Does daily reading make a difference?

A 2014-2015 study of Sustained Silent Reading (SSR) in Lao primary schools

Big Brother Mouse :: Laos :: July 2015
In 2006, Big Brother Mouse began publishing “Books that make literacy fun!” in Laos, a country where reading skills and education were weak, and most people had never read a book for enjoyment.

From 2006 until the middle of 2013, we had three main components to our work:

Creating more high-quality books. This was always our #1 goal. We didn’t have enough books. If we got children truly excited about reading, they would soon read everything available to them.

School book parties. Half-day book parties at rural primary schools were our main technique for distributing books and promoting literacy. At each event we read stories aloud, played games, sang songs about books, and talked about how to care for books. At the end, every child got a book of their own, and we left more books (initially 50, then 80) with the school. Virtually no schools had a library; we simply encouraged them to make
the books available through whatever mechanism worked best for them.

**Trying other ideas.** Most notably, we also experimented with reading rooms in the homes of volunteers in villages. Many of these were successful for a time, but it required regular visits, training workshops, and restocking with books, to keep them active. We found that working through schools got more results from our limited resources.

## Sustained Silent Reading

Sustained Silent Reading (SSR) is a common name for the practice of having students read every day. There have been many studies of SSR in developed countries; we have not found any examples from less-developed countries.

We originally began thinking about an SSR program because of descriptions by Jim Trelease in his *Read-Aloud Handbook*; then the work of Stephen D. Krashen, a retired professor at the University of So. California and a strong advocate of SSR, provided further help.

SSR programs typically include certain key characteristics:

- It takes place at the same time every day, usually for about 15 minutes.
- Students choose for themselves what to read. Many schools let students read anything at all, including comic books.
- The goal is that they read for enjoyment, build up reading skills, and acquire a habit of daily reading.
- In developed countries, often students are told to bring something from home or from the library; in Laos, they won’t have anything to read unless the school provides it.
- Teachers are encouraged to sit at their desk and read something that they themselves enjoy, setting an example.
- Students are not tested on what they read, nor are they required to write or report about it.

Like any new program, SSR has faced objections. Some parents complain that teachers are paid to teach, not to read a book. Others say that students can read at home (although those who most would benefit probably do not do so). Financial interests may come into play; some commercial reading programs sell workbooks that must be bought fresh for each student.

Prof. Krashen’s book *Free Voluntary Reading* summarizes many studies of this technique. He found:

- Overall, SSR is successful: 51 of 54 studies found that SSR students did as well or
better than comparison students.

- Long-term programs (7 months or more) were most successful: In 8 of 10 such studies, the SSR students did better; the other 2 studies found no difference.
- It is most effective for students who have some reading skills but who are not strong readers.
- SSR students also improve in the areas of grammar, writing, and vocabulary.

**First evaluation of SSR in Laos, 2013-14**

In Sept. 2013, we began an SSR program in Laos.

We continued our half-day book parties, but at the end, we gave a set of books to each classroom, so they can have a daily reading time.

That month we also did the baseline testing for a controlled, randomized evaluation to compare reading improvement in schools that got this program, with schools that did not.

**Results of the first evaluation**

At the end of the school year (after about 7 months of SSR) we tested again. In schools that had agreed to provide the daily reading time, students in grades 3 to 5 showed an average of 26% more improvement than those in schools without the program. There was no difference for students in grades 1 and 2. Their reading skills were so poor that having books, in itself, did not help.

To some, this may simply confirm the obvious: Children who read every day will get better at it, than those who do not -- just as daily practice would also make them better at football, computer games, origami, or cooking. In Laos, however, “education” is generally considered to mean a teacher and a blackboard. Many people do not see books and reading as an important part of education, so it was helpful within Laos to have data that showed the connection.

In addition, the evaluation gave us some valuable insights:

1. Reading levels were lower than teachers had reported. Repeatedly, teachers in varied locations had told us, “Grade 2 students can read a word but not a sentence.” In reality, in the 40 schools we tested, we found that less than 20% of students entering grade 3 could read a single, simple word.

2. Our first follow-up, two months after the program started, revealed that many teachers did not actually provide the reading time. They believed that “just” letting students read wasn’t doing their job. Or perhaps force of habit won out. We held workshops, at which Department of Education officials reinforced the message that this was something teachers should do, but by then half the school year was gone.

3. At these workshops, teachers asked for aids in helping children learn to read. We
developed a 7-book series, “I Can Read!” which introduced the letters one by one, with enjoyable practice at every step along the way.

