Chicago Public Schools Evaluative Report

August 8, 2012

Eric A. Hagedorn, Ph.D.

Hagedorn Evaluation Services
El Paso, TX

Introduction

The purposes of this study are to: 1) measure the effectiveness of a program called Money Savvy Kids® on the attitudes and knowledge of children in Chicago Public Schools.

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 2011-2012 school year, over 245 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 a 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 first 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 past year, results from the 2011-2012 Chicago study, were used in a psychometric evaluation of the instrument (factor analysis and subscale reliability). On the basis of these results, two psychometrically unsatisfactory items were substantially changed. The "saving money is greedy" item was replaced with "I'm too young to need a long term goal for my money." The "business people help others by providing them with goods and services" was replaced with "I can save money when I spend my money very carefully." The object here was to remove confusing items and add items similar to others in intent, but with different wording. The first item about knowing how to handle money was reworded negatively to ensure that students thought about it. Finally, the simple 3-point Likert scaling (agree, unsure, disagree) was replaced by the standard 5-point Likert scale. A full quantitative psychometric analysis of these results will be completed this fall.

This extensively revised survey was given to students before receiving their Money Savvy Pigs and after they had completed the curriculum. The analyses this academic year were based upon 2,733 matched pre and post-tests. This is one of the largest samples we have analyzed over the years. This report presents the interpretation of the results of those surveys followed with extensive supporting analysis. In other words, our conclusions begin now.

Executive Summary of Results and 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 students indicate statistically significant improvements on all 10 items.

The item for which there was the most change (as indicated by the largest effect size) was item #4. This item suggests that it is best to keep the money you save in your room. After instruction, 22.5 % more students disagreed that you should keep money in your room at home The percentage of students who were unsure increased slightly by 1.4%.

The remainder of the statistically significant changes had small Cohen 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 to 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 in the Chicago area. In addition, changes to procedures for administering and collecting completed tests, seem to be yielding far more and far better data: matchable tests are always better than independent samples. Overall, these results are consistent with those found in previous studies in Washington State (over several years); in Cleveland, Ohio; in North Dakota; and even in the Chicago area for the 2004-2005, and 2011-2012 school years. 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, 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 Wilcoxon Signed Ranks test is the non-parametric equivalent to the paired or dependent samples t-test. It is used because the paired samples t-test assumes a normal distribution in 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 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. One of the most well-known effect size calculations derives from 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 was 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" (Cohen, 1992; 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 $1/10^{th}$ of a standard deviation. This effect size is considered "small," even though the likelihood that such a change occurred by chance is very unlikely.

The eta squared statistic is considered as an effect size as well (it is most commonly used in analyses of variance but can certainly be used with paired samples analyses), but it is more commonly thought of as a measure of the amount of variance explained in the post-test 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.

More pre-tests were received than post-tests. This is a common occurrence in year -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 2,773 individual pre and post-tests from 51 distinct schools and 124 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

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=2733). These are provided in Table 1. When the sample size ("N") is less than 2733 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 from Chicago area students.

	N	Pre	SD	Post	SD	Desired change/achieved
Item 1	2733	2.55	1.53	2.02	1.44	Decrease, yes
Item 2	2712	4.15	1.22	4.41	1.09	Increase, yes
Item 3	2697	2.42	1.55	1.92	1.38	Decrease, yes
Item 4	2695	3.82	1.47	2.95	1.67	Decrease, yes
Item 5	2711	4.23	1.20	4.44	1.09	Increase, yes
Item 6	2694	3.26	1.21	2.88	1.52	Decrease, yes
Item 7	2687	2.59	1.52	1.99	1.41	Decrease, yes
Item 8	2692	4.30	1.18	4.48	1.03	Increase, yes
Item 9	2696	1.85	1.36	1.67	1.21	Decrease, yes
Item 10	2710	4.35	1.14	4.56	0.98	Increase, yes

