



Does daily reading make a difference?

Final report of a 2013-2014 study of Sustained Silent Reading (SSR) in Lao primary schools

This includes our preliminary report, written in October 2013 with a January update, telling the details of how we did the study, the tests used, and some observations from the baseline tests.

The final report, which includes results of the comparative study, begins on page 20.

A shorter (12 page) version, with less detail about the test and methods, is also available. On our website menu, please click “What We Do” then scroll down to “Evaluations.”

Big Brother Mouse :: Laos :: June 2014



About this report

We did this study in cooperation with the Department of Education and Sports in three provinces, for the purpose of measuring the impact of our daily reading program.

We studied 40 schools in 3 districts, each in a different province. We tried to get a representative sampling, but these figures do not represent all of Laos. (We did baseline testing in a fourth district in April, which scored much higher than any of these.)

The school system here struggles against considerable obstacles. We're not trying to highlight problems, but to identify them so we can work with school officials to fix them. We're releasing this so our supporters can see what we're doing, you are welcome to share it.

Thank you for your interest in our work.

Sasha Alyson
for Big Brother Mouse
Luang Prabang, Laos
www.BigBrotherMouse.com

Sustained Silent Reading in Laos

In Sept. 2013, Big Brother Mouse began a new program in Laos.

We hold a half-day book party at a rural primary school, as we've been doing for nearly 7 years. We get kids excited about books and reading, and let every child choose a book of their own, nearly always the first book they ever owned.

Then, we do something new: We give a set of books to each classroom, so they can have daily reading period. This is the first time we've promoted daily reading, a technique often called "Sustained Silent Reading" (SSR). In the past, we didn't have enough books to keep students reading day after day. In fact, as of September we were still a bit short; not until the end of 2013 did we have as many beginning-reader books as we wanted.

We did something else new in 2013: We began a controlled, randomized evaluation to compare reading improvement in schools that got this program, with schools that did not.

As the person who planned it, and the only native English speaker at Big Brother Mouse, I am writing the report, and I've occasionally referred to my personal experiences or opinions. Please don't let this distract from the fact that as I write this, the real work is being planned and carried out by a 100% Lao staff.

Why do an evaluation?

Some programs obviously need a rigorous evaluation of their impact. If you decide to use text messaging for HIV-awareness, you need to evaluate whether that's a better use of resources than other techniques. But if you're getting books to children who never owned a book, and to schools that had nothing to read, isn't that a worthwhile end in itself? Should you also take time from what you want to do, to measure whether kids who have books read better than those who do not?

All else being equal, an evaluation would certainly be good. But it consumes resources. The main resource isn't money, but time from the people whose skills are most needed for other things.

But advice from two places convinced us to do it: Mark Cubit, at Planet Wheeler, made a strong argument that this kind of evaluation was important; GiveWell.org, arguing that most "aid" money is wasted, makes a strong case that rigorous evaluations are the best antidote to this. Their site has been valuable and I hope their voice will be widely heard in the years ahead.

Here are our reasons for deciding that a thorough evaluation was worth the time:

1. “Obvious” isn’t enough. It seems obvious that daily reading, in schools where children have little or no other opportunity to read, will improve reading skills. But even the obvious sometimes isn’t right. More plausibly, perhaps that assumption is correct, but the implementation is weak. Teachers don’t do it, the books get lost or stolen or fall apart, or the books we provide aren’t as suitable as we thought. An evaluation helps us see that there’s a problem to address.

2. Evaluations should be standard practice. Good evaluations are rarely done in international aid work. We hope to contribute in a small way toward making them the norm.

3. We can evaluate other things. Once the testing framework is set up, it’s easy for us to test variations in the program. The point of our book parties is to get kids excited about books and create “consumer demand” for reading time, but we can greatly reduce costs if we bring together teachers in each district for a day-long training, then send them back to their school with books. Will that work as well? What happens if we help create a small library at each school, as well as putting books in the classrooms? What if we give some books to every family in the village?

4. We learned a lot. Even at this stage, the evaluation process has been unexpectedly valuable. As we set up the SSR program, we largely relied on two sources to decide what books to give to each grade level -- we looked at the textbooks for that grade, and we talked with teachers. As detailed below, we got misled. The baseline testing helped us identify that error and make changes. We gained other insights of value, as well.

Why so much detail, at this point?

This report contains more detail than some readers may need or want. (Details of specialized interest are indented, in smaller type.) We included the detail for several reasons:

1. This may help others who want to set up similar evaluations.
2. We’re doing a self-evaluation. Ideally, it would be done by an independent researcher, but nobody’s offered. All this detail doesn’t prove that we haven’t fudged any numbers, but will let others make an informed judgment about our technique and approach.
3. Along those lines, if you think we did make errors, please tell us so we can make improvements.
4. It’s a well-recognized phenomenon that research with positive or interesting results is more likely to be reported than research that finds nothing new. But that skewed information affects our ability to make good decisions. If 19 studies find that eating a piece of chocolate every day has no affect on cancer, but 1 study finds a 95% likelihood that it will ward off cancer and that study gets widely reported (and it will!), that doesn’t mean you should eat more chocolate, it means you need to know about the other 19 studies. But

those researchers didn't send out press releases.

We can't prevent this phenomenon, but by explaining in advance, in detail, what we're doing, we are removing the option for us to quietly bury the results if they aren't what we wanted. We hope to find a strong impact, but if we don't, we'll report anyway and those results should help us figure out where things went wrong, and plan something better.

Brief background

In 2003, I visited Laos for the first time. On that trip I never once saw a book written in Lao, and never saw anyone except tourists reading a book, magazine, or newspaper. That gave me the idea for a publishing project. For three years I came to Laos from a base in Thailand, usually once a month, to meet with government officials, NGOs, artists, students, and others. Two of those students, Khamla and Siphone, took great interest. Together we wrote five books that demonstrated what we had in mind with the slogan "Books that make literacy fun!"

