



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

September, 2008

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 2007-9 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). This survey was given to the Washington students before receiving their Money Savvy Pigs and after they had completed their training. This report presents the analysis and interpretation of the results of those surveys.

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 7 out of the 8 items appropriate for this type of analysis. The independent (unmatched) samples data indicated statistically significant improvements on 8 out of 10 items. The only item not changing significantly from pre to post in either the matched or unmatched data was Item #5 (The thing I enjoy most about earning money is getting to spend it right away). This may be offset by the students' response to Item #4 (I believe it is important to save money for the things I want to buy in the future). The majority of students agreed with Item #4 to begin with and a small, but significant number more agreed with this item after instruction. The students may honestly enjoy having money to spend right away, but they are at least aware of the value of saving.

The items for which there was the most change (as indicated by medium effect size) were items #6 (matched samples analysis) and #8 (unmatched samples analysis). Item #6 suggests that it is best to keep the money you save in your room. After instruction 22.3% fewer students agreed with this and 20.7 % more disagreed with this. The percentage of students who were unsure increased slightly by 1.5%. Item #8 states the belief that business people help others by providing goods and services. Instruction really seems to have impacted students who were unsure: 16.6 % fewer unsure. 3.2 % of those originally disagreeing with this item joined the 16.6% for a total increase of 19.8 % more students agreeing with this item.

The remainder of the statistically significant changes had small effect sizes, but the students began with fairly good responses to items 1, 2, 3, 4, and 8 (see Table 3), so this is not something to be concerned with. In addition, as many of the positive attitudes towards saving money and making donations, are counter-popular culture, reinforcing them is certainly warranted.

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. 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. An additional assumption for using either the paired-samples t-test or its non-parametric equivalent is that the item means for pre-tests correlate significantly with the item means for post-tests. While students should improve from pre to post, students with less ability will typically improve less than those of higher ability and vice versa. This should lead to a correlation in scores. When this correlation does not occur, it is more statistically appropriate to use independent samples techniques such as those described in the next paragraph for unmatched tests.

Because collected pre and post-surveys include a significant number that cannot be matched (probably due to students missing class on a particular day), an independent samples analysis of all the pre-tests compared to all the post-tests, will be done as well. If the data are normal, an independent samples t-test will be used. If not, a Mann-Whitney U test will be used.

(Test and survey data can often deviate from the normal distribution due to floor effects on pre-tests and ceiling effects on post-tests. Another factor which can cause deviations from normality

are outliers – test scores that are very low or very high. There are two general approaches to dealing with non-normality: data cleaning and transforming or using non-parametric statistics. Data cleaning includes removing outliers. Data transformations involve mathematical transformations of data, such as taking the logarithms of the data, and if this generates a normal distribution, doing statistical tests on the transformed scores. This evaluator prefers to accept the data as they are and use the appropriate non-parametric tests as needed.)

Any statistically significant changes from pre to post, using whatever method, 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. Of the many pre and post-tests received 313 could be matched by name. Of these 313 students, 39 reported being in 1st grade, 111 in 2nd grade, and 163 in 3rd grade. An additional sample of 117 pre and post-tests that could not be matched were added to the 313 pre and post-tests from the matched sample and combined into an unmatched, or independent sample of 430.

Results

Matched Tests: Mean Item Changes

Item response means and standard deviations were calculated for the combined group of elementary students (N=313). These are provided in Table 1. When the sample size (“N”) is less than 313 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. An additional assumption for paired samples data is that the pre-test means for a given item correlate significantly with the post-test item. The means for Items #4 and #8 did not so correlate, so the change from pre to post for these items will be found in the unmatched, or independent samples analyses to follow.