Table 2. 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 don't know very much about how to handle my money.	-14.56	0.000	0.29	Small
2. I can save money when I spend my money very carefully.	-9.15	0.000	0.18	Small
3. It is important to have the things I want when I want them.	-14.39	0.000	0.29	Small
4. It is best to put the money you save in your room at home.	-20.80	0.000	0.44	Small
5. It is important to save for the things that I want to buy in the future.	-7.70	0.000	0.15	Small
6. When I invest in stocks, I will always make money and never lose money.	-10.54	0.000	0.21	Small
7. I'm too young to need a long term goal for my money.	-15.73	0.000	0.32	Small
8. It is important for families to keep money in real banks.	-6.76	0.000	0.13	Small
9. I want to spend the money I earn right away.	-6.25	0.000	0.12	Small
10. When I donate money it helps other people and it helps me too.	-8.08	0.000	0.16	Small

Table 3. Eta squared

Item	Eta squared	Interpretation
1. I don't know very much about how to handle my money.	0.08	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.	0.08	Medium
4. It is best to put the money you save in your room at home.	0.16	Large
5. It is important to save for the things that I want to buy in the future.	0.02	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.	0.09	Medium
8. It is important for families to keep money in real banks.	0.02	Small
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.	0.02	Small

What Tables 1, 2 and 3 tell us about student responses to individual items.

In general the first two tables show that there were statistically significant improvements in student understanding/attitude 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: 5 indicating total agreement, 4 indicating agreement, 3 indicating uncertainty, 2 indicating disagreement and 1 total disagreement. Following these written descriptions is a table and a series of bar charts showing how student responses changed from pre to post in terms of percentages of students responding to each response category before and after instruction. 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 2.55, leaning towards uncertainty to 2.02, 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 significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The 0.29 effect size indicates that this improvement is 29% of an average standard deviation in size. Cohen considers this a "small effect." The eta-squared is medium.

The average response of the students to Item #2 changed from 4.15, leaning towards agreeing, to 4.41, 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 0.18 effect size indicates that this improvement is roughly 18% an average standard deviation in size. Cohen considers this a "very small effect." The eta-squared is also small.

The average response of the students to Item #3 changed from 2.42, on the disagreeing side of unsure, to 1.92, 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 13 in 1000 times. The 0.29 effect size indicates that this improvement is 29% 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 eta-squared is medium.

The average response of the students to Item #4 changed from 3.82, leaning towards "kind of agreeing" to 2.95, which is on the disagreeing 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 .44 effect size indicates that this improvement is one tenth of an average standard deviation in size. Cohen considers this a "small effect," but note that 0.50 is considered a "medium effect." The eta-squared is large, which suggests persistence of earlier views, despite instruction.

The average response of the students to Item #5 changed from 4.23, leaning towards agreeing, to 4.44, 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 0.15 effect size indicates that this improvement is about one seventh of an average standard deviation in size. This is a small effect size. The eta-squared is also small.

The average response of the students to Item #6 changed from 3.26, on the agreeing side of unsure, to 2.88, 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 0.21 effect size indicates that this decrease in score is about 21% an average standard deviation in size. Cohen considers this a "small effect." The eta-squared is small.

The average response of the students to Item #7 changed from 2.59, on the disagreeing side of uncertain, to 1.99, 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 0.32 effect size indicates that this improvement is almost one third of an average standard deviation in size. Cohen considers this a "small effect." The eta-squared is medium.

The average response of the students to Item #8 changed from 4.30, leaning towards agreement to 4.48, more strongly agreeing. This indicates an improvement in students' perceptions that their families should keep their money in banks. The two-tailed significance implies that this improvement in average score could only have occurred by chance less than 1 in 1000 times. The .13 effect size indicates that this improvement is about one eighth of an average standard deviation in size. Cohen considers this a "small effect." The eta-squared is also small.

