

Assessing the Reliability and Validity of the Jump\$tart Survey of Financial Literacy

Thomas A. Lucey

The University of Memphis

ABSTRACT: Financial education represents an area of popular interest, owing largely to the Jump\$tart surveys of financial literacy. However, while the surveys represent indicators of financial knowledge among high school seniors, these measures have not been statistically validated. This article describes an assessment of the surveys' reliability (internal consistency), and validity. It reports a moderately high degree of consistency overall, however, discloses low to moderate internal consistencies among subscales. It also finds significant response differences to one quarter of comparable items between surveys. The researcher observes challenges to affirming the surveys' validity and offers statistics suggesting social bias among survey items. He calls for further research into measures of financial literacy.

KEY WORDS: assessment; financial education; Jump\$tart Coalition; social bias; validity.

Reliability and validity represent important concepts in evaluating statistical measures. Unless measures possess these qualities, researchers inaccurately interpret phenomena and policymakers erroneously justify associated decisions.

A reliable measuring tool (or instrument) provides the same result over repeated assessment efforts. For example (presuming it has been scored properly), a wooden ruler reliably measures feet and inches, ensuring repeated use without endangering accuracy. Internal consistency closely resembles reliability, representing an instrument's (or its subscales') ability to prompt similar responses when measuring a phenomenon (or series of phenomena).

An instrument must be reliable to be valid; but if the measure lacks validity, it is not reliable. Validity involves an instrument's ability to measure its targeted concept. Validity lacks an intrinsic basis, because it requires contextual relationships to be established. Validity depends on the instrument's relationship to its setting.

Thomas A. Lucey. The University of Memphis, 419A Ball Hall, Memphis, TN 38152; email: Tlucey@memphis.edu.

The author appreciates the guidance and support of Dr. Duane Giannangelo, Dr. Jean Steiz, Dr. Lou Franceschini, and Dr. Steven Ross.

Research generally recognizes five types of validity: face validity, content validity, construct validity, congruent validity, and predictive validity. Ascertaining face validity and content validity requires expert interpretations of instrument items and their representation of the measured concept. Determining construct validity necessitates interpretation of documents and procedures supporting instrument construction. Interpreting congruent validity involves comparisons with other instruments measuring similar phenomena. Predictive validity represents an instrument's ability to accurately forecast outcomes. A complete validity interpretation addresses all five dimensions.

Financial education represents a popular topic due to publication of the Jump\$tart Coalition and its three surveys (Jump\$tart Coalition, 1997, 2000, 2002). However, the reliability and validity of the Jump\$tart surveys have yet to be determined (Mandell, personal communication, June 11, 2002). This article describes the reliability (consistency) and validity of the Jump\$tart Coalition's 1997 and 2000 surveys. It summarizes the surveys, their development, and their findings before describing the research methodology and data analysis. Finally, the article concludes with a summary of results and recommendations for future research.

In December of 1995, the Jump\$tart Coalition organized to improve K-12 students' personal finance skills (Mandell, 1998). In early 1997, a national sample of 1,532 high school seniors in 12th grade English classes completed the initial Jump\$tart Survey. The findings documented alarmingly low scores, indicating poor levels of financial knowledge. Successive surveys (Jump\$tart Coalition, 2000, 2002) confirmed findings of low financial literacy.

Findings included large differences among scores by seniors of different ethnicities. These disparities prompted Mandell's (2002) observation, "Performance differences were more closely related to race than any other background variable" (p. 15). While such disparities prompt concerns about financial knowledge among underrepresented students, it may result from survey bias.

Methodology

Internal Consistency

I procured the 1997 and 2000 survey datasets from their author (Mandell, 1997 survey data set, personal communication, October 20, 2002; Mandell, 2000 survey data set, personal communication, June, 17, 2002). The 2002 survey dataset was unavailable at the time of study proposal.

The analysis employed the Formula KR20 (Kuder-Richardson 20) to measure internal consistency. Fraenkel and Wallen (1990) described the Kuder-Richardson formulas as the most frequently employed methods for determining “internal consistency” (Fraenkel & Wallen, 1990, p. 135). The procedure interpreted the binary (right and wrong) coded student responses in the two datasets.

