



**Evaluative Report
Department of Financial Institutions Program
Washington State**

September, 2009

Eric A. Hagedorn, Ph.D.

Hagedorn Evaluation Services
El Paso, TX

Introduction

The purpose of this study was to measure the effectiveness of a program called Money Savvy Kids[®] on the attitudes and knowledge of children in schools in Washington State.

Money Savvy Kids is curriculum developed by Money Savvy Generation of Lake Bluff, Illinois. The curriculum includes eight lessons:

- The History of Money
- Where Does Money Come From?
- Kids Can Earn Money Too!
- Saving Money and Bank Field Trip
- Spending Money
- Donating Money
- Investing Money
- Family Money Press Conference

An important part of Money Savvy Kids curriculum is the Money Savvy Pig.[®] This is a four slot piggy bank. It provides teachers and parents with a fun and interesting way to introduce children to saving, spending, investing, and donating. Each child participating in the program receives a Money Savvy Pig. During the 2008-9 school year, over 100 elementary school classrooms received Money Savvy Kids[™] materials and curriculum training. Training was provided via the participants' use of self-study materials. They were asked to implement the program in their classrooms and to use a pre-and post test with the students.

To investigate the effectiveness of this program, Dr. Mark Schug of the Center for Economics Education at the University of Wisconsin – Milwaukee, developed survey (see Appendix A) measuring student beliefs about savings habits, handling money, the role of business, etc. This survey has been used in each subsequent evaluation study since the first such study at the end of the 2003-2004 school year. This study was featured in the academic journal *The Social Studies* in Spring 2005 (Schug & Hagedorn, 2005). Several items were reworded for greater simplicity and thus clarity in Fall 2008. This updated survey was given to the Washington students before receiving their Money Savvy Pigs and after they had completed their training. The analyses this academic year were based upon 988 pre-tests, 883 post-tests, and 709 tests where we could match individual pre-tests with individual post-tests. This is more than double the number of matched tests from 2007-2008.

This report presents the interpretation of the results of those surveys followed with extensive supporting analysis.

Conclusions

Overall, the aggregate data indicate that the Money Savvy Kids program was effective in positively affecting students' attitudes and knowledge about spending, saving and investing money. The paired (matched) samples data for these elementary students indicates statistically significant improvements on all 10 items. Items 4 and 5 were reworded for greater simplicity and clarity this past year. Interestingly this may have led to our detecting a statistically significant improvement on both items that was not detected during the 2007 – 2008 school year.

The item for which there was the most change (as indicated by medium effect size) was item #6. Item #6 suggests that it is best to keep the money you save in your room. After instruction 31.4% fewer students agreed with this and 25.6 % more disagreed with this. The percentage of students who were unsure increased slightly by 6.2%.

Students also improved with nearly a medium effect size (0.49) on Item 7 this year. Item 7 refers to the certainty with which one can make money on the stock market. Perhaps current events assisted student learning on this one

The remainder of the statistically significant changes had small effect sizes, but the measured changes were large enough to indicate statistically they were not likely to have occurred by chance and thus may be attributed to the curriculum and the teachers' use of it. In addition, as many of the positive attitudes towards saving money and making donations, are counter-popular culture, reinforcing them is certainly a value added.

In this evaluator's professional opinion, these data indicate that the Money Savvy Kids curriculum continues to positively impact the financial attitudes and understanding of the children who participated in this study from Washington State. In addition, changes to procedures for administering and collecting completed tests as well as the wording changes, seem to be yielding far more and far better data. Matchable tests are always better than independent samples and better worded instruments enhance validity. Overall, these results are consistent with those found in previous studies in both urban and suburban Chicago, in Cleveland, Ohio, and in North Dakota. What these studies collectively suggest is that the Money Savvy Kids curriculum is effective with a wide variety of English reading students.

Methodology

The Money Savvy Kids™ Assessment is a 10 item, Likert scale instrument. A three point response format was used: “agree” (with a value of 3), “unsure” (with a value of 2) and “disagree” (with a value of 1). Dr. Schug had a literacy expert check the questions for roughly a second grade reading level.