The average response of the students to Item #9 changed from 1.85, leaning towards disagreeing, to 1.67, leaning even 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 1 in 1000 times. The .12 effect size indicates that this improvement is roughly 12% of an average standard deviation in size. The eta-squared is also small.
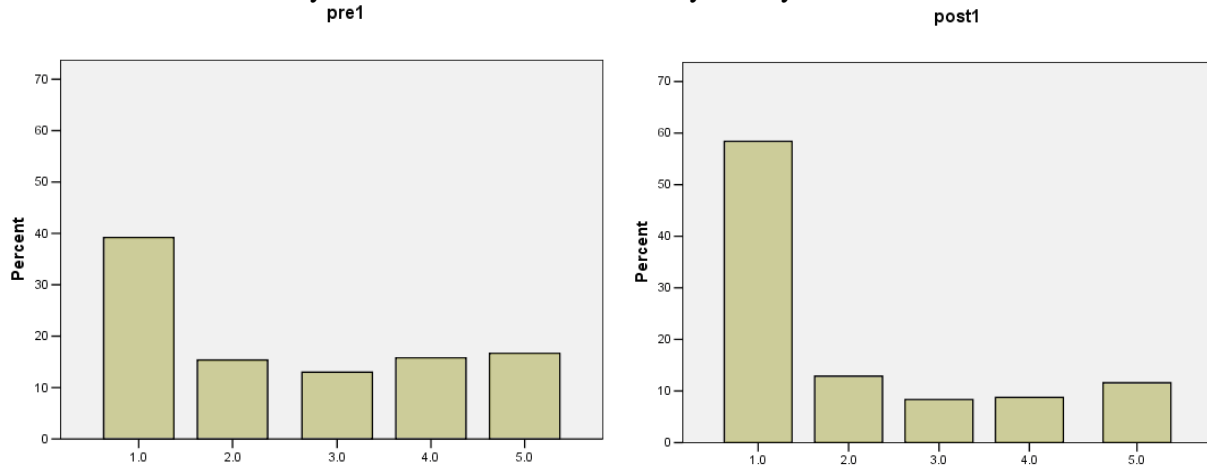
The average response of the students to Item # 10 changed from 4.35, learning towards strong agreeing, to 4.65, which leans towards total agreement. This indicates an improvement in student understanding because more students should agree that donating money helps the recipient and the donator. The two-tailed significance implies that this change in average score could only have occurred by chance 1 out of 1000 times. The 0.16 effect size indicates that this improvement is 16% of an average standard deviation in size. Cohen considers this a "small effect." The eta-squared 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.

Table 4. Percentages of chosen responses to selected items on matched pre and post-measures.

| Item	Response Category`	% students pre-test	% students post-test	Comments
1	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	39.2 15.4 13 15.8 16.6	58.4 12.9 8.3 8.7 11.6	Responses to this negatively worded item indicate that students believed that they did know about how to handle their money after participation. 19.2% more disagreed totally after, and 4.7% fewer were unsure.
2	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	6.7 5.3 11.8 18.5 57.6	5.3 2.4 8.3 14.1 69.9	While a majority (76%) believed they could save money by saving carefully before instruction, even more (84%) believed so after instruction.
3	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	45.8 12.6 12.8 11.9 16.9	61.6 12.1 8.6 7.9 9.8	More than half (58%) disagreeing with immediate gratification to begin with, jumped to almost 74% disagreeing after. Uncertainty about this dropped by 4.2% and those agreeing dropped by 11%.
4	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	13.8 8.8 10.1 16.9 50.4	33.5 11.6 11.5 13.1 30.1	22.5 % more students disagreeing that you should keep money in your room at home. 1.4 % more unsure. 24.1 % fewer agreeing with this after participation.
5	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	6.5 4.6 10.0 16.6 62.3	5.0 3.0 7.3 11.8 72.8	1.5 % fewer students disagreeing after instruction that saving for future is important. 4.7% fewer unsure about this and 5.7 % more agreeing.
6	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	11.6 7.3 46.1 13.4 21.6	29.6 11.1 23.3 13.6 22.4	21.8 % more disagreeing that you always make money on stocks. 22.8% fewer unsure. The percentage agreeing with this seem to persist after instruction
7	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	38.3 12.4 18.9 13.4 17.0	60.3 10.3 10.6 7.9 10.8	20% more strongly disagreeing that they are too young to set up long term goals for their money. Uncertainty dropped by 8.3%
8	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	6.2 4.6 8.5 14.3 66.4	3.8 3.6 7.2 12.0 73.4	3.4 % fewer disagree that it is important for families to keep money in real banks, although a small percentage do so. 1.3 % less unsure and 4.7 % more agreeing. While there is a small change in the correction direction, almost 81% knew the correct response before participation.
9	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	65.6 10.1 6.9 8.0 9.3	70.5 10.6 6.6 6.2 6.1	5.4 % more students disagreeing, 0.3 % fewer unsure, and 5% fewer agreeing with the statement that they want to spend their money right away.
10	1 total disagree 2 kind of disag 3 unsure 4 kind of agree 5 totally agree	5.6 3.4 9.4 12.7 68.7	4.0 2.1 5.3 10.8 77.8	7.2% more agreeing that donating money helps both others and themselves, although even before instruction 80% of the students agreed with this. 4.1% fewer are uncertain about this.