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

	N	Pre	SD	Post	SD	Desired change/achieved
Item 1	313	2.70	.528	2.84	.398	Increase, yes
Item 2	310	1.48	.742	1.23	.565	Decrease, yes
Item 3	311	1.40	.715	1.20	.526	Decrease, yes
Item 4	313	2.62	.654	2.84	.448	Increase, yes
Item 5	312	2.42	.781	2.41	.824	Slight decrease, yes?
Item 6	312	2.52	.761	2.05	.918	Decrease, yes
Item 7	308	2.20	.712	1.85	.863	Decrease, yes
Item 8	309	2.58	.622	2.82	.478	Increase, yes
Item 9	310	2.77	.525	2.84	.443	Increase, yes
Item 10	312	1.96	.850	1.76	.874	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.	-4.119	0.000	0.30	Small
2. Saving money is greedy.	-5.180	0.000	-0.38	Small
3. It is important to have the things I want when I want them.	-4.150	0.000	-0.32	Small
6. It is best to put the money you save in your room at home.	-6.848	0.000	-0.56	Medium
7. When I invest in stocks, I will always make money and never lose money.	-5.879	0.000	-0.44	Small
9. It is important for families to keep money in real banks.	-1.960	0.050	0.14	Small
10. When I donate money it helps others but doesn't help me.	-3.314	0.001	-0.23	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 7 of the 8 items on the assessment which were analyzed using paired samples techniques. The change from pre to post on the one item that did not change with statistical significance (#5) did change in the correct direction, but bear in mind – not being statistically significant implies that this change could have occurred completely by chance. 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.

The average response of the students to Item #1 changed from 2.70, leaning towards agreement to 2.84, 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 .30 effect size indicates that this improvement is roughly 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 1.48, leaning towards disagreeing, to 1.23, 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 $-.38$ effect size indicates that this improvement is roughly 38% 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 1.40, leaning towards disagreeing, to 1.20, 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 $-.32$ 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 #6 changed from 2.52, on the agreeing side of unsure, to 2.05, 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 -0.56 effect size indicates that this decrease in score is roughly 56% of an average standard deviation in size. Cohen considers this a “medium effect.”

The average response of the students to Item #7 changed from 2.20, on the agreeing side of uncertain, to 1.85, 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 $-.44$ effect size indicates that this improvement is 44% of an average standard deviation in size. Cohen considers this a “small effect” (although approaching a “medium effect” at 0.50).

The average response of the students to Item #9 changed from 2.77, leaning towards agreeing, to 2.84, 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 5 in 100 times. The $.14$ effect size indicates that this improvement is roughly 14% of an average standard deviation in size. Cohen considers this a “small effect.”

The average response of the students to Item # 10 changed from 1.96, very close to uncertain, to 1.76, 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 $-.23$ effect size indicates that this improvement is roughly one quarter of an average standard deviation in size. Cohen considers this a “small effect.” These interpretations, as well as the statistically non-significant changes are presented in terms of 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	4.2 23.1 72.6	1.9 12.8 85.3	Fairly large percentages believing they know how to handle money before and after participation. 10.3 % fewer unsure after. 12.7 % more students agreeing that they can properly handle money after participation.
2	1 disagree 2 unsure 3 agree	64.2 20.6 15.2	82.7 11.0 6.3	18.5 % more students disagreeing that saving is greedy. 9.6 % fewer unsure. 8.9 % fewer students agreeing that saving is greedy after participation.
3	1 disagree 2 unsure 3 agree	76.2 9.3 14.5	86.0 8.4 5.6	Fairly large percentage disagreeing with immediate gratification to begin with. 9.8 % more students disagreeing with immediate gratification, after instruction. 0.9 % fewer unsure. 8.9 % fewer agreeing with immediate gratification.
4	1 disagree 2 unsure 3 agree	8.2 18.0 73.8	3.3 10.0 86.7	4.9 % fewer students disagreeing after instruction that saving for future is important. 8 % fewer unsure and 12.9 % more agreeing.
5	1 disagree 2 unsure 3 agree	23.0 18.3 58.7	19.1 17.9 63.0	3.9 % fewer students disagreeing, 0.4 % fewer unsure, and 4.3 % more agreeing with the statement that what they enjoy most about earning money is getting to spend it later on.
6	1 disagree 2 unsure 3 agree	17.3 16.4 66.4	38.0 17.9 44.1	20.7 % more students disagreeing that you should keep money in your room at home. 1.5 % more unsure. 22.3 % fewer agreeing with this after participation.
7	1 disagree 2 unsure 3 agree	15.7 53.5 30.8	44.5 26.9 28.6	28.8 % more disagreeing that you always make money on stocks. 26.6% fewer unsure. 2.2 % fewer agreeing with this. Most of those disagreeing with this after instruction came from the unsure category.
8	1 disagree 2 unsure 3 agree	6.1 37.0 56.9	3.3 11.7 85.0	2.8 % fewer disagreeing that business people help others with goods & services, although very few disagree with this before or after. 25.1 % fewer unsure. 28.1 % more students agree with this after participation.
9	1 disagree 2 unsure 3 agree	5.4 11.5 83.1	2.8 11.9 85.3	2.6 % less disagree that it is important for families to keep money in real banks, although a small percentage do so. .4 % more unsure and 2.2 % more agreeing. While there is small change in the correction direction, over 80% knew the correct response before participation.
10	1 disagree 2 unsure 3 agree	36.5 26.7 36.8	54.2 17.4 28.4	17.7 % more disagreeing that saving only helps the saving individual. 9.3 % fewer unsure. 8.4 % fewer agreeing.