In 2006, we published the first books, Khamla got his business licence, then a publishing license, and we opened a small shop and office in Luang Prabang. Several NGO's had expressed interest in our plans and said if we published good books for children, they'd buy large quantities.

They did not. So at the end of the year, we began developing our own distribution. A new member of the staff, 16-year-old Sonesoulilat, organized a book party at a rural school. At the end, we let every child choose a book of their own.

We were experimenting to see what worked, and that worked. As of early 2014, Sonesoulilat has organized another 3,500 book parties at rural Lao schools, and we've given some 400,000 books to children. Khamla, Siphone, and Sonesoulilat are now the three owners of Big Brother Mouse, which is a Lao-owned business operating on a not-for-profit basis (the only suitable option under Lao law); I act as a volunteer adviser.

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

Creating more books. This was always #1. We didn't have enough books. If we got children truly excited about reading, it wouldn't take them long to read everything.

School book parties. The book parties remained our primary technique for distributing books and promoting literacy. At the end, every child got a book of their own, and we left more books (initially 50, then 80) with the school. Almost none of these schools has a library; we simply encouraged them to make the books available through whatever mechanism they chose.

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

Publicity and fundraising

We've seen many organizations become so wrapped up in fundraising and promotion that they seem to lose track of their original goals. We agreed to try a different approach. We would focus on figuring out what was needed, and what worked, and doing it well. Would enough funding come in from people who saw what we were doing and liked it, so that we could keep our focus? I had enough savings from selling a small business to pay for all printing and overhead in the first year, as well as buying a large house that serves as our office and provides stability.

Largely, thanks in particular to a very small number of ongoing institutional supporters (Planet Wheeler; Global Fund for Children; and more recently, CLSA Chairman's Trust), and some generous individuals, we've spent well under 10% of our time on fundraising and promotion. We've put a lot of information on our website; we post an occasional update on FaceBook; and that's about all.

In the past year, for the first time, we feel we could do more if we had more funding. If the SSR program has as much impact as we hope, we'll take more time to look for that funding, while trying to keep our focus. If SSR doesn't show good results, we won't need increased funding, we'll need to spend some time and effort figuring out how to improve it.

Sustained Silent Reading

Sustained Silent Reading (SSR) is the most 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 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



An exciting day: This photo was taken on Sept. 16, the very first day we introduced SSR anywhere in Laos.

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 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 it). Financial interests may come into play; some commercial reading programs, with workbooks that must be bought fresh for each student, bring in steady sales.

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 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.

His book has been valuable to us, and is available online.

Overview of the study

In September and October, we sent a team of 7 people into a total of 40 schools in 3 districts in 3 provinces:

District #1: About 1 to 2 hours from Vientiane, the capital of Laos, with a major road running through it, 65% minority population.

District #2: A somewhat remote district in the mountains, though it does have a major road. (Many villages in the district, of course, aren't on that road and access is difficult.) It is 97% minority.

District #3: Similar location to #2, but with only 73% minority population.

We selected these because they are in different provinces and thus give us a greater variety in our sample; these 3 provinces have been very cooperative about working with us; and the districts are not too far apart, which reduces our travel time and expense. They are broadly representative of the rural villages that make up 80% of Laos, although they have a higher average minority population.

Minority levels are significant because these students speak an ethnic language at home. Schools are taught entirely in Lao, but usually these students enter school unable to speak or understand Lao.

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 but did not leave any books or start the SSR program. (But later, we will!) 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. (In fact, we did a 3-way rather than 2-way split; more on that later.)

Randomized means we randomly chose which schools got SSR and which did not. That's important, too. If, for example, we just started the program at schools that were close to the good road because they were easier to service, and then their scores improved the most, we wouldn't know if it was because of SSR, or because living on a good road meant their families had more income and children studied instead of helping on the farm.

We conducted two tests or evaluations in each grade (1-5) of each school.

What we've learned already

In the past we were not directly involved in everyday school activities. We go into schools to hold book parties, but that's a special day. We publish books for all ages and reading levels, and let students choose what to read; we never tried to determine reading levels of each grade.

Even now, our goal has been to measure the impact of the daily reading program, not to explore the education system. Nonetheless, we've gained some insights and new information. This is more complicated than one might expect because there may be some disparity between official policy, versus what a rural teacher believes is official policy, versus actual practice, versus what people think we should hear, versus what they want us to hear. And all of this may vary between villages and districts.

How well can students really read?

When we ask teachers about reading levels, we often hear the same thing: That in grade 2, students can read one word but not an entire sentence.

What we found was not so good. We don't think teachers were trying to mislead us; they simply were drawing broad inferences from the best and most visible students, and didn't think about whether those students were typical of the class. In the 40 schools we tested, only 19% of students entering 3rd grade could read a single word. Most 5th graders could not answer simple comprehension questions about a short Aesop's Fable story.

Enormous variation.

We found much greater variation between schools than expected. Here, in ascending order, is the percentage of students entering third grade who could read a common one-

syllable word (on the oral test, described below) in the 40 schools we tested:

0, 0, 0, 0, 0, 0, 0, 0, 2, 3, 3, 3, 3, 5, 5, 5, 6, 11, 12, 12, 15, 15, 15, 17, 17, 18, 19, 23, 23, 24, 26, 27, 28, 31, 38, 50, 50, 56, 67, 67

We collected some additional data about each school, simply to look for correlations that might be noteworthy. Nothing accounted for this great a spread. A logical guess is that it's related to teaching quality, but we don't know any simple way to measure teaching quality. (If you do, a lot of school systems are eager to hear from you.)

Preliminary look at correlations

Ethnic Minorities: About half the population of Laos speaks an ethnic language at home. Most of these children don't begin to learn the Lao language until they enter school. We expected this to be a big hurdle as they tried to learn to read Lao.

We did find a correlation between the percentage of students who are native Lao speakers in a school, and overall reading levels of the school, but it was relatively small. Some schools that were 100% minority had the best reading scores, and some 100% Lao schools had very poor scores.