The completed pre and post-tests include the participating students' names. This allows for matching individual pre and post-tests. Once matched and recorded, either a paired-samples t-test or the non-parametric Wilcoxon Signed Ranks test would be performed on the data to determine if student responses changed from pre to post in a statistically significant manner. The paired samples t-test is appropriately used if the data did not differ significantly from a normal distribution. Normality is determined using the Kolmogorov-Smirnov test of normality (with Lilliefors correction) and the Shapiro-Wilk test. If the data do differ significantly from the normal distribution, one uses the non-parametric Wilcoxon Signed Ranks test.

Any statistically significant changes from pre to post, will be identified and interpreted. A statistically significant difference in means from pre to post indicates the likelihood that such a difference in mean in the population would occur by chance. For instance, an increase of mean score on Item #3 of .31 (on a scale of 1 to 5) occurs by chance only once in a thousand, as indicated by a p value equal to .001. While this information implies statistical significance (likelihood of occurring by chance), it says nothing about “how big” or “how important” a change of .31 is. To begin to understand these issues, one calculates effect sizes. The effect size

is essentially the ratio of the change to the standard deviation of the change scores. If the standard deviation of the change scores for Item #3 were around .3, the effect size would be about 1, indicating the change was roughly one whole standard deviation. In the literature, such an effect size is considered “large” (Kirk, 1995). If the standard deviation of the change scores was around 3 (indicate great variability in student responses to Item #3), the effect size would only be .10 – representing a change of about 1/10th of a standard deviation. This effect size is considered “small,” even though the likelihood that such a change occurred by chance is very unlikely.

More pre-tests were received than post-tests. This is a common occurrence in year long external curriculum projects: by the end of the academic year teachers may forget to administer post-tests or feel to pressed for time to do so. Nevertheless, this academic year we could match 709 individual pre and post tests from 29 distinct schools and 51 individual teachers. This number of matched tests is more than double the total number of matched tests from 2007-2008. Because of this, there was no need to further investigate the data using independent samples tests (such as the independent samples t-test or the Mann-Whitney U test).

Results

Matched Tests: Mean Item Changes

Item response means and standard deviations were calculated for the combined group of elementary students (N=709). These are provided in Table 1. When the sample size (“N”) is less than 709 for a particular item, it indicates that a certain number of students left this item blank. The data for every single item differed from normal with a statistical significance less than 1 in 1000 as determined by both the Kolmogorov-Smirnov and Shapiro-Wilks tests. Because of this, the Wilcoxon Signed Ranks test was used to determine if there were any statistically significant changes from pre to post.

Table 1. Descriptive statistics for paired samples data elementary students.

	N	Pre	SD	Post	SD	Desired change/achieved
Item 1	709	2.69	0.569	2.86	0.414	Increase, yes
Item 2	707	1.63	0.817	1.32	0.648	Decrease, yes
Item 3	705	1.54	0.792	1.28	0.613	Decrease, yes
Item 4	709	2.62	0.646	2.81	0.492	Increase, yes
Item 5	707	1.60	0.835	1.49	0.790	Decrease, yes
Item 6	707	2.61	0.718	2.04	0.911	Decrease, yes
Item 7	708	2.18	0.677	1.81	0.846	Decrease, yes
Item 8	706	2.62	0.613	2.79	0.505	Increase, yes
Item 9	709	2.77	0.541	2.83	0.464	Increase, yes
Item 10	709	2.05	0.826	1.75	0.876	Decrease, yes

Table 2. Significantly changed item response averages and effect size of changes.

