



**Evaluative Report
Chicago Public Schools Program**

July 28, 2014

Eric A. Hagedorn, Ph.D.

Hagedorn Evaluation Services
El Paso, TX

Executive Summary of 2013-2014 Analysis of Money Savvy Kids® Program

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 (N = 649) for these students indicate statistically significant improvements on 9 of 10 items.

The item for which there was the most change (as indicated by the largest effect size: $d = .47$) was item #4. Twenty-one percent more students disagreed that you should keep money in your room at home after instruction. About 1% fewer were uncertain about this. Item #1 had the second most change (as indicated by the next largest effect size: $d = .34$). Almost 18% more students totally disagreed with this negatively worded item: "I do not know very much about how to handle my money."

The remainder of the statistically significant changes had small Cohen effect sizes (a common method of quantifying changes in test/survey scores), but the measured changes were large enough to indicate statistically that 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 contrary to popular culture, reinforcing them is certainly a value added. The only item that did not change with statistical significance, was item #8: "It is important for families to keep money in real banks." The reason this is not troubling is that 83.6% of the students agreed with this statement to begin with.

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 in Chicago Public Schools. Overall, these results are consistent with those found in previous studies in Chicago, Washington State (over several years); 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.

Introduction & Background

The purpose of this study is to continue to measure the effectiveness of a financial literacy program called Money Savvy Kids to positively impact the attitudes and knowledge of children throughout the Chicago metropolitan area. Money Savvy Kids is a curriculum developed by Money Savvy Generation of Lake Forest, 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 the 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 2013-2014 school year, 41 teachers at 41 schools 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 we have used a revised attitudinal survey originally developed by Dr. Mark Schug of the Center for Economics Education at the University of Wisconsin – Milwaukee. Both surveys (original and revised) solicit student beliefs about savings habits, handling money, the role of business, etc. The original survey, first used for the 2003-2004 school year and until the 2011-2012 school year, was featured in the academic journal, *The Social Studies* in Spring 2005 (Schug & Hagedorn, 2005). Based upon a psychometric analysis of a large sample collected with the original survey, the survey itself was extensively revised – not in content, but wording, Likert scaling (3 to 5 point response format), item order, etc. See Appendix for the instrument. The analyses this academic year are based upon 649 matched pre- and post-tests, using this revised survey.

Methodology

The original 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 revised Money Savvy Kids Assessment is still a 10 item, Likert scale instrument, but uses a 5-point scale (“totally agree” = 5, “kind of agree” = 4, “unsure” = 3, “kind of disagree” = 2, and “totally disagree” = 1).

The completed pre and post-tests include the participating students' names. This allows for matching individual pre and post-tests. Once matched and recorded, the frequency of each response to each Likert category for each item is converted to the percentage of all responses to that item. In other words, for item 1, “I do not know very much about how to handle my money,” 37% of the students who responded initially totally disagreed with this statement (pre-test). Improvement is obvious because 67% of the students totally disagreed with this statement on the post-test. These results are provided in Table 1. Bar charts of the percentage of responses to all items, pre- and post- follow table 1. These bar charts use the same scale (a maximum percentage of 80), to allow unambiguous visual comparisons.

While this descriptive presentation of student results is valuable, a careful series of statistical analyses were conducted to determine if the changes in responses after instruction were likely to have occurred by pure chance or not. If the changes did not occur by chance, it is reasonable to infer that the instruction had an impact.

The first statistical analysis used to determine if student responses changed from pre to post in a statistically significant manner is the non-parametric Wilcoxon Signed Ranks test. The Wilcoxon Signed Ranks test is the non-parametric equivalent to the paired or dependent samples t-test. The non-parametric test is used rather than the parametric (and better known) paired samples t-test, because the paired samples t-test assumes a normal distribution of the

data, which these data are not (based upon Kolmogorov-Smirnov and Shapiro-Wilk tests of normality). This is not unusual in data coming from a 5-point Likert scale.

Any statistically significant changes from pre to post, will be identified and interpreted. A statistically significant difference in means from pre to post indicates that the likelihood that such a difference in mean in the population would occur by chance less than 5 times in a 100. For instance, a decrease of the mean score on Item #3 of .26 (on a scale of 1 to 5) occurs by chance only once in a thousand, as indicated by a p value equal to .000. While this information implies statistical significance (likelihood of occurring by chance), it says nothing about “how big” or “how important” a change of .26 is. To begin to understand these issue, one calculates effect sizes. One of the most well-known effect size calculations derives from statistician Jacob Cohen. The Cohen Effect Size is essentially the ratio of the change in mean to the standard deviation of the change scores. If the standard deviation of the change scores for Item #3 were around .26, the effect size would be about 1, indicating the change was roughly one whole standard deviation. In reality, the standard deviation was around 1.53, hence a ratio of $0.26/1.53 = 0.17$, and an effect size essentially 17% of one standard deviation. In the literature, such an effect size is considered “very small,” while an effect size of 1 is considered “large” (Cohen, 1992; Kirk, 1995). The rules of thumb are a “small” effect size is .20, a “medium” effect size is .50, and a “large” effect size is .80.