A simple word, but not a sentence

Here's an example of a useful insight we've gained by looking at the results so far.

Every Lao syllable starts with a consonant, then a vowel follows it. If the spoken word starts with a vowel sound, the written word starts with a silent consonant. Some syllables also have a final consonant. (Any consonant can begin a word. Only some consonants can end a word, and they have slightly different sound in that position.) Some vowels are written with a single symbol, others (the complex vowels) require that several vowel symbols combine.

I took a detailed look at the multiple-choice picture-vocabulary that we gave to grade 3 in District #3, the best of the three districts we studied. About 60% of all 3rd graders could read the very simplest words, composed just one consonant and one simple vowel. Typically only 10-30% could read words that had a complex vowel, a final consonant, or multiple syllables.

This shows that most students have learned the basics: they know what sounds correspond to what individual letters. They're still weak on some nuances, particularly the complex vowels, and they desperately need practice.

At the same schools, only 40% of 4th graders could read "Where do you live?" well enough to check Laos, rather than China, India, or Thailand.

So it seems that the teachers were right when they described a large body of students who could read a word, but not a whole sentence. But those students are in grades 3-5, not grade 2. The process of going from a basic knowledge of the alphabet, to being able to comprehend a simple sentence, takes them a couple of years. They simply don't get any practice.

Three groups of schools

We divided schools into three groups:

1. A control group, that got no books or program at this time.

2. A standard group, that got the SSR program. This costs about \$350 US per school.
3. An “extra attention” group, which are getting at least one more follow-up visit, plus phone calls to ask how things are going, and more books. This will probably cost about twice as much.

We have two reasons for the 3-way split:

First, at the end of the year, we’re eager to have identified something that works. If the standard program doesn’t work, but the “extra attention” approach does, we’ve identified what is still a very low-cost way to improve reading and education levels.

Second, whatever the results, we’re likely to learn more and get better insights about what’s needed. Our increased interaction with group 3 may give us ideas about how to strengthen the standard SSR program.

A note about the cost: Effectiveness is almost meaningless, without also considering cost. It costs us \$350-370 to set up this program in a typical school (average 100-110 students). However, we also get sponsorships for the books that we print, which helps us keep down book prices. Without those sponsorships, the cost would be about \$440.

Testing and measurement techniques

We used 8 specific types of tests, which broadly fall into 3 or 4 categories. I’ve given a lot of detail here, in case it is helpful to others planning similar evaluations.

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.

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

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), and for common one-syllable words.

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.

On our first evaluations, we used easier words for grade 2 than for grade 3, because we weren't trying to compare different grades. We wanted to measure improvements in one group of second-graders against improvements in another group. We've decided in the future to use the same words for each grade, so that a comparison between grades is meaningful.

Another change in the oral word test: Lao has both simple vowels, which consist of one simple symbol; and complex vowels, in which 2 or 3 of these simple vowels combine to make a new vowel. On the first tests, we used a random sample of common one-syllable words. But we now see a problem with that: If the tester hit a sequence of three complex vowels early in the test, the class score might seem lower than it should be. We're trying a new system now: First we'll show words that have a complex vowel. When no more children can read those, we'll switch to words with a simple vowel. We'll thus end up with 2 scores for each test.

To measure improvement, we'll use this second number (how many could read a simple word). In a few cases, this will show greater improvement than was actually the case. However, there's no way for us to go back and know which cases those are. Using the same test format for the follow-up tests would not fix this, it would simply mean we had two numbers that were less accurate than they might have been.

Type II: Written multiple choice (10-20 questions per test)

We use several types of multiple-choice test. At this time, all have exactly one right answer, there is no "none of the above" or "all of the above" choice. Most of them have a choice of 4 answers; an "Aesop's Fables" comprehension test that we'd prepared before the others has just 3 choices. The tests are:

Vocabulary: There's a single word. The student selects the closest synonym. After the first week, we replaced this with:

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

Information: A typical multiple-choice question, as the term is usually used: A question requests the right answer. This information might come from (1) general experience; (2) a textbook used at that grade level; or (3) a book in our SSR set for that grade level.

"Aesop": In the future, we hope to add a longer type of question (2-3 sentences) with answers that test comprehension of those sentences. This year, the Aesop's Fables tests measure comprehension of a short (1 page) story.

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



Identify a letter. Students standing behind our team member have already identified a letter. She'll keep going until those left cannot do so.

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.

Each test has an ID number at the bottom. The first three digits refer to the pool of questions. Then a unique identifier distinguishes this from other tests in the same set, drawing from the same pool. Finally is a check digit to guard against errors when our staff types in the students' answers.

When we enter the responses into our database, we write a unique number on each sheet, so we can find it later if needed.

Recording answers: In our office, each response to each question is recorded in our database. Two people enter them, at different times; if the second person enters something different from the first, they're later asked to enter it again (without seeing what they already entered.) Cases where the first two entries are different get flagged, so that a supervisor can look for overall patterns of errors. In reality, so far we've found that entries are 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.

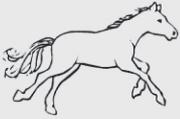
If a student marks more than one box, or none at all, their answer is recorded as "5", meaning "Don't know". If a student checks the same box on every question (typically the first box), we also consider that to be a "5".

With two-sided tests, students sometimes miss the back side. We're trying to monitor and be sure they see it. If they clearly just missed that side entirely, we enter a "9" (does not count) for those

ຮູບ

ຈົ່ງໝາຍຕົກເອົາຂໍ້ທີ່ຖືກຕ້ອງຕາມຮູບພາບ

1.



- ມ້າ
- ໝູ
- ສາມ
- ຕາ

2.



- ບູ
- ເຫັດ
- ໃບໄມ້
- ເຮືອ

3.



- ຊຸ່
- ຈຸດ
- ເມຍ
- ໂມງ

4.