Publications (Mandell, 1998, 2000, 2002) presented both overall survey results and detailed subscale findings. Therefore, the study included calculations of internal consistencies for the overall survey and the four subscales.

Between Survey Reliability

The study compared the binary (correct or incorrect) responses for the 1997 and 2000 surveys. Since significant changes were made to five common survey items, and because the 1997 survey contained one item excluded from the 2000 survey analysis, the study interpreted between-survey reliability of only 25 items.

The comparison employed a series of χ^2 calculations. These analyses interpreted the frequencies of correct responses among the surveys' common or similarly worded items. Because the number of respondents to the 1997 survey exceeded twice the number for the 2000 survey, the processes employed a random sample of cases from the 1997 dataset. The number of random cases equaled the number of analyzed cases from the 2000 dataset. I developed the random sample using the software's (SPSS 11.5, 2002) case selection menus.

Validity

The study interpreted the surveys' validity by reviewing prior financial literacy measures, considering literature about the surveys' development, communicating with the Jump\$tart Coalition, and reviewing related research. Analysis also assessed the surveys' face, content, construct, congruent, and predictive validities.

Social Bias

The researcher randomly selected eight (or 24%) high schools in a large southern city public school system to survey opinions of 27 social studies teachers. For each item, the respondents indicated whether they thought students of different backgrounds would similarly interpret content. Respondents considered the following issues in their decisions: whether students of different races or ethnicities would interpret the items similarly; whether students of different family incomes would interpret the items similarly; whether students of different family wealth would interpret the items similarly; whether students of different living circumstances (e.g. living at home with parents or living on own, possibly with children) would interpret the items similarly; and other considerations they thought important.

Instrument. The instrument contained all items from the 1997 survey (including the alternative responses for each item). Response alternatives for each instrument involved a 4 level Likert-style format, ranging from Strongly Agree (4) to Strongly Disagree (1). The study calculated mean agreement with each survey item.

Results

The article describes internal consistency and between-survey reliability results. It then explains the surveys' validity and possible bias.

Internal Consistency

The study calculated both surveys' internal consistencies. Two analyses occurred. The first determined the consistencies of the surveys, including all items. The second adjusted the findings for similar items between the surveys. The process only included cases involving valid responses to all survey items.

Unadjusted consistencies. The statistics in Table 1 indicate both instruments contained moderately high internal consistencies overall; however, the subscales possessed low to moderate consistencies. The researcher found higher consistencies for the 2000 survey overall and for three subscales (income, spending and credit, savings and investment). The internal consistency was higher for the 1997 money management subscale.

Interpretation. The low subscale alphas result from at least two causes. First, internal consistency generally increases with the numbers of items. Greater numbers of items potentially increase subscale consistencies. Second, the subscales, as defined by the surveys' author, contain items categorized too narrowly. For example, budgets (or spending plans) represent both money management and spending and credit issues. Jumpstart's original Personal Financial Management Guidelines (1999b) listed several benchmarks categorized in more than one area. Overlapping financial tenets possibly affected subscale interpretations.

Adjusted consistencies. The 2000 survey contained one less item than the 1997 survey and five of the 30 items differed significantly. Table 2 presents adjusted statistics.

TABLE 1
Inter-item Consistency: 1997 and 2000 Jumpstart Surveys

	1997 (N=1,404)		2000 (N=665)	
	Number of Items	α	Number of Items	α
Income	7	0.38	7	0.58
Money Management	5	0.31	4	0.23
Savings and Investment	8	0.35	8	0.43
Spending and Credit	11	0.59	11	0.59
Overall Survey	31	0.74	30	0.78

TABLE 2
Adjusted Internal Consistency: 1997 and 2000 Jumpstart Surveys

	Items	1997 α	2000 α
Income	5	.32	.50
Money Management	3	.23	.24
Savings and Investment	7	.30	.36
Spending and Credit	10	.56	.54
Overall Survey	25	.69	.74

Interpretation. The adjusted results yielded similar statistics to those determined by the unadjusted analyses. The statistics indicate the surveys held moderately high internal consistencies overall, with the subscales involving much lower consistencies. The 2000 survey boasted mostly higher alphas than the 1997 survey, although the alpha associated with 2000 spending and credit subscale was somewhat lower than that for the 1997 survey.

