. *Feasibility of long term study*: Would a remote partnership between the school and our research group be a feasible model for a longer-term controlled study?

B. Methodology and Implementation

The selection of the school for the pilot test was done after meeting with the headmasters and teachers at both schools and assessing the state of the computer labs and the potential for a successful study. Subsequently, the pilot study consisted of conducting interviews with teachers at the selected school, providing training for the teachers, and making qualitative observations of teachers and students as they used the Reading Tutor. We also explored with the teachers possible details of a longer controlled study.

C. Results and Observations

Computing infrastructure: Both schools had a computer room with about 20 computers, most of which were 266 Hz Pentium II machines with 64-128MB of RAM and 10GB hard drives. Some teachers from each school had taken a computer skills training course when the computers were initially donated, but had not had the opportunity to reinforce their skills through additional training or guided practice. The schools did not have internet access and the computers were being used primarily for tasks such as typing exam questions, rather than as educational tools for students or for accessing online teaching resources. Given the large class sizes—which ranged from 50 to over 100 students—and strict timetable, the teachers also faced challenges in feasibly using the one lab to teach their students from grades one to nine about computers. Some computers at both schools were not functioning due to broken keyboards, mice, and power strips; some computers were not being protected from dirt during the dusty winter months; power outages were a daily occurrence in Mongu; and for one school, even maintaining electricity for the computer lab was a challenge due to limited financial resources. For the pilot study, we selected the school with better maintained equipment and a higher likelihood of maintaining communication by telephone and email since the school had a telephone and the headmaster had a working email address that he accessed weekly.

Training Teachers: We trained three teachers to use the Reading Tutor. The hour-long session covered the educational features of the tutor as well as administrative tasks such as managing users. After basic instruction and some time to practice on their own, the teachers were able to guide students in using the tutor. They would often give feedback to the children as a complement to the tutor when the students had difficulty reading stories. For example, they would instruct the

children to click for help when they needed it. The teachers would sometimes correct a mispronounced word that the tutor did not detect. We think this involvement of teachers in the early stages as the child gets used to the tutor could be an important part of longer-term use of the Reading Tutor in this school setting. A similar role was played by project staff in the Ghana studies. The teachers especially appreciated the ability to track their students' performance using the tutor, including the number of new words the children had read and the number of words they read per minute.

Children's response: We observed one group of eleven children in grades two through four as they read one or two stories from the tutor. The students were selected by the teachers and had varying levels of reading ability, English comprehension, and speaking fluency. As in Ghana, all but one student were completely new to computers. We introduced them as a group to the basic components of the computer and then gave them the opportunity to demonstrate use of the mouse and keyboard to each other. We provided verbal instruction to students individually as they began reading stories from the Reading Tutor. Just as in Ghana, the speech recognition capability appeared to work adequately with the Zambian students' and teachers' accents. Some Grade 2 students seemed to have difficulty understanding English. This is because English is introduced as a language of instruction only in Grade 2, as a part of the Zambian Ministry of Education's initiative to encourage basic literacy by teaching in a familiar language in the first year of school [21]. The students enjoyed using the Reading Tutor; when we returned to the school on another day to meet with the teachers, the group of students was using the word building exercises in the Reading Tutor on their own time.

Figure 3 –Teachers and a child test the Reading Tutor in Mongu

Figure 4 – Students in Mongu using word-building exercises as a group

Test instruments and Reading Tutor content: We discussed with the teachers other reading assessment options for a longer term study, in addition to the fluency test and Test of Written Spelling. We learned that through the Zambia Primary Reading Programme (PRP), the students' literacy levels are measured by reading standardized story books aloud to their teachers. Each color-coded book is associated with a given reading level. Once the teacher determines that the child can successfully read at a given level, he/she moves the child up to a higher level reading group with a different set of books. The teachers were interested in engaging the students with the computers in a way that would support curricular requirements from the Ministry of Education, and this was also emphasized to us by the Ministry of Education officials. Accordingly, we discussed incorporating the standardized PRP reading material into the tutor, which will also have the advantage of providing additional content choices for students. We saw that many of the existing Reading Tutor stories, such as those related to baseball, zoos, or recycling, had little relevance to children's lives in Mongu.