- ກ້ອນ
- ຮ່າງກາຍ
- ກະຕ່າຍ
- ລີງ

5.



Example of Picture-Vocabulary test, Grade 3.

Percentages show how many students chose the right answer out of all those who saw the question (including those who left it blank); and how many got it right out of those who checked exactly one answer. By chance, 25% of the latter group should have gotten it right.

Originally this was intended to test vocabulary, and deliberately included some words such as *dinosaur* and *whale* that children might not know, but that they would have an opportunity learn about from our books. By tallying only questions about objects that everyone knows, the test can also measure reading ability.

1. (Horse picture)

- _Horse
- _Pig
- _Three
- _Eye

[56%, 65%. Scores were even higher for most other very simple words, such as fish, house, and chicken. Some children may not have seen a horse; they've all seen chickens.]

2. (Leaf picture)

- _Crab
- _Mushroom
- _Leaf
- _Boat

[35%, 43%. This is a word they should all know, but the Lao word for leaf has two syllables. From the very quick analysis we've done so far, that seems to be why it scored so much lower.]

3. (Clock picture)

- _Name
- _Point
- _Wife
- _Clock

[26%; 32%. I had wondered if all children would recognize this object. Others on the staff were confident that they would, but these low scores suggest otherwise. The question is still a useful measurement of general learning, but not of reading ability.]

4. (Rabbit picture)

- _Block
- _Body
- _Rabbit
- _Monkey

[33%; 39%. My first thought was, "Body" shouldn't be there, that's technically right, too." But in fact that was chosen less often than any other answer.]



We explain how multiple-choice works before the test, but even so, some students do not understand it.

questions and they are not included in the scoring. A “5” will lower the final score because it represents a question the student couldn’t answer; a “9” will not affect the score.

Analysis: Recording every reply separately means that later we can investigate:

- Are some questions more difficult than we expected, with a higher-than-random percentage of wrong answers?
- Are there ambiguous questions, revealed when a lot of students chose the same wrong answer?
- From the beginning to the end of each grade, how much do students learn from textbooks? Do they learn more from textbooks if their school had a reading program, compared to schools that did not? How much did they learn from the books provided by Big Brother Mouse? (But in fact, comprehension levels are so low that it was premature to start trying to measure knowledge.)

Type III: Multiple choice comprehension: Aesop’s Fables

Students read a short Aesop’s Fable story. We have 3 stories; students sitting adjacent get different stories. Then they answer 3 multiple-choice questions to test their comprehension, each with 3 possible answers. Eventually we’ll probably standardize this to have 4 questions, like the other multiple-choice tests.

ຂໍ້ມູນ

1. ປະເທດ ຂອງ ເຮົາ ແມ່ນ ຂໍ້ ໃດ?

- ໄທ
- ຈີນ
- ລາວ
- ອິນເດຍ

2. ເລກ ຫົກ ແມ່ນ ເລກ ໃດ?

- 5
- 6
- 7
- 8

3. ສັດໃດແມ່ນສັດເລືອຄານ?

- ແມງກອກ
- ງົວ
- ໝູ
- ເຕົ່າ

4. ແມວ ແມ່ນ ຫຍັງ?

- ແມ່ນ ດາວ
- ແມ່ນ ປາ
- ແມ່ນ ຕົ້ນໄມ້
- ແມ່ນ ສັດ

5. ແຕ່ກ່ອນເຄີຍມີລົດໄຟຄັນໜຶ່ງຢູ່ແຂວງ ໃດ?

- ຄຳມ່ວນ
- ຜົ້ງສາລີ
- ຈຳປາສັກ
- ສະຫວັນນະເຂດ

Example of information test, Grade 4.

As before, if students were picking an answer randomly, the second score for each test would be 25%. On two questions below, children did worse than random. (These percentages represent answer sheets from about 75 students. The number of questions that fell well below 25% was too high to be attributed to statistical fluctuations.)

These test included both easy questions, which merely measure reading comprehension, as well as questions about facts some students would not know. By looking at results from only one type of question, we can use the same test to measure improvements in reading; or increases in knowledge.

1. What is the name of our country?

- Thailand
- China
- Laos
- India

[40%, 46%. Since 25% would have gotten it correct by chance, we infer that about 20% of the grade-4 students who saw this question were able to read, understand, and answer correctly. No other information question scored significantly higher.]

2. Which number is a six?

- 5
- 6
- 7
- 8

[31%, 36%. This was supposed to be an easy question! It was not. The same question for other single-digit numbers got similar scores.]

3. Which animal is a reptile?

- Rhinoceros beetle [widely known here]
- Cow
- Mouse
- Turtle

[11%, 13%. A grade 3 textbook teaches about reptiles and identifies turtles as a reptile.]

4. What are cats?

- They're stars
- They're fish
- They're trees
- They're animals

[15%, 17%; the wrong answers were fairly evenly divided among the three wrong choices. Some research has suggested that Asians categorize things differently from westerners; even so, we were all surprised. So surprised that I pulled out the original student answer sheets to make sure the problem wasn't somewhere in the computer program. Of the first 19 sheets that included this question, only 2 had the right answer.]

5. In the past, which province had a railroad?

- Khammuan
- Phongsali
- Champasak
- Savannakhet

[32%, 37%. It was Champasak, in the south, in case you're wondering. The French, when they hoped to develop the Mekong as a trading route to China, built a very short railroad to bypass the rapids. A rusted steam locomotive still sits there, and is shown in one of our books along with a reference to the railroad. But who would have expected this to score better than "What are cats?"]

This year, we've given this test only to grade 5, and even at that level it was too hard for the vast majority. District #2 had the lowest average score, at 35%; (remember that they could have scored 33% by choosing answers at random.) District #3 was best, at 48%.

Type IV: 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.

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. So if one student reads 6 lines, and the next nine students can't read at all, that's a class score of 0.6 lines per minute.