Independent Samples (Un-matched)

A sample of 428 pre-tests were compared with an unmatched sample of 430 post-tests. This was done for two reasons: 1) to investigate responses to Items #2 and #4 that did not sufficiently correlate to justify using a matched sample method and 2) to use more of the graded exams and thus increase the sample size. Table 4 includes the means and standard deviations for each item in the unmatched pre and post-tests. Table 5 provides the results of the Mann-Whitney U test (the non-parametric equivalent of the independent samples t-test), and the effect sizes for any significant results. Table 6 provides an extensive description of the changes in student responses for each item from pre to post.

Table 4. Descriptive statistics for paired samples data for elementary students.

	N _{pre}	Pre	SD	N _{post}	Post	SD	Desired change/achieved
Item 1	428	2.68	.548	430	2.83	.419	Increase, yes
Item 2	427	1.51	.743	428	1.24	.554	Decrease, yes
Item 3	428	1.38	.726	428	1.20	.520	Decrease, yes
Item 4	428	2.66	.624	430	2.83	.451	Increase, yes
Item 5	427	2.36	.830	430	2.44	.793	Slight increase, yes?
Item 6	428	2.49	.773	429	2.06	.905	Decrease, yes
Item 7	426	2.15	.666	427	1.84	.841	Decrease, yes
Item 8	427	2.51	.610	427	2.82	.464	Increase, yes
Item 9	426	2.78	.531	429	2.83	.448	Slight Increase, yes
Item 10	427	2.00	.857	430	1.74	.872	Decrease, yes

Table 5. Significantly changed item response averages and effect size of changes.

Item	U value	2-tailed significance	Effect size	Described Effect Size
1. I know a lot about how to handle my money.	80065	0.000	0.31	Small
2. Saving money is greedy.	74109	0.000	-0.41	Small
3. It is important to have the things I want when I want them.	81968	0.000	-0.29	Small
4. I believe it is important to save money for the things that I want to buy in the future.	79925	0.000	0.31	Small
5. The thing I enjoy most about earning money is getting to spend it right away.	87201	0.145		
6. It is best to put the money you save in your room at home.	68478	0.000	-0.45	Small
7. When I invest in stocks, I will always make money and never lose money.	71161	0.000	-0.41	Small
8. I believe business people help others by providing them with goods and services to buy.	66001	0.000	0.57	Medium
9. It is important for families to keep money in real banks.	88877	0.276		
10. When I donate money it helps others but doesn't help me.	76666	0.000	-0.30	Small

Rather than interpret changes for every item, the results for Items #4 and #8 will be given (as they could not be analyzed as matched) and then a summary of any differences between the matched and unmatched data in how items changed from pre to post.

The average response of the students to Item #4 changed from 2.66, leaning towards agreeing, to 2.83, which leans even more towards strongly agreeing. This is an improvement of student understanding because it is appropriate for students to believe that saving money for future purchases is important. 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 31% of an average standard deviation in size. Cohen considers this a “small effect.”