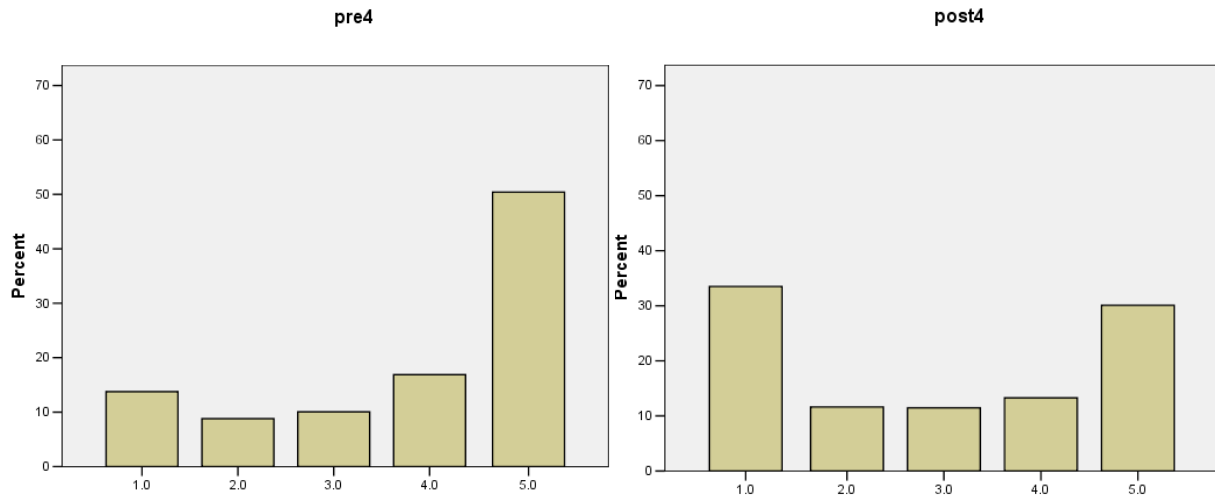Item 1: I don't know very much about how to handle my money.