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. We at this writing haven't decided whether to re-use it (so as to be sure of having no change in difficulty level), or to change, when we return; we'll probably decide it's better to take something new, that they haven't seen.

Technique: 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.

We record other information but at this writing we don't use it:

- How many students couldn't read at all
- How many started, but gave up
- Total time used
- Errors made while reading
- Ethnic composition of the class

We set up this test when we thought reading skills would be better than they actually are. With our second set of evaluations, we're replacing this with the multiple-choice picture vocabulary test.

Which tests are used for which grades

We use two tests for each grade.

Grade 1: Identify a number. Identify a letter.

Grade 2: Identify a letter. Identify (read aloud) a single word.

Grade 3, all schools: Identify (read aloud) a single word. Then:

District 1: Multiple-choice information questions

Districts 2 and 3: Multiple-choice picture vocabulary



Reading aloud: Each student stands and reads for one minute. Everyone else gets a copy of the same book and tries to follow.

Grade 4, District 1: Identify (read aloud) a single word. Multiple-choice vocabulary word
Districts 2 and 3: Multiple-choice picture vocabulary. Multiple-choice information questions

Grade 5: First evaluation: Aesop's Fables comprehension. Reading aloud.

Second: Aesop's Fables comprehension and multiple-choice picture vocabulary.

Issues with multiple-choice

We're concerned that many students don't seem to understand the multiple-choice concept. Very likely, their reading skills are so poor that this doesn't greatly affect results, but we're not certain.

Switching to a fill-in-the-blank approach has three big drawbacks:

- It takes more time to score;
- It's difficult to keep scoring consistent; inevitably there will be iffy answers that one scorer considers correct, and another does not.
- Many students have nearly illegible handwriting.

Timeline

We conducted the initial “baseline” tests from 16 Sept. to 11 Oct, 2013. We plan to return in late April and early May for the year-end tests. At the start of the next school year in September, schools in the control group that didn’t already get the SSR program will get it.

For the year-end tests, we’ll use the same formats and systems as we did at the beginning of the school year. Changes and improvements in the tests will be done only as we do more testing, in new districts.

Most schools will have the program for about 7 months, between the two evaluations. In studies of SSR, 7 months is sometimes used 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’re right on the cusp.

These studies are largely (or entirely) in more developed countries, where children have other opportunities to read. Here, they usually do not. So if SSR works here, we would expect to find some measurable results after this 7-month period. We could have planned a study that extends over 2 school years, but we felt the obvious drawbacks outweigh the likelihood of any benefits.

— January 2014 Update —

In January we returned to 3 districts. We held day-long workshops for teachers at selected schools that had agreed to start the daily reading program, to ask how it was going, and what problems if any they had encountered, and how we could help things keep moving ahead.

Teachers loved having the books. Children sometimes came to school early to read, or would read during play time. But, as we talked with teachers, we soon realized that many teachers weren’t doing SSR at the same time each day. They grew up in a world in which “education” meant a teacher and a blackboard. They knew reading skills were weak, but that was all the more reason not to give up true teaching time - a teacher and a blackboard - for students to just have fun reading a book.

We had wondered if implementation would be a difficulty, but decided it probably would not be. The program is very easy: For 10 to 15 minutes a day, the teacher simply lets children choose a book and read. Now that we know it didn’t go so well, we’re taking a number of steps to address the problem. Most useful, we think, will be for district-level education officials to tell the teachers that they should allow reading time.

We will continue with the evaluation, but we'll be seeing only the impact of a somewhat haphazard implementation. Whatever the measurable impact, we expect to learn from it. We're taking steps to improve implementation, and we're already doing the baseline testing for a second evaluation, which will run March-April 2014 to April-May 2015.

— FINAL REPORT, June 2014 —

We returned to these 40 schools for follow-up testing, from 28 April to 8 May, 2014. Once again, we got some unexpected news and insights, as well as an opportunity to measure the impact of the reading program.

As reported in our newsletter:

(The following is taken verbatim from our newsletter; if you read that, you needn't re-read it here.)

Brief background

In September and October, we tested reading levels at 40 schools in 3 provinces. Then we randomly divided the schools into 3 groups:

Groups A and B got our standard book party and children all got to choose a book of their own, then we left books in every classroom, and the teachers agreed to let children read a book of their choice for 10-15 minutes each day.

Group A got extra attention: We held workshops in early January to give teachers more books and see how they were doing, and we did other followup.

Group C didn't get any books or reading program at all. It was the control group, to help us judge whether any improvements in Groups A and B were likely to have been from the books, or would probably have happened anyway.

Results

Briefly: For grades 3 to 5, reading abilities in Group A increased 26% more than in Group C. Group B, as expected, was in between.

That's a step in the right direction. At an average cost of about \$4 per student, I don't recall hearing about any other technique that gets that type of result. The much-publicized "One Laptop Per Child" project about six years back cost \$200 per child just for the laptop, plus overhead, and was clearly a lot of fun and an ego-boost for the people who proposed it but with largely negligible impact.

But... 26% is not enough. It's far less than what is needed, and less than what we think is possible. (Imagine that your first child didn't learn to read until grade 4. Your next child can read by the end of grade 3. Are you happy now?)

In grades 1 and 2, the program made no measurable difference. Nor did we expect it to after we saw reading scores from these schools. Only 4% of the children entering second grade could even read a simple word. That's why we began creating our new "I Can Read!" series: So the teachers will get some help as students learn to read.

Observations

We learned at the January workshop that many teachers hadn't fully implemented the program; they were reluctant to give up the teacher-and-blackboard system that they grew up with, and which they believed was the only real way to teach. With support from district officials, we encouraged them to start doing it, even though the year was half-over.

If schools started the program immediately, they would have done it for an average of 7 months and 1 week, between our testing. That's the longest we can do in a single school year. Prof. Stephen Krashen has been a long-time advocate of SSR. He summarizes 54 studies of SSR in his book "Free Voluntary Reading." He found that SSR often needs a long-term program (which he defined as more than 7 months) to show results.