Feasibility of long term study: The teachers suggested that a random subset of the children in grades 2 through 4 could be selected to participate in a controlled study. They explained that students in these grades have 60 minutes of reading class per day and suggested that the intervention group of students could spend 30 minutes of that time working with the Reading Tutor, with teacher supervision, while the control group of students would remain in class. We anticipate several challenges in conducting this project as a remote field study. Although we have email and telephone contact information for the key collaborators at the school, we expect communication to be difficult given their limited internet access and the frequent disruptions in telephone communications due to daily power cuts. Success would depend on the teachers taking ownership of the project as a result of their enthusiasm. We hope that communication with local contacts will help address some of the expected challenges. For example, we plan to involve the Mongu District Education Board Secretary's office in evaluating the project's progress at the school. We also engaged a local ProjectEDUCATE technical contact who joined us at our meetings with the schools, received training from a member of our team, and assisted us with installation and use of the Reading Tutor software. We hope that he will stay in touch with us remotely, help incorporate new stories into the tutor, check in periodically with the school, and be available to the teachers to assist with any technical challenges.

VI. Lessons Learned

From the experiences in Ghana and Zambia, we can glean many useful lessons for our future work as well as for others implementing similar studies. These can be broadly categorized into lessons about building relationships and support, lessons about running the study, and lessons about the viability of automated tutoring.

A. Building Relationships

An important requirement for running these studies was the process of developing partnerships at each stage in the project. Significant time must be devoted to building relationships and developing a shared vision. Phone conferences can be a useful tool, but some face-to-face meetings are essential, particularly in the early stages of project planning. A staged implementation, as in a pilot study followed by a controlled study, helps distill key questions to ask, refine design and implementation decisions, and identify potential problems.

B. Logistics

We learned that it was important for someone involved in the local day-to-day running of the project to have significant decision-making power and to feel ownership of the project. This individual must have the ability, for example, to replace equipment or interface with representatives of the school administration. Otherwise, problems can easily stall progress. It is also essential to have access to some local technical expertise to troubleshoot and repair problems—building this local support base during the pilot studies is crucial.

For the controlled study in Ghana, we found that the remote collaboration required regular, sometimes daily, communication among the project partners; email, instant-messaging and voice-over-IP were a cost effective and feasible means of achieving this.

Not surprisingly, we learned that a process that requires a significant change in behavior and extra work on the part of parents, such as having the children come to school or participate in a study during a vacation, is hard to sustain and should be avoided if possible. However, it was also clear that unexpected situations are bound to arise for any study that involves cross-continent collaboration and so flexibility and the ability to adjust the study design if necessary is essential. For example, in the controlled study in Ghana, the second half of the study had to be held during the school vacation although this was not the original plan.

In running the Ghana studies, transportation of the children was the greatest expense and so although scheduled out-of-school use in an ICT center was the model chosen for the controlled study, it is not a viable long-term usage model. However, as we learned in Zambia, there are many challenges to be addressed for an in-school usage model for public under-resourced schools, even in a school that already has a computer lab. These issues include the limited availability of computers, inadequate experience on the part of the school in maintaining computer labs, scheduling challenges with respect to the school timetable, and also the need to work within the constraints of the existing literacy curriculum.

C. Viability of Automated Tutoring

A key lesson was that even without prior computer experience, the participating children were quickly able to acquire the skills needed to use the tutor. In general, they were excited about using the computer and about the interactive features of the Reading Tutor. Assistance from the project staff (in Ghana) or from the teachers (in Zambia) helped those children who were initially nervous about computer use or about reading in English. In both locations, the children's natural curiosity overcame any initial apprehension of the unfamiliar technology.

In observing the children using the Reading Tutor, we noticed that children with a basic foundation in English and not much prior experience with computers, such as the S2 children in Ghana, were easily engaged with the tutor. It was clear that insufficient familiarity with the English language was a challenge for some, particularly the S3 children in Ghana and some of the Grade 2 students in Zambia. For children who did not yet understand English well enough to benefit from an automated tutor that uses only English, a tool that bridges between the local language and English, e.g. by giving prompts or explaining words in the local language, might have been better. Finally, we noticed that the S1 children who were already fluent readers and experienced with computers seemed to get bored and distracted easily when using the Reading Tutor. This might have been because the Reading Tutor displays the story being read one sentence at a time, and there would sometimes be a short delay in loading the next sentence. For a fluent reader, even this short delay was noticeable and sometimes frustrating. There were also instances that indicated that some children got bored of repeated activities and enjoyed variety. For example, many children enjoyed the word-building exercises in the Reading Tutor but would sometimes complain that there were too many of them.