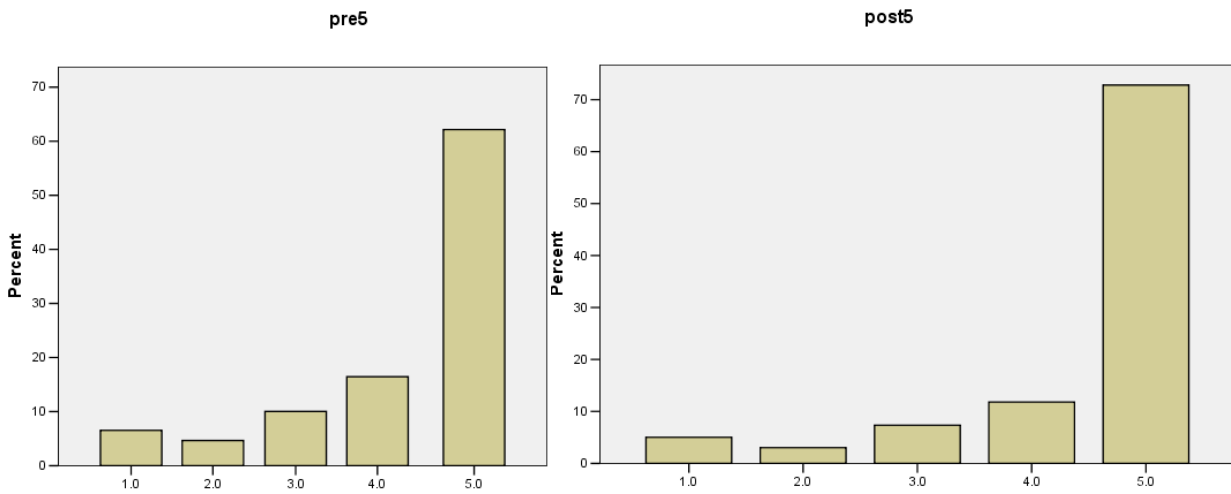Item 2: I can save money when I spend my money very carefully.
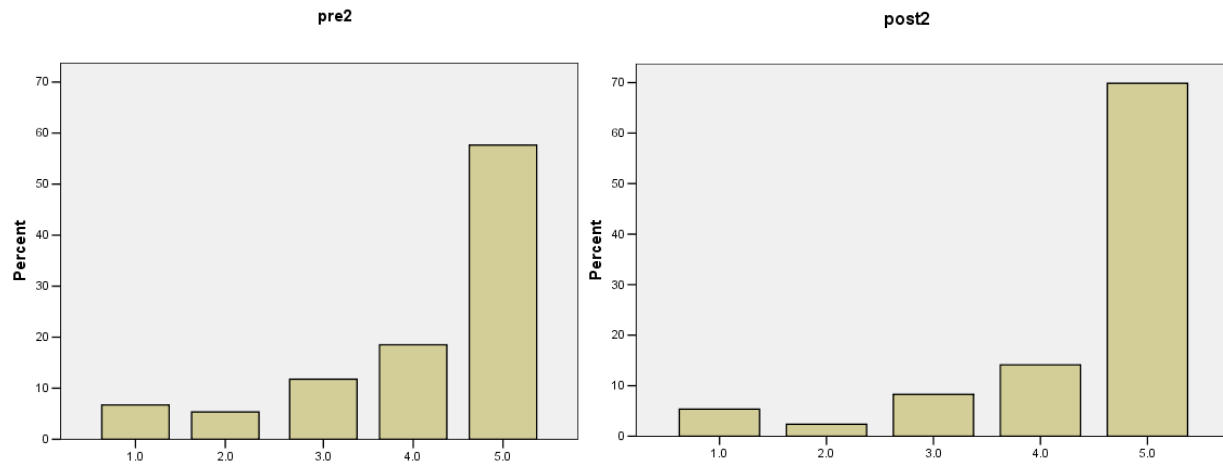
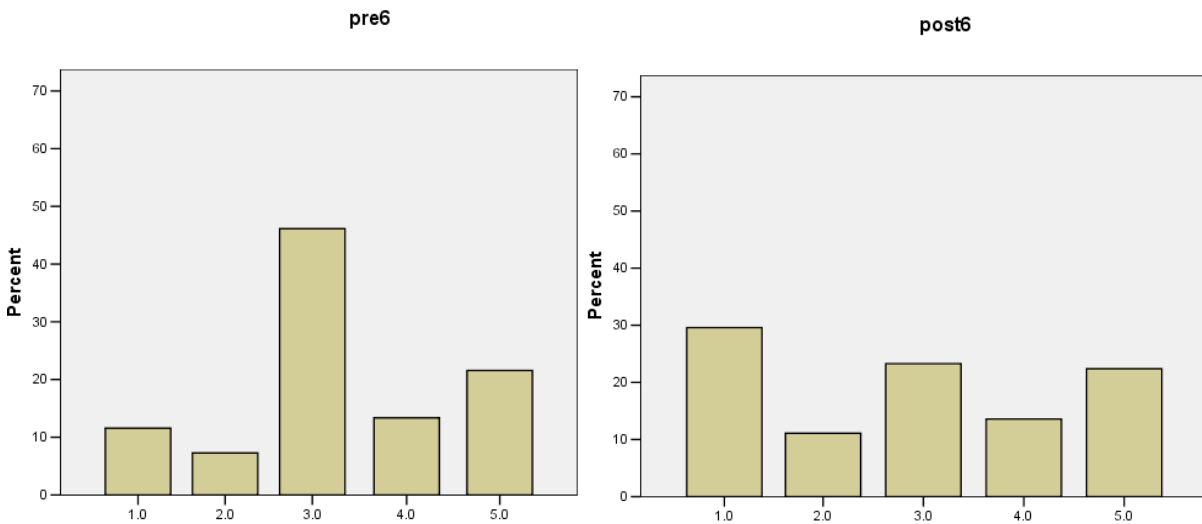Item 3: It is