Krashen's work suggests that even where the program didn't produce measurable results (in some schools it did not; and generally in grades 1 and 2 it did not), that exposure to books may be laying the groundwork that will lead to bigger improvements in the long term. That's good to keep in mind, but we still believe, with improvements in the program, we can get some clearly-measurable results.

Krashen also found that SSR was most effective for children who had basic reading skills already, matching what we found.

Even in grades 3-5, a majority of children in these 40 schools can't read well enough to enjoy reading a simple story. Some have such weak reading skills that books wouldn't have helped them at all. We'll provide the "I Can Read" series for all grade levels, to help non-readers benefit from reading time.

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 aren't going to disappear immediately, but it should gradually become easier to get them to actually implement the program.

We will continue testing. We're making two major changes in the program (by

Lowest, Highest Range	0 97	14 86	5 57	12 69	0 70		
N	40	39	39	40	37		
	Oral Words	Picture Vocab	Information	Aesop	Read Aloud		
A 14	Raw data	17.8 48.3	43.7 61.4	21.5 35.3	42.4 54.1	21.7 42.1	
	Increase	30.5	17.7	13.8	11.7	20.4	94.0
	Adjusted	18.3 49.8	41.2 65.8	31.8 58.3	53.3 73.8	31.0 60.1	
	Increase	31.4	24.6	26.5	20.6	29.1	132.1
B 13	Raw data	17.4 48.5	43.2 57.1	25.1 37.0	42.8 50.7	18.0 40.0	
	Increase	31.1	13.9	11.9	7.8	22.0	86.8
	Adjusted	17.9 50.0	40.5 59.8	38.6 61.5	54.1 67.9	25.7 57.1	
	Increase	32.0	19.3	22.9	13.8	31.4	119.5
C 13	Raw data	17.0 42.8	39.9 55.5	22.9 31.3	37.0 48.9	22.5 36.2	
	Increase	25.8	15.6	8.4	11.9	13.6	75.3
	Adjusted	17.5 44.1	36.0 57.6	34.5 50.6	43.9 64.8	32.2 51.6	
	Increase	26.6	21.6	16.1	20.9	19.5	104.7

Grades 3, 4, and 5

TABLE OF RESULTS

The three rows represent the three groups we created: in Group A (14 schools) we left books for a reading program and gave extra support; Group B (13 schools) got books but no extra support; Group C (13 schools) got no books.

The numbers represent the averages of the average for each school, they are not weighted by number of students.

Tests: Tests are described earlier in this report (pp. 11-17). In the paired numbers for each test, the first number is the baseline score, in September-October 2013. The second number is the followup, in April-May 2014.

Raw data: For each group (A, B, C), the top line is the raw scores. For question #1, this number is the percentage of students who could read a common one-syllable word. For the multiple-choice tests (#2 through 5) these scores are the percentage of questions that were answered correctly. Blanks were counted as wrong answers. For question #5, it was 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 numbers on the next line show how much the raw scores increased. The number at the end of the line is the sum of these increases.

Adjusted scores: Using the raw data is simple and clear, but it gives too much weight to some tests. For example, “Oral Words” scores increased by 25 to 30 points, but “Information” scores never increased by more than 14 points. A one-point increase in an “Informa-

tion” score represents more improvement than a 1-point increase in “Oral Words”. So we adjusted these numbers for a more useful comparison.

The gray numbers at the very top show the very lowest and highest school-wide average from all 40 schools in this test. The next line shows the “range”, defined as the difference between these two. Then we adjusted the raw data to show where they stood within this range, using the formula:

$$([\text{RawScore}] - [\text{Minimum}] * (100 / [\text{Range}])$$

The third line for each group shows these adjusted scores, with the increase on the line below, and the total increase in the last column.

Lowest, Highest Range	19 100 81	0 96 96	0 96 96	
N	39 Numbers	38 Consonants	39 Oral Words	Total
A 14 Raw data	59.2 88.9	46.9 75.5	2.8 26.1	
Increase	29.7	28.6	23.3	81.6
Adjusted	49.6 86.3	48.8 78.6	2.9 27.2	
Increase	36.7	29.8	24.3	90.8
B 13 Raw data	55.5 85.9	46.8 76.6	4.9 25.5	
Increase	30.4	29.7	20.6	80.7
Adjusted	45.1 82.6	48.8 79.8	5.1 26.6	
Increase	37.6	31.0	21.4	90.0
C 13 Raw data	54.9 85.9	30.5 65.3	3.2 26.8	
Increase	31.0	34.8	23.7	89.5
Adjusted	44.3 82.6	31.8 68.1	3.3 28.0	
Increase	38.3	36.3	24.7	99.2

Grades 1 and 2

We originally expected to test grades 1 and 2 for reading ability, but our trial tests showed there was no point in doing so. Instead, for grade 1 we measured how many children could identify numbers and consonants. In grade 2 we tested consonants, and also measured how many could read a one-syllable word.

It seems clear that the program had no impact on reading levels in these grades. The control group got the largest total increase, largely because by chance, at the beginning of the year far fewer students in group C knew the consonants than in groups A and B. They had a bigger increase, though they still ended up behind groups A and B.

Where reading levels are this low, we see no benefit to introducing the reading program. We expect that our new “I Can Read” series will help these students, and we’ll continue these evaluations, to see if it does.



Traditionally, rural Lao children learn the alphabet from a teacher and a blackboard. It can be difficult for westerners to grasp just how rare books have been here but this picture makes the point: The girl in the middle is holding her book upside down, and doesn't seem to realize it yet.

taking several steps to improve implementation; and by adding “I Can Read” books that help children at all grade levels who cannot read enough to benefit from even the easiest books – which is a high majority of the total students). With these results, we think we can get a district to allow a test in which more time is allowed for reading, perhaps as much as 30 minutes in the morning, and 30 minutes in the afternoon.