Although the Reading Tutor is designed for use by a single user at a time, we noticed that in both pilot studies children would gather around a single computer and try to help each other. This also happened to a lesser extent during the controlled study in Ghana and has been observed in other technology interventions in developing communities [15]. It suggests that we should investigate multi-user scenarios.

Finally, there are several usability improvements that can be made to the Reading Tutor to reduce the required technical support, particularly in a developing community setting. These features would apply to any PC-based intervention.

- Installation must be easy and straightforward.
- Customization must be easy (both for adding local content and to control for the installation footprint for machines without much hard disk capacity).
- An easy administrative interface is needed for controlling options such as level of logging, again to deal with limited-capacity machines.

VII. Conclusions and Future Work

This paper presents our initial experiments in Africa with an automated reading tutor to improve child literacy. It investigates the viability and effectiveness of a computer-based reading tutor in improving the reading ability of children, particularly those attending under-resourced public schools. Although literacy and education for all are at the top of the global development agenda, not much work has been done regarding the role of technology in this process. This

work is a useful contribution both because it demonstrates that there is promise for the effectiveness of the approach and because of the practical lessons regarding the implementation of the study. We have presented an initial proof-of-concept investigation. Many additional questions would need to be explored to understand the potential for large-scale application of these technologies.

Next steps for our project include trying to set up a year long in-school study in Zambia. The purpose of the study would be to test the feasibility of an in-school usage model in a developing community setting. It would measure the children's reading gains over an entire school year of using the Reading Tutor and would compare these gains to those obtained with a non-technological approach to guided reading practice. An important goal of the year-long study will be to learn about sustainability and the feasibility of incorporating teachers closely into this work.

Additionally, we will continue to develop an understanding of what features would be required of a literacy-enhancing tool developed specifically for use by children in developing communities. What features would be needed, taking into consideration limited resources and equipment, as well as current instructional practices? What tutorial methods are most useful? Could a lighter weight tool be designed for alternative platforms, such as mobile phones, as has been done with the educational games for English-as-a-second language instruction in the MILLEE project [7]? There is significant scope for further research on the role of computing technology in improving child literacy in developing communities.

ACKNOWLEDGMENT

We are grateful to Steve Fienberg for his advice on the statistical analysis as part of the Statistical Practice course at Carnegie Mellon University. We express our appreciation to the Ghana-India Kofi Annan Centre for Excellence in ICT (AITI-KACE), Accra, Ghana, for their role in the controlled study, particularly to Dorothy Gordon for her support, Patricia Nyahe for her tireless management of the study, and the UNESCO Ghana Office for sponsoring their work. We also thank Leslie Casely-Hayford of Associates for Change, Ghana, for her advice and contributions to the pilot and controlled studies in Accra. We acknowledge the valuable contributions of Cybercity Internet Café and ProjectEDUCATE to the pilot studies in Ghana and Zambia respectfully. Finally, we are indebted to all the participating children and schools, and to all others who contributed to the project in various ways.

REFERENCES

[1] M. J. Adams, "The promise of automatic speech recognition for fostering literacy growth in children and adults," in M.C. McKenna, L.D. Labbo, R. D. Kieffer, & D. Reinking (Eds.), *International Handbook of Literacy and Technology, Volume 2*. Mahwah, NJ: Lawrence Erlbaum Associates, 2006.

[2] R. L. Allington and A. McGill-Franzen, "The Impact of Summer Setback on the Reading Achievement Gap," in *Phi Delta Kappan* 85(1):68-75, September 2003.

[3] J. Cohen, *Statistical power analysis for the behavioral sciences* (2nd ed.), Hillsdale, NJ: Lawrence Earlbaum Associates, 1988.

[4] J. Cohen, "A power primer," *Psychological Bulletin*, 112, 155-159, 1992.

[5] S. L. Deno, "Curriculum-based measurement: The emerging alternative," in *Exceptional Children*, 52(3):219–232, 1985.