important to have the things I want when I want them.
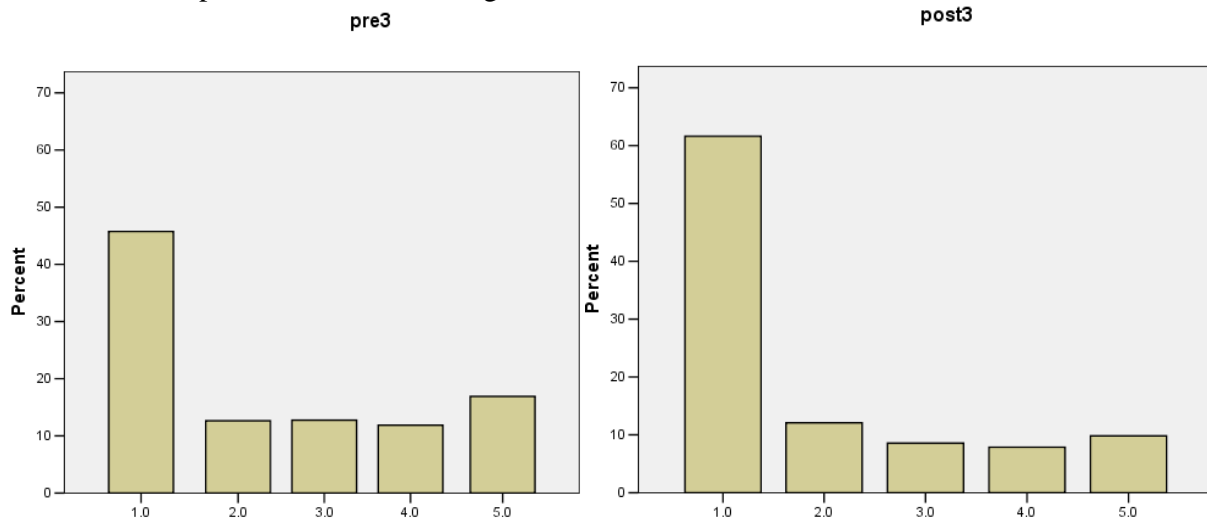
Item 4: It is best to put the money you save in your room at home.
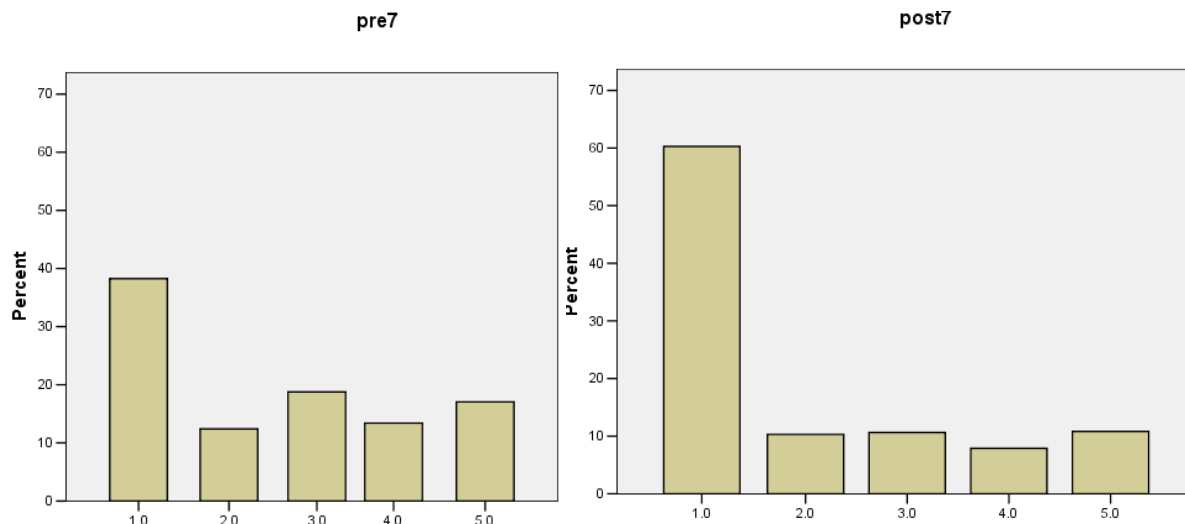