We will also revisit schools more often. We expect to go to fewer schools in the coming year than in the past year, and spend more time monitoring and adjusting the program to see what works best.

The “I Can Read” books, which offer a new way for students to learn reading skills, will be an additional challenge because again, we're asking teachers to do something different. However, we've told a handful of teachers about it, and they're excited to see the books. They know the present system isn't working, they just don't know anything else.

Further detail

Now here's a level of detail and numbers that was too much for the newsletter, but we're glad you're interested.

Putting a number on the results

Scoring a multiple-choice test is pretty straightforward. (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 get a 10% score. But no alternative seems better, and since these situations should arise equally in each group, we scored by following the obvious route, and can see no reason that should affect the comparative results.)

If that was unexpectedly complicated, compiling the results was even more so. There were several choices to make, and each one affected the results.

For example: Should we measure improvement by increase by subtracting the baseline score from the followup? Or as a percentage increase over the baseline scores? The percentage approach sounds good, but at the extremes, it gets misleading. A school that went from 1% to 10% has a 900% increase, while one that goes from 2% to 10% shows only a 500% increase, but in reality, there's no real difference.

I tried out a few options and it was reassuring to see that in every case, group A improved between 24% to 30% more than group C, which means the big picture was not dependent on these choices.

Changes from our original plan

Several things didn't go as expected. For those who truly love detail, here are things that changed, with more significant ones first:

Written-vocabulary test dropped: Our first vocabulary test gave a word, and asked students to select the best synonym from a list of 4 words. We sensed a lot of confusion from students about this, and indeed there was. They had probably never seen such a test before. Often they picked a word with an opposite meaning instead of a synonym: For the synonym of "worried" 35% of students chose "comfortable" and only 33% chose "concerned". (The tests were in Lao of course; these are the closest English equivalents.) We changed to a picture vocabulary test after that week, using a picture as the "question". (Sample on page 14.) This was intuitively much easier for students to understand. I did not include the original written-word vocabulary test in our tabulations because, given this confusion, it didn't provide an accurate measurement of reading ability. (For the record: On written-word vocabulary test, Group A improved 10% more than Group C, while Group B improved 33% less. We would expect Group B to fall in between A and C, and usually it did, suggesting that this test wasn't providing an accurate measurement.)

We changed the way that we count oral reading of a single one-syllable word. For the baseline tests, we held up common one-syllable words, in random order. After three words in a row that nobody could read, we considered that the remaining students could not read a word.

From the baseline testing, we saw that students are much slower to learn the complex vowels than simple vowels. The words we showed were common, but about 40% of them used complex vowels. If we happened to show 3 complex vowels in a row, we'd reach the no-one-who-is-left-can-read point sooner than we should. Most likely this happened in about a half-dozen schools, but after the fact, there was no way to adjust for it. Repeating the same technique would have just compounded the error.

For the follow-ups, we had two sets of cards. First we showed those with a complex vowel, and recorded how many could read at that level. Then we continued with easier words, and recorded how many could read one of them. We used this second, larger number when comparing to the baseline tests. So the follow-up tests, on average, will show a slightly higher reading level than if we had used the same technique as for the baseline.

Missed follow-up tests: Somehow, the testing teams didn't get the message to do the consonant test in grade 2, so in all 40 schools, we only have one test (oral vocabulary) for grade 2.

Through my error, we gave the wrong follow-up test to grade 4 in the first district (20 schools). We gave an oral vocabulary test in September, but a written test in May. We could have used the scores anyway, since the tests measured the same thing but with different methods, but we decided not to. Had we used them, the level of improvement would increase by about 0.5 percentage points.

In 9 cases (of 320 tests used), the team gave the wrong test, or missed a test, or didn't record information properly, so we had to compile results without those.

Students who didn't understand the multiple choice concept: Midway through entering the baseline tests, we decided that if a student checked the same box (usually #1) for every question, on a test with 10 or more questions, it meant they didn't understand the concept at all, or couldn't read well enough to even try. We gave these sheets a different coding, and we counted them the same as a wrong answer. This lowered the scores, but in a way that was randomly distributed among the three groups, so any impact on the comparative results should be minor, and we decided not to spend the time to re-enter tests that had already been entered.

Double-entry of answers: Every answer to every question got entered into our database: 87,161 answers in all. I set up the system so that a second person in our office could re-enter each test, as a way to ensure accuracy.

We began doing it that way, but didn't continue, for several reasons. (1) The error rate was very low; on only 0.4% of questions did the two staff members enter different answers. Since only one of those was typed wrong, that means about 0.2% were wrong, and often these were differences of interpretation: "is that check-mark erased, or not?" On another 2% of the tests, students who didn't seem to understand the multiple choice concept were coded differently because we changed our procedure, as noted above, but the effect on the final comparative scores would be negligible. (2) Longnee, who would be doing the second entry, has been busy working on some books for the "I Can Read!" series, and we're eager to have them done for September. (3) There were other problems with this evaluation, most notably that we don't really know how good implementation was at each school. So even with perfect entry, we won't know just what we've measured, and the program will be different next year anyway. (4) This is a self-evaluation, so whatever we do, it won't have the

credibility of an independent study. It may help convince a researcher that this program merits such a study.

How well did these tests measure reading?

We deliberately looked for several different ways to measure reading levels, since we weren't sure what would be most accurate. We were told by various sources, as we made plans, that children had experience with multiple choice tests, so for grades 3-5, we used multiple-choice more than other approaches. But many students did not understand the concept, and marked every answer, or always marked the same one, or put marks in between the boxes.

Despite this, it seems clear that taken together, the tests did a reasonable job of showing relative differences in improvement. For 4 of the 5 tests, the average score in the schools that got a reading program (groups A and B) was higher than the control group.