Item	Z value	2-tailed significance	Effect size	Described Effect Size
1. I know a lot about how to handle my money.	-6.258	0.000	0.33	Small
2. Saving money is greedy.	-7.957	0.000	-0.41	Small
3. It is important to have the things I want when I want them.	-6.899	0.000	-0.35	Small
4. It is important to save for the things that I want to buy in the future.	-5.843	0.000	0.33	Small
5. I want to spend the money I earn right away.	-2.332	0.013	-0.13	V. Small
6. It is best to put the money you save in your room at home.	-11.987	0.000	-0.68	Medium
7. When I invest in stocks, I will always make money and never lose money.	-9.260	0.000	-0.49	Cusp of medium
8. Business people help others by providing them with goods and services.	-5.709	0.000	0.31	Small
9. It is important for families to keep money in real banks.	-2.488	0.015	0.13	V. Small
10. When I donate money it helps others but doesn't help me.	-6.650	0.000	-0.35	Small

What Tables 1 and 2 tell us about student responses to individual items.

In general these two tables show that there were statistically significant improvements in student understanding on all ten of the items on the assessment. The individual item changes are described below in terms of averages based on the rating scale: 3 indicating agreement, 2 uncertainty, and 1 disagreement. Following these written descriptions is a table showing how student responses changed from pre to post in terms of percentages of students.

The average response of the students to Item #1 changed from 2.69, leaning towards agreement to 2.86, more strongly agreeing. This indicates an improvement in student self-confidence regarding the proper handling of money. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The .33 effect size indicates that this improvement is one third of an average standard deviation in size. Cohen considers this a “small effect.”

The average response of the students to Item #2 changed from 1.63, leaning towards uncertainty, to 1.32, which is leans towards disagreeing. This indicates an improvement in student understanding, because it is appropriate for students to disagree with the notion that saving money is selfish. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The -.41 effect size indicates that this improvement is roughly 41% an average standard deviation in size. Cohen considers this a

“small effect.” The minus sign indicates that the average score decreased from pre to post (which is appropriate for this item).

The average response of the students to Item #3 changed from 1.54, leaning towards unsure, to 1.28, which is more strongly disagreeing. This indicates an improvement in student understanding, because it is appropriate for students to disagree with the notion that saving money is selfish. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The $-.35$ effect size indicates that this improvement is just under one third of an average standard deviation in size. Cohen considers this a “small effect.” The minus sign indicates that the average score decreased from pre to post (which is appropriate for this item).

The average response of the students to Item #4 changed from 2.62, leaning towards agreement to 2.81, more strongly agreeing. This indicates an improvement in students’ perceptions that you should save for the future. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The $.33$ effect size indicates that this improvement is one third of an average standard deviation in size. Cohen considers this a “small effect.”

The average response of the students to Item #5 changed from 1.60, leaning towards unsure, to 1.49, which is more strongly disagreeing. This indicates an improvement in student understanding, because it is appropriate for students to disagree with the notion you should spend your money immediately. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 13 in 100 times. The $-.13$ effect size indicates that this improvement is about one eighth of an average standard deviation in size. This is a very small effect size. The minus sign indicates that the average score decreased from pre to post (which is appropriate for this item).

The average response of the students to Item #6 changed from 2.61, on the agreeing side of unsure, to 2.04, which, on average, indicates uncertainty. This actually indicates an improvement in student understanding, because the average dropping indicates more students disagreeing with this item which is appropriate for this item (saving money in your room is the best method of saving). The exact two-tailed significance implies that this change in average score could only have occurred by chance less than 1 out of 1000 times. The -0.68 effect size indicates that this decrease in score is roughly 68% of an average standard deviation in size. Cohen considers this a “medium effect.”

The average response of the students to Item #7 changed from 2.18, on the agreeing side of uncertain, to 1.81, which is now on the disagreeing side of uncertain. This indicates an improvement in student learning because it is more appropriate for students to disagree with the idea that investing in the stock market always pays off. The two-tailed significance implies that this change in average score could only have occurred by chance less than 1 out of 1000 times. The $-.494$ effect size indicates that this improvement is roughly half of one average standard deviation in size. Cohen considers this a “small effect,” although it is very close to the cutoff for a “medium effect” at 0.50.

The average response of the students to Item #8 changed from 2.62, leaning towards agreement to 2.79, more strongly agreeing. This indicates an improvement in students’ perceptions that business people help others and how they do so. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times.

The .31 effect size indicates that this improvement is just less than one third of an average standard deviation in size. Cohen considers this a “small effect.”