4. Although we hadn’t been looking for it, the tests revealed a substantial obstacle for children learning to read: The “complex vowels,” in which 2 or 3 simple vowels that have one sound when used alone, combine to create different sound. We used this insight in preparing the “I Can Read!” series.

The second evaluation, 2014-15

In September and October 2014, we sent our teams to a total of 45 schools in 3 districts in 1 province. We chose this province because, after our first evaluation, we wanted to compare the impact of one versus two reading periods a day.

This was a controlled, randomized evaluation. For those unfamiliar with these terms:

**Controlled** means there is a control group, consisting of schools where we tested reading levels at the beginning but we did not leave any books or start the SSR program at that time. (We will do so in Sept.-Oct. 2015.) This is the best way to evaluate whether a program actually makes a difference. We measure reading levels in each group at the beginning of the year. Then we start the program in some schools. At the end of the school year we return and measure again. Hopefully, reading skills improved everywhere, whether they got books from us or not. If the SSR program made a difference, then scores in the schools that had SSR will show more improvement than those that didn’t have it.

**Randomized** means we randomly chose which schools got SSR and which did not.

**Differences from the first evaluation**

In 2013-14, 2 of the 3 districts we chose were relatively remote and poor. This year, only one district would allow us the additional time to test the effect of adding another reading period each day, so we went to three districts in that district. This meant we had a less varied mixture of schools and students. Overall, these districts are less remote and more prosperous than the average.

In addition, as evidenced by their willingness to let us allow more reading time in some schools, this province seems to us to be unusually active in trying to improve education. Being less remote, officials may simply have more experience trying new things. Our baseline tests showed that more than twice as many students entering grade 3 could read a word here, as compared to the first year. Unfortunately, the districts we went to the first year probably are more representative of Laos as a whole.
Three groups of schools

We divided schools into three groups:

A. 30 minutes a day: 15 schools agreed to have 15 minutes in the morning, and again in the afternoon.
B. 15 minutes a day: 15 schools agreed to have one 15-minute reading time every day.
C. The control group: This was 15 schools that got no books or program at this time.

Testing and measurement techniques

We used 8 tests, which broadly fall into 3 categories. Below, small indented type is used for detail that may not be of interest for many readers.

Type I: Oral test: Identify a number, letter, or word

For students entering grade 1 we wanted something very easy. We started with numbers: We held up one large number printed on a sheet of paper, and tried to record how many knew it.

This was not as easy as it might seem. At first we tested students individually. Shyness or nervousness seemed to keep some of them from responding, even if they knew it. So we tried a different method, making it more of a game: The tester holds up the number, and asks everyone to shout out what it is. She identifies one person who shouted first, and clearly does know it and wasn’t just repeating what was already said; that student goes behind her and can’t see the next numbers. This continues until no one can identify a number.

We use the same technique for individual letters (consonants only; in Lao, vowels are complicated, and as a marker of reading, testing just for consonants seems suitable) in grades 1 and 2.

We did the same, with common one-syllable words, in grades 2 and 3, and at the end of the year for grade 1.

If the fastest student knows the letter (or number, or word) but is very slow to get it, that’s recorded as half a point.

We don’t stop after one word or letter that none of the remaining students can read; we try to more, and only stop when they’ve been unable to answer three in a row.

This year, we attempted to measure reading abilities at two levels: First, we showed words that have a complex vowel. When no more children can read those, we switched to words with a simple vowel. We thus got 2 scores for each test. The second number (how many could read a simple word) is the one we’ve used in our comparisons.
Type II: Written multiple choice (10-20 questions per test)

We use several types of multiple-choice test. All have exactly one right answer, there is no “none of the above” or “all of the above” choice. They have a choice of 4 answers.

Pictures and vocabulary: There’s a picture. The student selects one of four words below that identifies it.

Information: This has typical multiple-choice questions. When we originally designed this, we categorized questions depending whether students might know the answer (1) from general experience; (2) from a textbook; or (3) from a Big Brother Mouse book. In the end, however, we found that access to and use of the books in 2 and 3 varied too much for these to be useful, creating an additional unknown variable, so we only tallied category 1. So while we still refer to this as the “information” test, it really serves as a simple reading comprehension test.
Details of test creation and scoring: We begin by creating about 80 to 150 questions for each type of test. Some are most suitable for one particular grade, some for several grades. To create a test, we define a pool of suitable questions, typically about 50-100, for that grade. Then we make at least 5 different tests, each drawing randomly from this pool, trying to use each question equally often. We print and collate the five tests so that when they’re distributed, students next to one another will have different tests. This reduces copying; it also improves accuracy by giving us a wider base of questions being asked.