Where it worked best

The program worked dramatically better in some schools than in others. Any study of this sort is likely to show some variation, but it seems like there's a X-Factor that affects results, and which we haven't identified yet. Some possibilities:

- **What percentage of students speak Lao at home?** Obviously, those who speak Lao at home will learn to read it more readily. For reasons that aren't so obvious, it also seems that these students get more benefits from SSR, but it's not a strong correlation.

- **Do neighboring villages speak Lao?** The poorest results for SSR came from ethnic minority schools in areas where there are few native Lao speakers anywhere close. Ethnic minority villages that are close to Lao-speaking villages generally showed better results. Probably these villagers have more opportunity to learn Lao in the course of daily interactions.

- **Implementation.** Knowing whether a school implemented the program is more difficult than it might seem. We cannot reach by many of them by telephone, and teachers often give the polite or expected answer. Two members of our staff were driving near two of the villages a few weeks ago, so they stopped to talk with some children, planning to ask casually if they had read books during class. They got nowhere: the children couldn't speak Lao well enough to communicate, thus emphasizing the language issues involved, but not helping on this question.

- **Quality of implementation.** Beyond the issue of whether they did it, some teachers might have found ways that were especially effective. We know that some teachers let children to go to school early and read. Whether they provided 5 to 10 minutes of classroom time, or 20 minutes, would make a big difference. We can identify the schools and

classrooms that got the most benefits from the program, and we'll try to learn more about what they might have done that was special.

▪ **Location.** Some schools are about a 2-hour drive from Vientiane; others are several hours from even a small town. SSR was more helpful for those closer to Vientiane. These villages are also more likely to speak Lao, or to have neighboring villages that do, so it's difficult to break out whether location itself matters. But it could. Being closer to Vientiane, teachers probably have more exposure to new ideas, and might have been more willing to implement the program with enthusiasm.

That's a lot of variables, some of which are hard to quantify, and by the time we break down a sample size of 40 schools, we just don't have enough data to draw good conclusions. Rather than trying to squeeze more out of this study, we're going to focus on several major improvements in the program, and some minor tweaks in our testing procedures, and watch for more insight about "the X-Factor".

Other observations

A big surprise during this process is that while the original point was to measure the impact of SSR, the greater value, so far, has been to help us improve the program. A few other notes:

Why were results poor for the Aesop's Fables reading test? We'd like to know, too. It's very plausible that in fact, SSR did *not* help significantly with comprehension of stories of this length. Given weak reading skills, 10-15 minutes a day, perhaps for just a few months, perhaps simply wasn't enough to make a big difference in comprehension of a multi-paragraph story.

In addition, with these sample sizes, some numbers will just fluctuate randomly. Group B did worse than groups A and C on this test; but better than groups A and C on the Read-Aloud test. Both tests were given to the same group of 5th-grade classrooms. That's why we went to 40 schools, and gave many tests in each school: A bigger sample and bigger variety of tests gives a more accurate picture.

Complex vowels - again. Our initial tests hinted at something we hadn't considered before: That the complex vowels are a big stumbling block for young readers. (Complex vowels consist of 2 or 3 simple vowels that combine to make a new sound. Often these vowels are not together. Several complex vowels consist of a symbol before one or two consonants, a symbol above them, and a symbol after them. Each symbol has a vowel sound individually, but together they make a new sound, which is heard after the consonant.)

In May, we found that 26% of 2nd-grade students could read a word with a simple



How the evaluation helped: We made a last-minute change to our “I Can Read” series. The vowel appears in red here, so the student can see that three characters are working together.

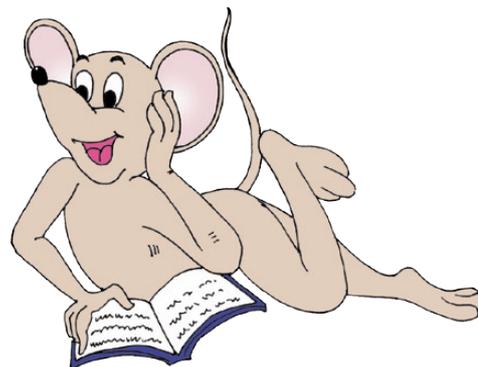
vowel, but only 12% could read a word with a complex vowel. These results gave us another idea: In the “I Can Read” books, the complex vowels will print in red ink when they first appear, so the child can see that they are a single unit.

Next evaluations

We will continue these evaluations next year. Now we can show that in grades 3 to 5 SSR will help rather than hurt, so we hope to get permission for longer reading times. We’ll also require a firm commitment from each school to implement the program, possibly written instructions from the district telling them to do so, and we’ll monitor more closely to be sure it happens.

We may add a “Lao language” test for some grades. Right now we ask how many students speak Lao at home. Knowing how many can understand Lao will help us better understand this important factor.

Thank you!



We're not done yet, but we feel this is an important milestone. We cannot thank all the supporters who have made it possible, but we'd like to highlight a few who not only provided significant and repeated financial support, but also understand what we're doing, and why it's important:

Planet Wheeler

Global Fund for Children

Manoj Paul

Cubit Family

CLSA Chairman's Trust

Peggy Horn

Denis and Helen Weily

Bengier Foundation

Charles Felsenthal

Christopher Thomas and Connie Potter, Lyle and Agnes Schaller, Jane Burren, Yatai Food Cart, Chris Ashton & Emma Morton, Brian & Gerry Warren, Charlie Weeks, Save the Children, Tammachat Natural Textiles, Terre des Hommes Germany, Doug Ewart & Judith Keene, Fiona Douglas, the Robert family, Joan Kelly, John and Jean Burton, Glenn Auve, Hugh Hunter, Julie McIntyre, basmati, Deborah Deitsch-Perez, Melbourne Girls Grammar School, U.S. Embassy, John & Robin Millar, Leah Camhi, Nick Hubbard, Josephine Ong Kui Hua and Ong So Eng, Raymond Faulkner, Lin Palmer, Miki & Anja Striewski-Bonizzato, Jackie Cunliffe. Zena Carter