Another effect size calculated for these data is called the eta squared statistic. While it is most commonly used in analyses of variance, it can certainly be used with paired samples analyses. This effect size can be thought of as a measure of the amount of variance explained in a post-test variable (a mean post-test item score) as based upon the information from the pre-test. In this respect it is similar to the r-squared from correlational or regression studies. In light of this interpretation, a smaller eta squared might be more desirable if our desired outcome was for the students to learn more, and for their post-test scores to go up, reflecting this. Why? If how the students did on the post-test depends heavily on how they did on the pre-test, one could assume that this is due to personal characteristics (e.g. reading ability) that have not changed due to this curriculum. If the eta squared is smaller, whatever changes occurred from pre to post do not depend on factors related to the pre-test but to something that happened between the pre and post-test, namely, the Money Savvy Pig intervention. Eta squared values can be categorized, as suggested by Cohen, similarly to r-squared values: 0.01 small, 0.06 medium, and 0.14 large.

As commonly occurs, more pre-tests were received than post-tests. This often happens in year-long external curriculum projects: by the end of the academic year teachers may forget to administer post-tests or feel too pressed for time to do so. In addition, students could be absent on either the pre- or post-test administration day and in some cases the students might even use a nickname or first name only and thus be hard to match. Nevertheless, we could match 649 individual pre and post-tests from 13 distinct schools and 21 individual teachers. Because paired samples data are more informative than unmatched data from the same source, we did not analyze the unmatched data. In such a large sample, this was deemed unnecessary. In previous studies with other school districts, where both paired and independent samples data were available, the results were very similar.

Results

Percentages of students responding to each category on test

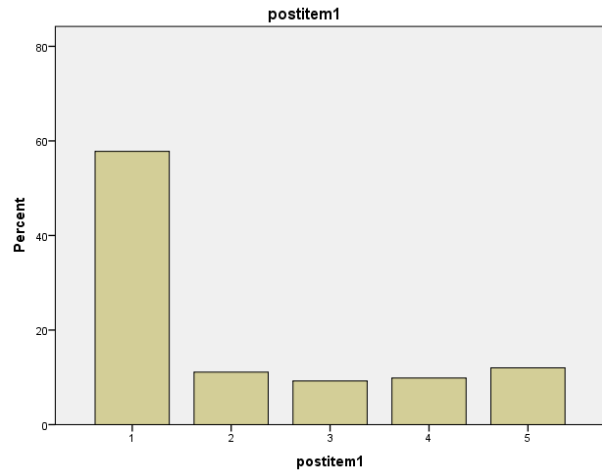
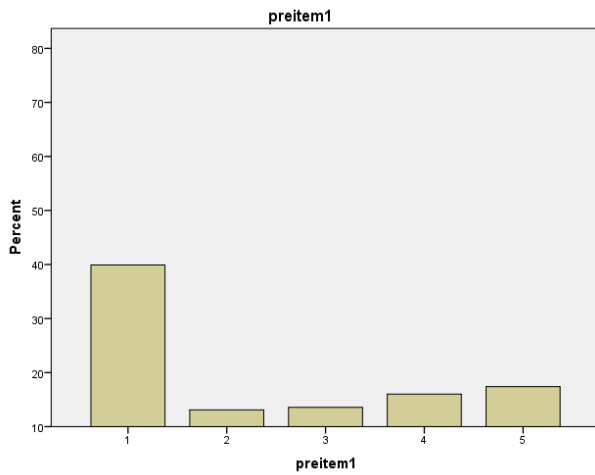
Table 1 below and the series of bar charts following, show how student responses changed from pre to post in terms of percentages of students responding to each response category (totally agreeing, kind of agreeing, unsure, kind of disagreeing, and totally disagreeing) before and after instruction.

Table 1. Percentages of chosen responses on matched pre- and post-tests.