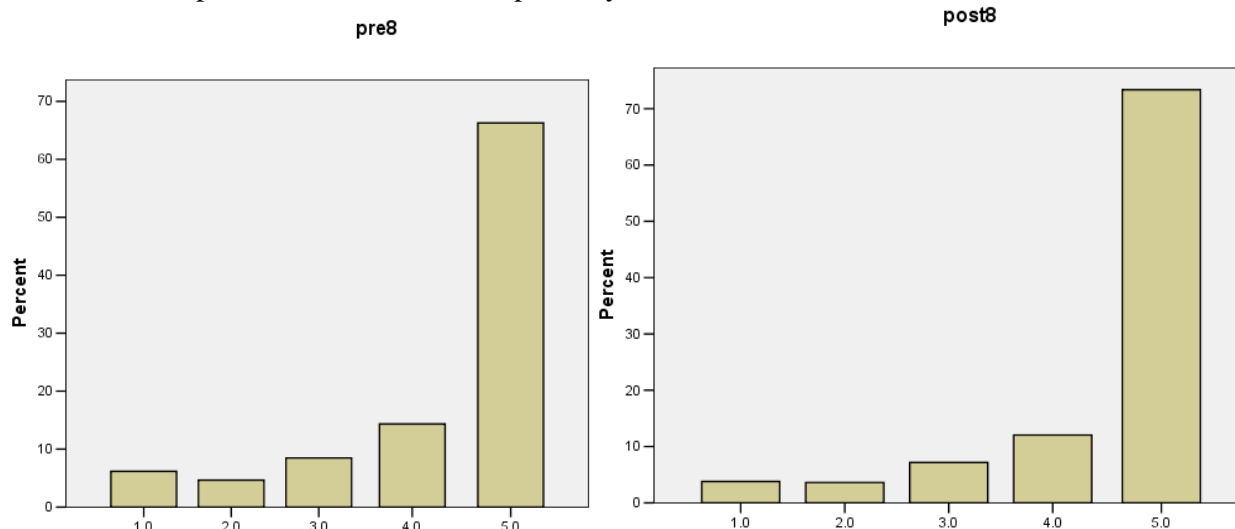Item 5: It is important to save for the things that I want to buy in the future.