Recording answers: In our office, each response to each question is recorded in our database. We designed the system so that two people would enter each sheet, at different times, and the computer would check to be sure they were the same. In reality, we found that entries were quite accurate (less than 0.5% error rate), differences usually arise about interpreting, for example, whether a student marked two answers, or just one and erased the other one but didn’t erase it very well, so we cut back to just a single entry of each test.

Type III: Reading Aloud

Students stand in turn, and read aloud for one minute. The class score represents how many lines the average student can read in one minute.
Example of Picture-Vocabulary test, Grade 4.
We explain how multiple-choice works before the test, but even so, some students do not understand it.

If they can’t read, we don’t make them stand there for a minute, but the score is calculated as if that minute were used.

The text we use is a traditional fairy tale, enjoyed by both children and adults. The average line has about 20 syllables. (Lao is best measured in syllables rather than words.) In the initial round of evaluations, the same text was used in all 40 schools.

**Technique:** Everyone gets a copy of the book. Children read in a random order: They each get a card with a unique number, then we call a number from a random list of numbers. When their number is called, the student stands and reads for one minute, if they can, beginning where the last student left off. If they cannot, they sit down. Some make the effort, but give up mid-way. When the minute is up, they get to finish the phrase or sentence they’ve begun.

To score, we simply divide the number of lines read by the number of students who had a chance to read. We ignore any extra time that was used by students, as they finished a phrase.

**Issues with multiple-choice**

One concern is that some students don’t understand the multiple-choice concept. Testing methods vary in Laos; often teachers write questions on the board and students copy
them into their notebooks. Multiple-choice may seem intuitively obvious, and we explain it before each test, but we expect some students are still confused by it. For purposes of comparing relative levels of improvement, we believe this provides useful data. This issue is one reason, however, that we also used other types of tests.

**Tests used, by grade:**

We use two tests for most grades, and an extra one in first grade.

**Grade 1:** Identify a number. Identify a letter. Read a simple word.

We originally planned to use just the first two tests; previously we had found so few first graders who could read, that it didn't make sense to test that. But this year, after we saw that reading levels were much higher than at the schools we tested the first year, so at the end of the year we also tested them on oral reading. Therefore, we did not get a baseline for this, but we have no reason to think that many students entering grade 1 could read a word. (And if some did, then the control group -- which scored better at identifying
numbers -- was probably also better with words, and our results would under-estimate the impact of the reading program.)

**Grade 2:** Identify a letter. Read aloud a single word.

**Grade 3:** Read aloud a single word. Multiple-choice picture vocabulary

**Grade 4:** Multiple-choice picture vocabulary. Multiple-choice information questions.

**Grade 5:** Reading aloud. Multiple-choice information questions.

**Timeline and follow-ups**

We conducted the initial baseline tests at the end of Sept. and first week of October, 2014; and the final tests in late April 2015. Thus schools had the program for about 7 months, between the two evaluations. In his overview of SSR, Prof. Krashen uses 7 months as the dividing line between long-term and short-term programs. Long-term SSR had a strong record of success; short-term programs sometimes show no difference. So we believe a 2-year evaluation might show even stronger results, but we have no plans for it at this time.

We did repeated follow-ups with each school that got the SSR program, to be sure it was implemented. These included:

- Workshops, with many teachers from one area together for a day;
- Brief half-day visits by two members of our staff;
- Week-long visits, in which two members of our staff stayed with a family in the village and worked with each classroom each day.

**Problems, mistakes**

Number skills are generally very poor in Laos. We trained some of our office staff to enter answers to the multiple-choice questions (test type II). Then they put aside the envelope for that school, and a different person, with strong number skills, entered types I and III, which required some simple calculations. Somehow, one box of these envelopes disappeared before he got to them, so for 15 schools, we have only results from the multiple-choice tests. The missing envelopes were well distributed among the three groups; we still had full information for at least 9 schools in each of the three groups, and we had multiple-choice test information for all schools.

We made some changes from the first year, as far as which grades got which tests. We didn’t anticipate how easily this could cause confusion. In 24 schools, grade 5 students were given last year’s follow-up test. We considering using the results anyway, since both
tests were intended to measure comprehension, but we decided that they use very different approaches and should not be included in the tally.