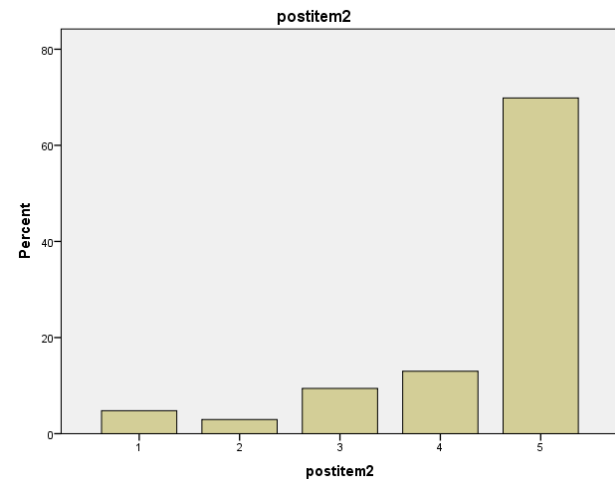
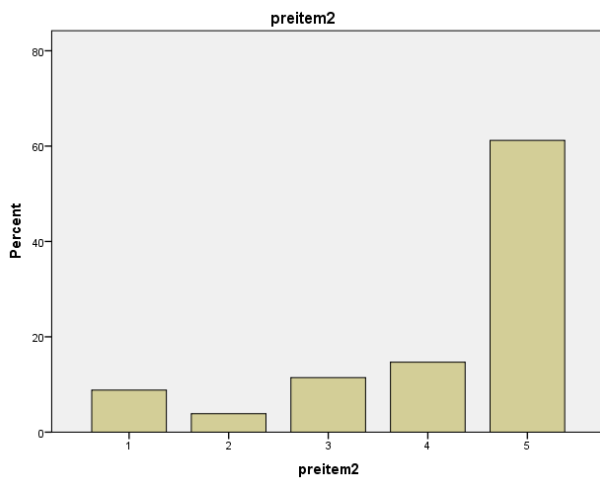
Item	Response Category	% students pre-test	% students post-test	Comments ("agreeing" is the combination of "totally agree" + "kind of agree"; "disagreeing" is the combination of "totally disagree" + "kind of disagree")
1	1 total disagree	39.9	57.8	Responses to this negatively worded item indicate that students believed that they did know about how to handle their money after participation. 17.9% more totally disagreed after, and 4.4% fewer were unsure.
	2 kind of disag	13.1	11.1	
	3 unsure	13.6	9.2	
	4 kind of agree	16.0	9.9	
	5 totally agree	17.4	12.0	
2	1 total disagree	8.8	4.8	While a majority (75.9%) believed (totally and kind of agreeing) they could save money by saving carefully before instruction, even more (82.9%) believed so after instruction.
	2 kind of disag	3.9	2.9	
	3 unsure	11.4	9.4	
	4 kind of agree	14.7	13.0	
	5 totally agree	61.2	69.9	
3	1 total disagree	47.9	56.2	More than half (60.3%) disagreeing with immediate gratification to begin with, increased 8.1% to 68.4%. Uncertainty about this dropped by 1.1% and those agreeing dropped by 7%.
	2 kind of disag	12.4	12.2	
	3 unsure	10.2	9.1	
	4 kind of agree	11.6	7.7	
	5 totally agree	18.0	14.9	
4	1 total disagree	15.1	29.5	21.1% more students disagreeing that you should keep money in your room at home. 0.7 % fewer unsure. 20.3% fewer agreeing with this after participation.
	2 kind of disag	6.2	12.9	
	3 unsure	10.1	9.4	
	4 kind of agree	12.8	11.9	
	5 totally agree	55.7	36.3	
5	1 total disagree	6.5	4.2	5.3% fewer students disagreeing after instruction that saving for future is important. 2.2% fewer unsure about this and 7.4 % more agreeing.
	2 kind of disag	5.3	2.3	
	3 unsure	10.7	8.5	
	4 kind of agree	17.9	13.8	
	5 totally agree	59.6	71.1	
6	1 total disagree	9.4	27.9	20.7 % more disagreeing that you always make money on stocks. 21% fewer unsure. The percentage agreeing with this increased slightly (0.3%) despite instruction
	2 kind of disag	5.8	8.0	
	3 unsure	43.3	22.3	
	4 kind of agree	13.3	13.6	
	5 totally agree	28.2	28.2	
7	1 total disagree	34.2	51.8	17.6% more strongly disagreeing that they are too young to set up long term goals for their money. Uncertainty dropped by 7%
	2 kind of disag	13.3	11.1	
	3 unsure	21.1	14.1	
	4 kind of agree	11.6	7.3	
	5 totally agree	19.8	15.7	
8	1 total disagree	6.9	6.5	While 83.6% agreed (strongly and kind of) with this to begin with, instruction only caused a 0.1% increase in agreement. The number totally disagreeing with this dropped .4 but the number kind of disagreeing with this actually increased 1.1%.
	2 kind of disag	2.9	4.0	
	3 unsure	6.7	5.7	
	4 kind of agree	10.4	9.1	
	5 totally agree	73.2	74.6	

9	1 total disagree	67.8	69.5	3.9 % more students disagreeing, 0.2 % fewer unsure, and 3.8% fewer agreeing with the statement that they want to spend their money right away.
	2 kind of disag	7.8	10.0	
	3 unsure	7.8	7.6	
	4 kind of agree	6.2	5.4	
	5 totally agree	10.5	7.5	
10	1 total disagree	6.8	5.3	4.3% more agreeing that donating money helps both others and themselves, although even before instruction 82.5% of the students agreed with this. 1.8% fewer are uncertain about this.
	2 kind of disag	2.8	1.7	
	3 unsure	7.9	6.1	
	4 kind of agree	12.4	10.5	
	5 totally agree	70.1	76.3	