The average response of the students to Item #8 changed from 2.51, just leaning towards agreeing, to 2.82, leaning much more towards agreeing. This indicates an improvement in student learning because it is more appropriate for students to agree with the idea that business people help others by providing goods and services. 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 .57 effect size indicates that this improvement is almost 57% of an average standard deviation in size. Cohen considers this a “medium effect.”

The only differences in matched versus unmatched data occurred for Items #6 and #9. Item #6 refers to it being best to keep your money in your room. The increase in the number of students who disagreed with this led to a medium effect size change in the mean for the matched data and a small effect size change for the unmatched data. As the effect size for the unmatched was 0.45 and the boundary between small and medium is 0.50, this is not a real noteworthy difference.

Item #9 refers to the importance of families keeping money in real banks. In the matched data, the change from a mean of 2.78 to 2.83 was not statistically significant. The change in the unmatched data was from 2.77 to 2.84, and this was statistically significant, although with small effect size. Bear in mind, when one sets the level of significance at 0.05 a value of 0.049 is considered significant, while one of 0.51 is not. In both cases, the majority of students knew both before and after instruction that banks were where families should keep their money.

Table 6. Percentages of chosen responses to selected items on unmatched pre and post-measures.

Item	Response	% students pre-test	% students post-test	Comments
1	1 disagree 2 unsure 3 agree	3.5 22.4 74.1	1.3 13.1 85.6	Fairly large percentages believing they know how to handle money before and after participation. 2.2 % fewer disagreeing after. 8.3 % fewer unsure after. 11.5 % more students agreeing that they can properly handle money after participation.
2	1 disagree 2 unsure 3 agree	67.3 17.6 15.0	84.2 8.7 7.1	16.9 % more students disagreeing that saving is greedy. 8.9 % fewer unsure. 8 % fewer students agreeing that saving is greedy after participation.
3	1 disagree 2 unsure 3 agree	74.1 12.5 13.4	85.9 8.4 5.8	Fairly large percentage disagreeing with immediate gratification to begin with. 11.8 % more students disagreeing with immediate gratification, after instruction. 4.1% fewer unsure. 7.6 % fewer agreeing with immediate gratification.
4	1 disagree 2 unsure 3 agree	9.6 18.5 71.9	3.2 9.9 86.9	6.4 % fewer students disagreeing after instruction that saving for future is important. 8.6 % fewer unsure and 15 % more agreeing.
5	1 disagree 2 unsure 3 agree	18.3 21.5 60.2	21.7 16.0 62.3	3.4 % more students disagreeing, 5.5 % fewer unsure, and 2.1 % more agreeing with the statement that what they enjoy most about earning money is getting to spend it later on.
6	1 disagree 2 unsure 3 agree	16.3 15.7 68.0	39.4 15.7 44.9	23.1 % more students disagreeing that you should keep money in your room at home. 23.1 % fewer agreeing with this after participation.
7	1 disagree 2 unsure 3 agree	17.7 45.3 37.0	45.8 23.5 30.6	28.1 % more disagreeing that you always make money on stocks. 21.8 % fewer unsure. 6.4 % fewer agreeing with this.
8	1 disagree 2 unsure 3 agree	7.1 27.6 65.4	3.9 11.0 85.2	3.2 % fewer disagreeing that business people help others with goods & services, although very few disagree with this before or after. 16.6 % fewer unsure. 19.8 % more students agree with this after participation.
9	1 disagree 2 unsure 3 agree	4.8 13.5 81.7	2.9 10.6 86.5	1.9 % less disagree that it is important for families to keep money in real banks, although a small percentage do so. 2.9 % less unsure and 4.8 % more agreeing. These changes, while aligned with the instructional goals, are not statistically significant. Importantly, over 80% of students agree with this item, as they should, both before and after.
10	1 disagree 2 unsure 3 agree	38.1 27.9 34.0	53.4 17.9 28.8	15.3 % more disagreeing that saving only helps the saving individual. 10 % fewer unsure. 5.2 % 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).