In two schools, grade 4 was given the “information” test that was intended for grade 5, and so students presumably got slightly lower scores than they should have. One of these schools was in group A, one was in group B. The result is that students who had the reading program perhaps improved very slightly more than the numbers show.

Results

We returned to these 45 schools for follow-up testing, in late April and early May, 2015. Once again, we got some unexpected news and insights, as well as an opportunity to measure the impact of the reading program. Detailed tables are on pages 16-19.

Briefly:

• On average, reading abilities in Group A (which was to have 2 reading periods a day) increased about 45% more than those with no reading (Group C).
• Group B, with one reading period per day, improved about 34% more.
• Last year the program did not help grades 1 and 2. This year it helped all grades, in roughly similar degrees.

Important details:

A number of factors should be considered when looking at these numbers. In particular, some conditions here perhaps produced higher results than could be expected in different conditions.

• How much reading time? We did our best to find out how much actual reading time was provided. In fact, not many schools did allow 2 reading periods per day. On average, Group A only had one more reading period per week than Group B. (Averages are shown in the last column of the results table.) That seems very unlikely to account for the difference between 45% and 34%. Much more likely is that most of this difference is just statistical fluctuation from a relatively small sample of schools. While it would be nice to use the higher figure, we think it’s most accurate to say a once-a-day program made about a 39% difference, and that the impact of a twice-a-day program still isn’t known.

• Follow-up was critical. These numbers measure the impact of a daily reading program with repeated follow-up to make it happen. As we discovered last year, to simply start the program, and not do the follow-up, will not get the same results.

• Book quality. Our goal, since 2006, has been to create “books that make literacy
Traditionally, rural Lao children learn the alphabet from a teacher and a blackboard, and they may have no experience with books. This can sound like an exaggeration but these pictures make the point: In each photo, a child is looking at a book upside down, and doesn’t seem to realize it yet. All of the books have pictures as well as text. While orienting a picture properly might seem intuitive and automatic, it is not.
fun!” We’re proud of the books we’ve produced. We think the fact that children are eager to read these books helped the program. But we’re biased. We have heard of programs that used donated books, which perhaps were donated because nobody wanted them, and if others try the SSR approach we hope they will keep a strong emphasis on finding books that children are excited about reading.

**Fifteen minutes was a big increase.** Children in Laos typically have no opportunity at all to read for enjoyment, and perhaps no opportunity to read books at all -- they may read only from the blackboard. Textbooks are generally available, but are often too hard for students at the level for which they are written. So even 15 minutes a day represented a giant relative increase. If this were done in a different country, where students already do some reading, we wouldn’t expect to see such a large percentage increase.

**Low comparison point.** The increase also came from a comparison with a weak increase in the control group. It’s easier to show a big percentage gain from a low starting point than from a high one. A small business might double its sales in a year; a big corporation rarely does so. Last year, we found many students in grade 5 who could not read a word. If a 15-minute SSR program were tried in a country where students learns to read by the end of grade 2, the results might be different.

**It’s a new program.** The two factors above suggest that the program might show weaker results, after students are reading more and improving skills faster. But working in the other direction: Most of us get better at things, as we do them more. That’s the whole point of daily reading. With more experience, it’s reasonable to expect that teachers will find ways to make SSR more effective, for example by noting which students need more support.

**SSR isn’t the point.** Our goal is that students grow to love reading, and that books become more widely available. We didn’t try to measure attitude changes, but some teachers told us that students came to school early, or came back early from lunch, so they could read. That enthusiasm -- not SSR -- is where we’ll get the real pay-offs.

**New beginning-reader books.** We added our new “I Can Read!” series this year, and indeed saw a strong increase in the number of first graders who learned to read, though the actual number was still too small.

Our tests confirmed that the complex vowels are a stumbling block. Entering grade 2, 31% of students could read a word with a simple vowel but only 10% if it had a complex vowel. Entering grade 3, the numbers were 47% and 23%.
TABLE OF RESULTS

The three rows represent the three groups we created, with 15 schools in each group.

Tests: Tests are described earlier in this report. In the grey, paired numbers for each test, the first number is the baseline score, in September-October 2014. The second number is the followup, in late April 2015. For grade 1, reading orally, we assume a baseline of 0 as explained above.