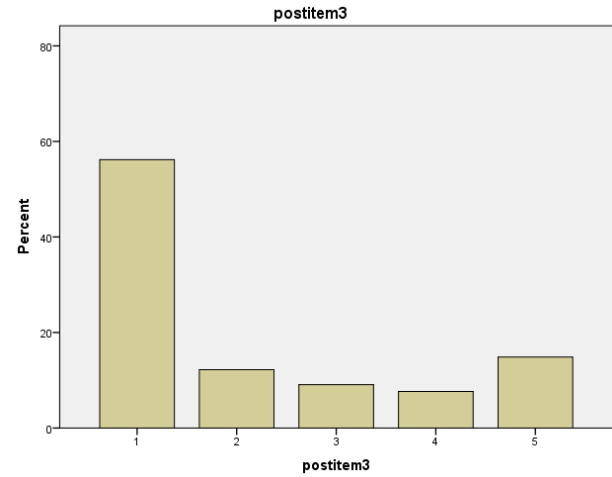
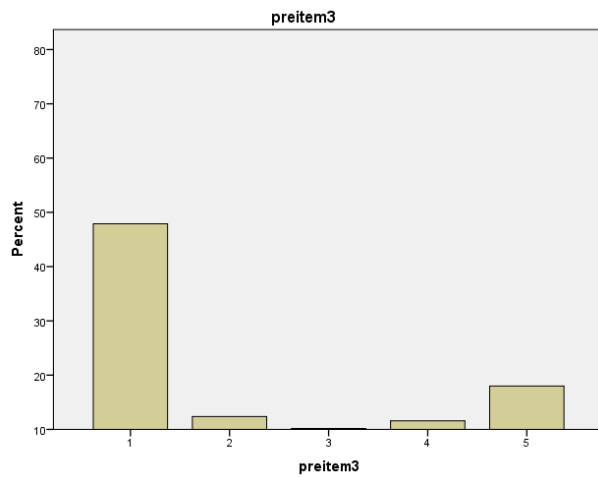
Item 1: I don't know very much about how to handle my money.



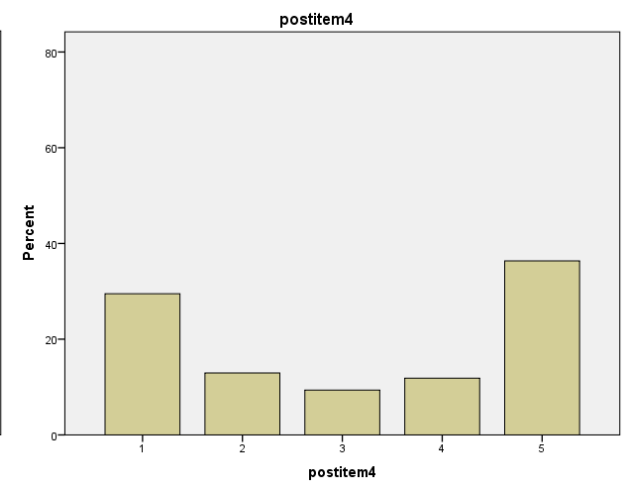
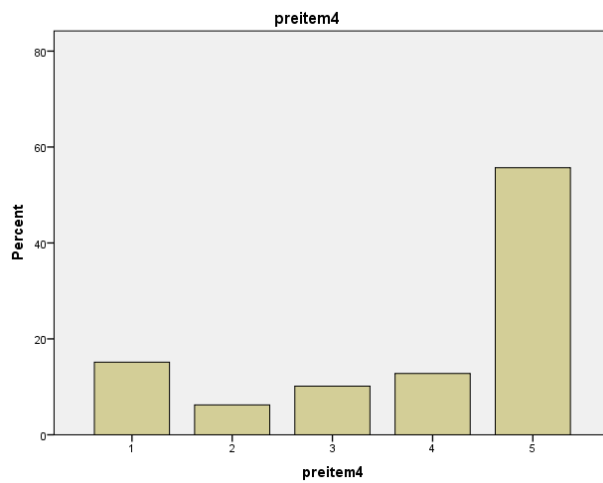
Item 2: I can save money when I spend my money very carefully.