Item 6: When I invest in stocks, I will always make money and never lose money.
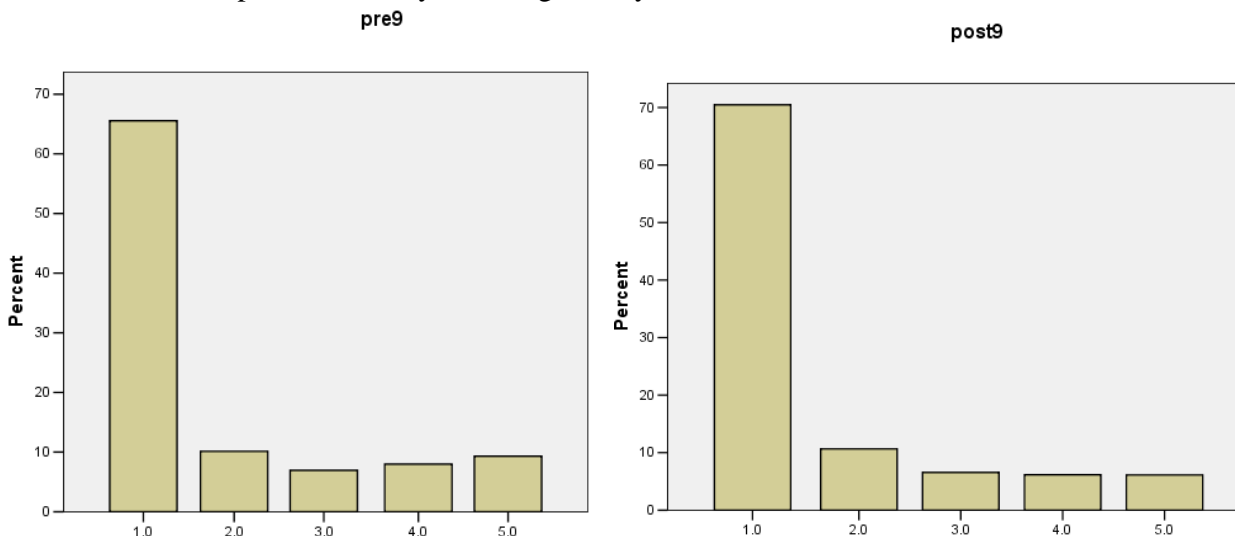
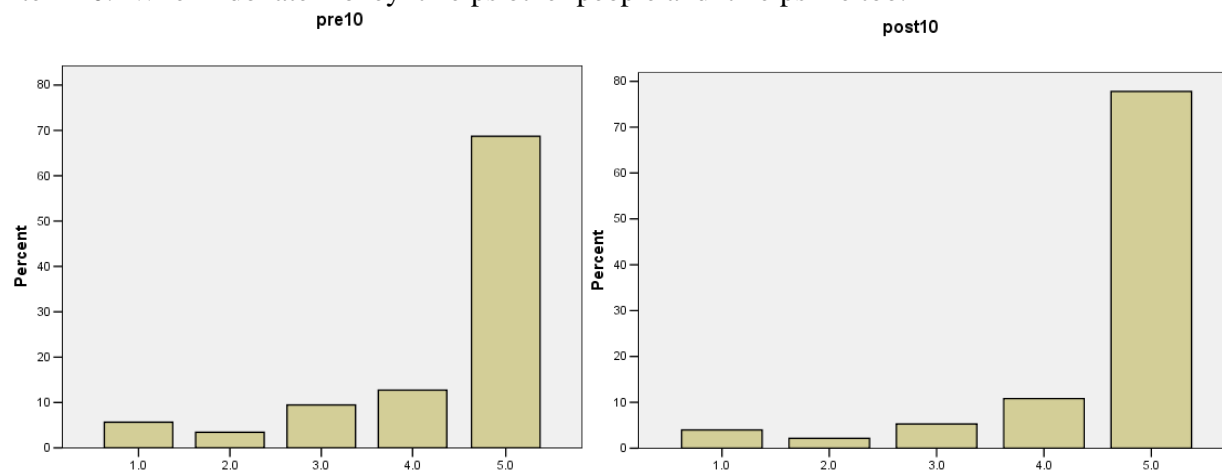Item 7: I'm too young to need a long term goal for my money.

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.

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 Pig[tm] goes to the big city: Testing the effectiveness of an economics curriculum for young children. *The Social Studies, 96*(2).