Before and After Scores: In the first 4 test columns, this number is the percentage of students who could identify a number or letter, or read a word. For the multiple-choice tests (columns 5 and 6) these scores are the percentage of questions that were answered correctly, with the number adjusted as explained below. For “Read Aloud” the number shows how many short lines of a story were read aloud by 10 students, each of whom could read for 1 minute, if they were able.

The single boldface numbers on the next line show how much these scores increased. The number at the end of the line is the sum of these increases.

For various reasons, it might make sense to weight some scores more than others. We tried doing this (as we did last year), but it made very little difference in the bottom line,
and is difficult to explain to anyone who has limited experience with such things; it can seem like we were just fudging the data. We decided to keep it simple.

**Scoring multiple-choice questions**

Scoring a multiple-choice test seems straightforward, but actually, it’s not. Do you treat a blank as a wrong answer? It seems obvious that you do, but that leads to an anomaly: A student who cannot read at all, but who guesses on every question, will get a 25% score. A student who can read a bit, and answers 10% of the questions correctly but leaves the rest blank will seem to do worse, with a 10% score.

If most students are scoring 70% to 100%, this isn’t very important. But here, a lot of students were scoring 25% to 35% on a multiple-choice test with 4 answers. That disguises a big difference. The first child probably cannot read at all and was randomly guessing; the second one actually was picking the right answer occasionally. We made two adjustments to avoid this problem: (1) We assumed that questions left blank were answered correctly 1/4 of the time, to get “AdjustedScore”; and then (2) we further adjusted with the formula (AdjustedScore - 25) x 4/3. So a student who scored 100% still gets 100%; a student who scored 25% through randomly guessing will get 0%, and a student who leaves half the questions blank and guesses on the other half will also get 0%.

<table>
<thead>
<tr>
<th>Comprehension</th>
<th>Read Aloud</th>
<th>Average</th>
<th>Reading times per week</th>
</tr>
</thead>
<tbody>
<tr>
<td>38.1</td>
<td>28.2</td>
<td>5.671</td>
<td></td>
</tr>
<tr>
<td>61.7</td>
<td>46.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.5</td>
<td>18.1</td>
<td>26.0</td>
<td></td>
</tr>
<tr>
<td>52.3</td>
<td>45.8</td>
<td>5.147</td>
<td></td>
</tr>
<tr>
<td>69.6</td>
<td>65.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.3</td>
<td>19.2</td>
<td>24.0</td>
<td></td>
</tr>
<tr>
<td>57.5</td>
<td>35.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>73.3</td>
<td>45.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.8</td>
<td>10.0</td>
<td>17.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
</tbody>
</table>

(44.7% = improvement in group A compared with group C)

44.7%
What’s next

The best news is that now we can show teachers that reading is not a “waste” of school time. Long-established habits and beliefs won’t disappear immediately, but it should gradually become easier to get them to actually implement the program.

Clearly, it’s going to take time for teachers to accept that letting students read can be an effective way to improve education. Much more useful that the data here, we expect, will be word of mouth, as teachers, students, and parents comment on the benefits they see from SSR. This is already happening.

The evaluations are expensive and time-consuming, and we do not plan more right now. If someone isn’t convinced by a 39% improvement, then another study by us isn’t likely to change their mind. We hope this might inspire an outside, independent researcher to conduct an evaluation, here in Laos or in another developing country.

Our plans are to use our limited funds and time to continue working with the schools that have the program, and to set it up in a some additional schools, to be sure it continues
to work and isn’t forgotten after a while because it’s too much trouble to keep the books from dissipating.

We have also begun asking schools to co-pay a share of the cost (currently 25%; we’d like to get that to 50% soon). That allows us to get SSR into more schools. It could also improve implementation, because we won’t go to schools unless they’ve made a conscious decision that they want this, and are willing to pay something for it.

The “I Can Read!” books clearly seem to have helped: Almost twice as many first-graders who had these books could read a word at the end of the year, as in the control group. Some teachers asked for something to help them explain how these books work, and we’d already been considering that. We’re making a large-format book that teachers can hold up while introducing the basic concepts.
Thank you!

Giving children a multiple-choice test is not as exciting for us, nor as appealing for most supporters, as giving them a book. But it has helped us get a better understanding of what’s effective. We’d like to thank the donors whose unrestricted support let us do the evaluation this year:

Manoj Paul
Peggy Horn
Planet Wheeler
Cubit Family
Tricia Sharpe and Ray Barker
Brian and Gerry Warren
Zena Carter
Colin Cotterill and Books for Laos
Chris Ashton
Robert Moyer
Denis and Helen Weily

Thank you!