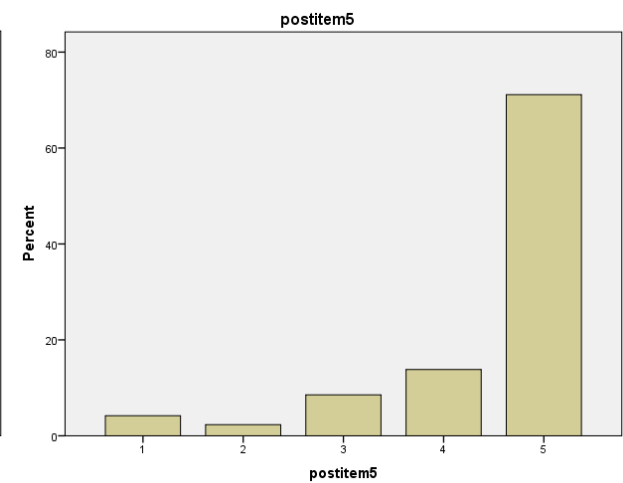
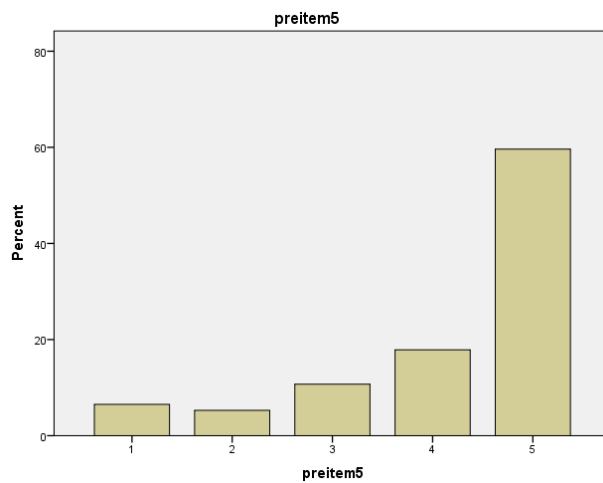
Item 3: It is important to have the things I want when I want them.



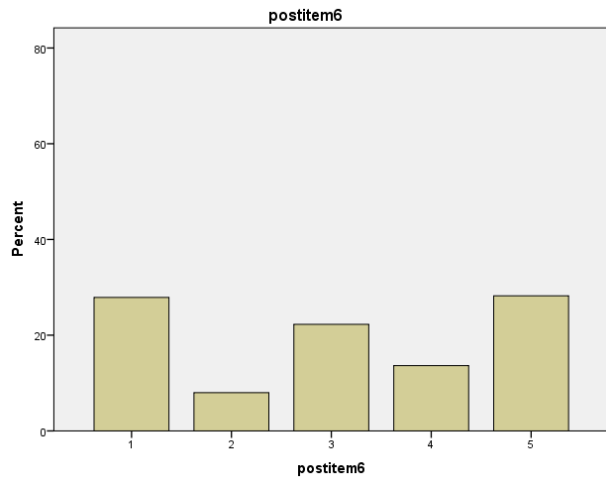
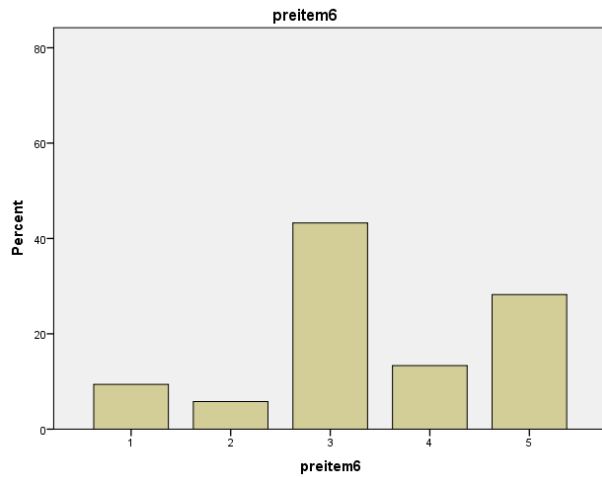
Item 4: It is best to put the money you save in your room at home.



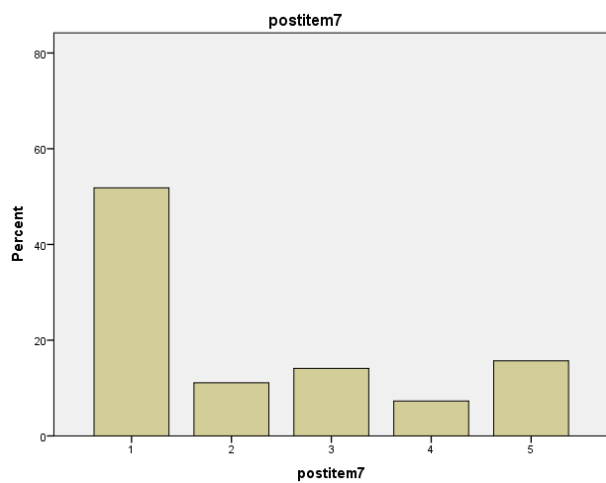
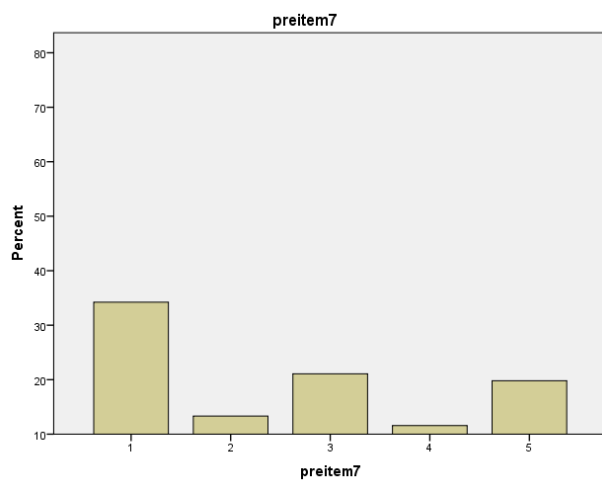
Item 5: It is important to save for the things that I want to buy in the future.