The average response of the students to Item #9 changed from 2.77, leaning towards agreeing, to 2.83, leaning even more towards agreeing. This indicates an improvement in student understanding because more students agree that it is important for families to keep money in real banks. The two-tailed significance implies that this change in average score could only have occurred by chance 15 in 1000 times. The .13 effect size indicates that this improvement is roughly 14% of an average standard deviation in size.

The average response of the students to Item # 10 changed from 2.05, very close to uncertain, to 1.75, which leans towards disagreeing. This indicates an improvement in student understanding because more students should disagree that donating money only helps the recipient. The two-tailed significance implies that this change in average score could only have occurred by chance 1 out of 1000 times. The -.35 effect size indicates that this improvement is just over one third of an average standard deviation in size. Cohen considers this a “small effect.”

These interpretations are also presented in terms of the changes in percentages percentages of students picking each possible choice before and after being taught in Table 3.

Table 3. Percentages of chosen responses to selected items on matched pre and post-measures.

Item	Response	% students pre-test	% students post-test	Comments
1	1 disagree 2 unsure 3 agree	5.5 19.7 74.8	2.4 9.6 88.0	Fairly large percentages believing they know how to handle money before and after participation. 10.1 % fewer unsure after. 13.2 % more students agreeing that they can properly handle money after participation.
2	1 disagree 2 unsure 3 agree Missing	59.2 19.2 21.6	78.4 11.1 10.2 0.3	19.2% more students disagreeing that saving is greedy. 8.1 % fewer unsure. 11.4 % fewer students agreeing that saving is greedy after participation.
3	1 disagree 2 unsure 3 agree Missing	64.9 16.1 18.8 .3	80.3 10.9 8.6 .3	More than half disagreeing with immediate gratification to begin with. 15.4 % more students disagreeing with immediate gratification, after instruction. 5.2 % fewer unsure. 10.2 % fewer agreeing with immediate gratification.
4	1 disagree 2 unsure 3 agree	9.0 20.0 70.9	4.2 11.0 84.8	4.8 % fewer students disagreeing after instruction that saving for 13.9 % more agreeing.
5	1 disagree 2 unsure 3 agree Missing	62.8 14.4 22.7 .1	69.3 12.0 18.6 .1	6.5 % more students disagreeing, 2.4 % fewer unsure, and 4.1 less agreeing with the statement that they want to spend their money right away.
6	1 disagree 2 unsure 3 agree Missing	13.8 10.9 75.0 .3	39.4 17.1 43.6	25.6 % more students disagreeing that you should keep money in your room at home. 6.2 % more unsure. 31.4 % fewer agreeing with this after participation.

7	1 disagree 2 unsure 3 agree Missing	15.5 51.1 33.3 .1	47.2 24.8 27.9	31.7 % more disagreeing that you always make money on stocks. 26.3% fewer unsure. 5.4 % fewer agreeing with this. Most of those disagreeing with this after instruction came from the unsure category.
8	1 disagree 2 unsure 3 agree Missing	6.9 24.5 68.3 .3	4.4 12.4 83.1 .1	2.5 % fewer disagreeing that business people help others with goods & services, although very few disagree with this before or after. 12.1 % fewer unsure. 14.8 % more students agree with this after participation.
9	1 disagree 2 unsure 3 agree	5.6 12.1 82.2	3.7 9.7 86.6	1.9 % less disagree that it is important for families to keep money in real banks, although a small percentage do so. 2.4 % less unsure and 4.4 % more agreeing. While there is a small change in the correction direction, over 80% knew the correct response before participation.
10	1 disagree 2 unsure 3 agree	31.9 31.6 36.5	54.2 16.9 28.9	22.3 % more disagreeing that saving only helps the saving individual. 14.7 % fewer unsure. 7.6 % fewer agreeing.

References

- Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (Third ed.). Pacific Grove: Brooks/Cole Publishing Company.
- Schug, M. C., & Hagedorn, E. A. (2005). The Money Savvy Pig™ goes to the big city: Testing the effectiveness of an economics curriculum for young children. *The Social Studies*, 96(2).