[6] J. Hasbrouck and G. A. Tindal, "Oral reading fluency norms: A valuable assessment tool for reading teachers," *The Reading Teacher,* 59(7), 636-644, 2006.

[7] M. Kam, A. Agarwal, A. Kumar, S. Lal, A. Mathur, A. Tewari, and J. Canny, "Designing E-Learning Games for Rural Children in India: A Format for Balancing Learning with Fun," *Proceedings of ACM Conference on Designing Interactive Systems* (DIS '08), Cape Town, South Africa, February 25-27, 2008.

[8] S. C. Larsen, D. D. Hammill, and L. C. Moats, *Test of Written Spelling*, Pro-Ed, Austin, Texas, 1999.

[9] M. Lipson and K. Wixson, "Evaluation of the BTL and ASTEP Programs in the Northern, Eastern, and Volta Regions of Ghana," Report prepared by the International Reading Association for The Education Office, USAID/Ghana, August 2004. Available online at http://www.reading.org/resources/issues/reports/ghana.html

[10] R. G. Miller, *Simultaneous statistical inference*, 2nd ed., Springer Verlag, pages 6-8, 1981.

[11] J. Mostow and G. Aist, "Evaluating tutors that listen: An overview of Project LISTEN," in K. D. Forbus & P.J. Feltovich (Eds), *Smart machines in education* (pp. 169-234), Cambridge, MA: AAAI Press/The MIT Press, 2001.

[12] J. Mostow, G. Aist, P. Burkhead, A. Corbett, A. Cuneo, S. Eitelman, C. Huang, B. Junker, M. B. Sklar, and B. Tobin, "Evaluation of an automated Reading Tutor that listens: Comparison to human tutoring and classroom instruction," *Journal of Educational Computing Research*, 29(1), 61-117, 2003

[13] J. Mostow, G. Aist, C. Huang, B. Junker, R. Kennedy, H. Lan, D. Latimer, R. O'Connor, R. Tassone, B. Tobin, and A. Wierman, "4-Month evaluation of a learner-controlled Reading Tutor that listens," In V. M. Holland & F. P. Fisher (Eds.), *The Path of Speech Technologies in Computer Assisted Language Learning: From Research Toward Practice* (pp. 201-219), New York: Routledge, 2008.

[14] National Institute of Child Health and Human Development, *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* (NIH Publication No. 00-4769), Washington, DC: U.S. Government Printing Office, 2000.

[15] U. S. Pawar, J. Pal, K. Toyama, "Multiple Mice for Computers in Education in Developing Countries," *International Conference on Information and Communication Technologies and Development*, ICTD 2006, pp.64-71, May 2006.

[16] R. Poulsen, P. Hastings and D. Allbritton, "Tutoring Bilingual Students with an Automated Reading Tutor That Listens," *Journal of Educational Computing Research*, 36(2), 191-221, 2007.

[17] Project LISTEN Videos, available at http://www.cs.cmu.edu/~listen

[18] K. Reeder, M. Early, M. Kendrick, J. Shapiro, and J. Wakefield, "The Role of L1 in Young Multilingual Readers' Success With a Computer-Based Reading Tutor," Talk at the *Fifth International Symposium on Bilingualism*, Barcelona, Spain, April 2005.

[19] M. R. Shinn, N. Knutson, R. H. Good, W. D. Tilly, and V. L. Collins, "Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading," *School Psychology Review*, 21:459-479, 1992.

[20] K. E. Stanovich, *Progress in Understanding Reading: Scientific Foundations and New Frontiers*, New York: Guilford Press, 2000.

[21] UNESCO, "Education for All Global Monitoring Report 2005 – Education for All: The Quality Imperative," United Nations Educational, Cultural and Scientific Organization (UNESCO) Publishing, 2005. Available online at http://portal.unesco.org/education/en/ev.php-URL_ID=35939&URL_DO=DO_TOPIC&URL_SECTION=201.html

[22] UNESCO, "The Plurality of Literacy and its Implications for Policies and Programmes." UNESCO Education Sector Position Paper. United Nations Educational, Scientific and Cultural Organization, Paris, 2004. Available online at http://unesdoc.unesco.org/images/0013/001362/136246e.pdf