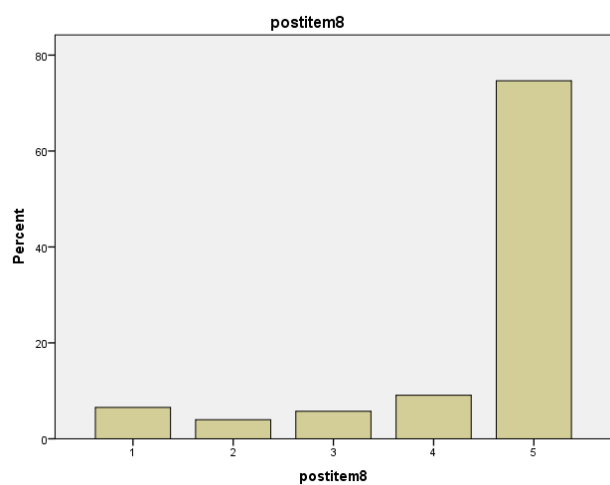
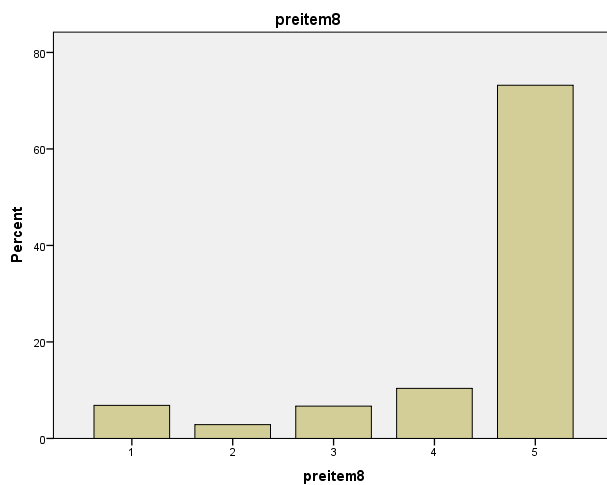
Item 6: When I invest in stocks, I will always make money and never lose money.



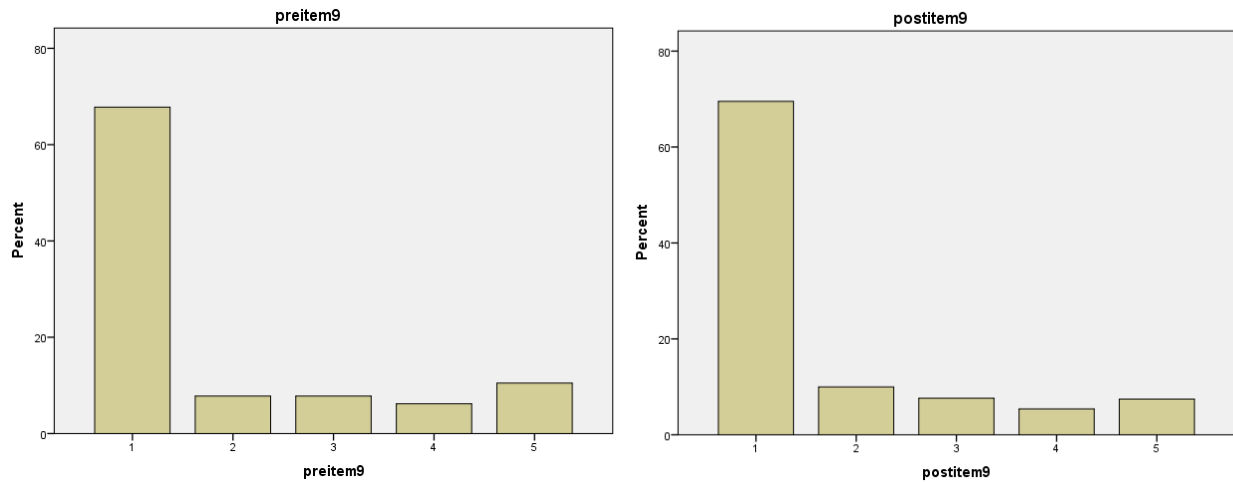
Item 7: I'm too young to need a long term goal for my money.



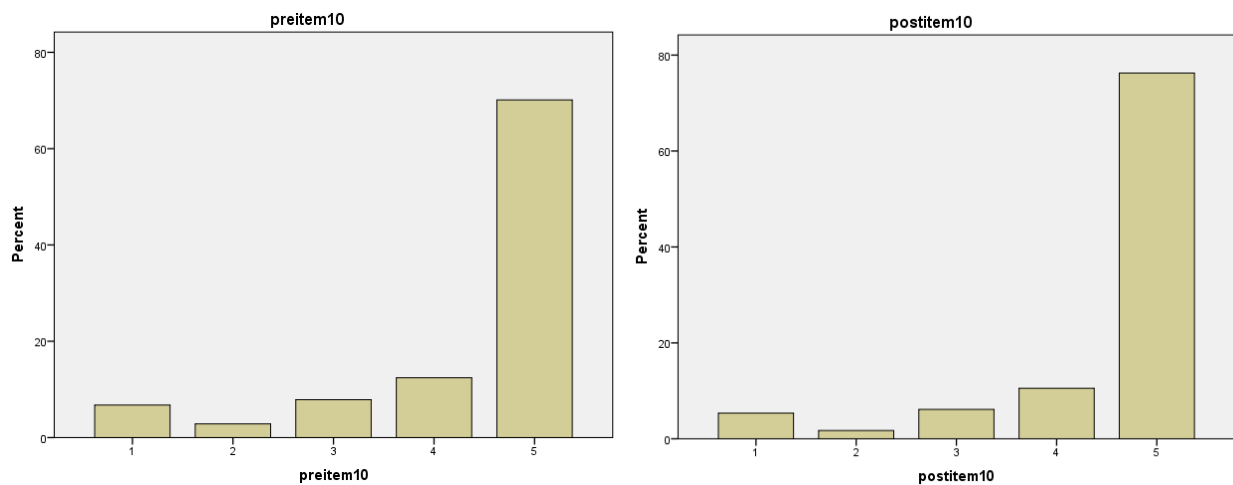
Item 8: It is important for families to keep money in real banks.



Item 9: I want to spend the money I earn right away.



Item 10: When I donate money it helps other people and it helps me too.



Statistical Analyses of Mean Item Changes

Item response means and standard deviations were calculated for the combined group of participating students for whom we had both pre- and post-tests (N=1020). These are provided in Table 2 along with an interpretation regarding desired learning change. When the sample size (“N”) is less than 1020 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 non-parametric Wilcoxon Signed Ranks test was used instead of the paired samples t-test to determine if there were any statistically significant changes from pre to post.

Table 2. Descriptive statistics for paired samples data from Chicago area students.

	N	Pre	SD	Post	SD	Desired change/achieved
Item 1	649	2.58	1.55	2.07	1.46	Decrease, yes
Item 2	647	4.16	1.29	4.40	1.09	Increase, yes
Item 3	639	2.39	1.58	2.13	1.51	Decrease, yes
Item 4	641	3.88	1.50	3.13	1.69	Decrease, yes
Item 5	644	4.19	1.21	4.45	1.03	Increase, yes
Item 6	638	3.45	1.22	3.06	1.57	Decrease, yes
Item 7	631	2.69	1.52	2.24	1.52	Decrease, yes
Item 8	627	4.40	1.17	4.41	1.18	Increase, yes
Item 9	630	1.84	1.39	1.71	1.26	Decrease, yes
Item 10	636	4.36	1.17	4.51	1.06	Increase, yes

The statistic calculated in the Wilcoxon Signed Ranks test is called the “Wilcoxon Z.” These values, along with their associated statistical significance, are included in Table 3. If significant, the Cohen Effect Sizes are calculated and interpreted. Table 4 presents the eta squared statistics and Cohen’s interpretation of these results.

Table 3. Wilcoxon Z (indicator of significant change) and Cohen effect size of changes.

Item	Z value	2-tailed significance	Cohen Effect Size	Described Effect Size
1. I do not know very much about how to handle my money.	-6.97	0.000	0.34	Small
2. I can save money when I spend my money very carefully.	-3.99	0.000	0.20	Small
3. It is important to have the things I want when I want them.	-3.87	0.000	0.17	Very small
4. It is best to put the money you save in your room at home.	-9.01	0.000	0.47	Small
5. It is important to save for the things that I want to buy in the future.	-4.87	0.000	0.23	Small
6. When I invest in stocks, I will always make money and never lose money.	-5.08	0.000	0.28	Small
7. I’m too young to need a long term goal for my money.	-5.99	0.000	0.30	Small (appr)
8. It is important for families to keep money in real banks.	-.275	0.783		Change not significant
9. I want to spend the money I earn right away.	-2.00	0.046	0.10	Very Small
10. When I donate money it helps other people and it helps me too.	-2.80	0.005	0.13	Very Small

Table 4. Eta squared statistic

Item	Eta squared	Interpretation
1. I do not know very much about how to handle my money.	0.07	Medium
2. I can save money when I spend my money very carefully.	0.03	Small
3. It is important to have the things I want when I want them.	0.02	Small
4. It is best to put the money you save in your room at home.	0.13	Medium
5. It is important to save for the things that I want to buy in the future.	0.04	Small
6. When I invest in stocks, I will always make money and never lose money.	0.04	Small
7. I'm too young to need a long term goal for my money.	0.06	Medium
8. It is important for families to keep money in real banks.	0.00	Change not significant
9. I want to spend the money I earn right away.	0.01	Small
10. When I donate money it helps other people and it helps me too.	0.01	Small

What Tables 2, 3 and 4 tell us about student responses to individual items before and after instruction. In general the first two tables show that there were statistically significant improvements in student understanding/attitude on nine out of ten of the items on the test. The individual item changes are described below in terms of averages based on the rating scale: "5" indicating total agreement, "4" indicating agreement, "3" indicating uncertainty, "2" indicating disagreement and "1" total disagreement. As mentioned in the methodology section, the third table of eta squared values indicate how much the pre-scores explain the variance in the post-scores.

The average response of the students to Item #1 changed from 2.58, leaning towards uncertainty, to 2.07 more strongly disagreeing, which is appropriate for this negatively worded item about knowing how to handle one's money. This indicates an improvement in student self-confidence regarding the proper handling of money. The two-tailed statistical significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The 0.34 effect size indicates that this improvement is 34% of an average standard deviation in size. Cohen considers this a "small effect." The eta-squared (.07) is medium.

The average response of the students to Item #2 changed from 4.16, leaning towards agreeing, to 4.40, which is further towards totally agreeing. This indicates an improvement in student understanding, because it is appropriate for students to know that savings occurs with careful spending. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The 0.20 effect size indicates that this

improvement is roughly 24% an average standard deviation in size. Cohen considers this a “small effect.” The eta-squared (.03) is also small.

The average response of the students to Item #3 changed from 2.39, on the disagreeing side of unsure, to 2.13, which is more strongly disagreeing. This indicates an improvement in student understanding, because it is appropriate for students to disagree with the notion that they should have things when they want them. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The 0.17 effect size indicates that this improvement is 17% of an average standard deviation in size. Cohen considers this a “very small effect.” The eta-squared (.02) is small.

The average response of the students to Item #4 changed from 3.88, leaning towards “kind of agreeing” to 3.13, which is on the agreeing side of uncertainty. This indicates an improvement in students’ perceptions that you should not save your money in your room. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The .47 effect size indicates that this improvement is 47% of an average standard deviation in size. Cohen considers this a “small effect,” but it is approaching a “medium effect,” at .50. The eta-squared (.13) is medium, which suggests persistence of earlier views, despite instruction.

The average response of the students to Item #5 changed from 4.19, leaning towards agreeing, to 4.45, which is more strongly agreeing. This indicates an improvement in student understanding, because it is appropriate for students to agree with the notion 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 0.23 effect size indicates that this improvement is 23% of an average standard deviation in size. This is a small effect size. The eta-squared (.04) is also small.

The average response of the students to Item #6 changed from 3.45, on the agreeing side of unsure, to 3.06, 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 (the stock market is not a certain way to make money). 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 0.28 effect size indicates that this decrease in score is about 28% an average standard deviation in size. Cohen considers this a “small effect.” The eta-squared (.04) is small.

The average response of the students to Item #7 changed from 2.69, on the disagreeing side of uncertain, to 2.24, which further towards disagreeing. This indicates an improvement in student learning because it is more appropriate for students to disagree with the idea that they are too young to have long term financial goals. 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 0.30 effect size indicates that this improvement is 30% of an average standard deviation in size. Cohen considers this a “small effect.” The eta-squared (.06) is medium.

The average response of the students to Item #8 barely changed from 4.40, leaning towards agreement to 4.41, slightly more agreeing. This indicates an improvement in students’

perceptions that their families should keep their money in banks. This slight change was not statistically significant, as this slight change could occur purely by chance roughly 78% of the time.

The average response of the students to Item #9 changed from 1.84, leaning towards disagreeing, to 1.71, leaning more towards totally disagreeing. This indicates an improvement in student understanding because more students should disagree with impulsive/rapid spending right after earning. The two-tailed significance implies that this change in average score could only have occurred by chance less than 46 times out of 1000. The .10 effect size is considered “very small,” and indicates that this improvement is only one tenth of an average standard deviation in size. The eta-squared (.01) is small.

The average response of the students to Item # 10 changed from 4.36, leaning towards strong agreeing, to 4.51 which just leans towards total agreement. This indicates an improvement in student understanding because more students should agree that donating money helps the recipient and the donator. The two-tailed significance implies that this change in average score could only have occurred by chance less than 5 out of 1000 times. The 0.13 effect size indicates that this improvement is 132% of an average standard deviation in size. Cohen considers this a “very small effect.” The eta-squared (.01) is also small.

These interpretations are also presented in terms of the changes in percentages of students picking each possible choice before and after being taught in Table 4. To better interpret Table 4, a series of bar graphs for each item are included as well. The vertical scales are the same for each pair of graphs: pre and post. The horizontal scale – which is meaningless – may vary as an artifact of the software used.

Conclusions – see page 1 Executive Summary

References

- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- 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 Pigtm goes to the big city: Testing the effectiveness of an economics curriculum for young children. *The Social Studies*, 96(2).