Between—Survey Reliability

The researcher calculated the probability of equal correct response representation among the two surveys' 25 common items. Table 3 presents the statistical findings.

Interpretation. The results indicate significant response differences to eight of thirty-one (32%) items. The data depict significant differences between survey responses for a high (80%) percentage of the income items. They also reveal significant differences in a moderately

TABLE 3
Between 1997 and 2000 Survey Reliability Analysis: Weighted Frequency Differences among Correct Item Responses

Income	χ^2	<i>p</i>
1. College vs. High School Earnings	1.843	.175
2. Academic vs. Socialite: Higher Income	12.410	.000**
3. Primary Income Sources: Ages 20–35	13.645	.000**
8. Paycheck deductions	5.861	.000**
9. Facts about sales taxes	31.683	.000**
<i>Money Management</i>		
4. Name for Retirement Income from Company	21.962	.000**
7. Greatest problem with inflation	.069	.793
12. Greatest Need for Life Insurance	2.645	.104
<i>Savings and Investments</i>		
11. Interest on Bank Savings	17.288	.000**
15. Safest place for college money	.199	.655
16. Highest growth over 18 years.	13.636	.000**
17. Least use for emergencies.	.725	.395
18. Unprotected by Federal Government	.398	.528
19. Best Protection in Inflation	1.021	.312
20. Sarah saves earlier than Ben, Who has more at the end?	.006	.939
<i>Spending and Credit</i>		
21. Instruments not associated with spending	.103	.749
22. Greatest Finance Charges	2.732	.098
23. Finance Charge on car versus vacation	2.639	.104
24. Not true about ATMs	.016	.900
25. Not reduce cost of college loan	3.557	.059
26. Credit company reduces risk with new borrower	.110	.740
27. Card loss if stolen	14.438	.000**
29. You can check your credit history	.881	.348
30. What can credit counseling do?	.230	.631
31. Debt is beneficial	.022	.882

p* < 0.05; *p* < 0.01.

low (29%) percentage of responses to savings and investment items, present significant differences for a moderately-low (33%) percentage of the money management items, and show significant differences for a low (10%) percentage spending and credit items.

Consider the following factors when interpreting these results. First, the analysis involved a random sample of cases from the 1997 survey. Different response patterns associated with different samples could have affected results. Second, the survey processes collected no data measuring student achievement. Achievement variations represent possible sources of differing responses. While the researcher recognizes challenges comparing achievement results among states,

future surveys should compare survey findings with student achievement.

Validity

Face validity. Published findings (Mandell, 1998) reported members of the Jump\$tart Coalition reviewed the survey items. While this procedure appeared to validate the survey, the reader should recognize financial professionals and educational leaders comprise the Coalition. Hence, the review lacked full consideration by all knowledgeable parties. Parents, teachers, and students were not sufficiently consulted. The researcher acknowledges parents are not always the best teachers (Bryant-Quinn, 1999) and do not necessarily agree when to teach financial concepts (Danes, 1994); however, the researcher also recognizes parents offer valuable insights into their children's financial education needs and represent important informational resources.

Content validity. Jump\$tart's surveys measured four areas of financial understanding: income, money management, savings and investments, and spending and credit. While survey content included items within each of these areas, the number of items was insufficient to cover all content. The researcher bases this observation on his review of the Coalition's original Personal Financial Management Guidelines (1999b) (those existing at the time of the administrations of the 1997 and 2000 surveys). These guidelines contained 49 benchmarks for high school seniors. After accounting for overlapping financial areas, the number of benchmarks was still much larger than the number of survey items.

Construct validity. Soon after its formation, the Jump\$tart Coalition developed curriculum guidelines and benchmarks for financial education. The Coalitions' original executive board developed the benchmarks (Bannister, personal communication, September 9, 2003). Aside from the disclosure of a "nationally selected panel of teachers completing the guidelines" (Jump\$tart, 1999a, p. 2), no publications or publicly available documents describe the overall methodology for developing Jump\$tart's 2001 curriculum standards or their 2001 revision (Duguay, personal communication, July 2, 2003; Duguay, personal communication, August 4, 2003). Supervisory responsibility for their revision is unclear (Bannister, personal communication, September 9, 2003; McCorkle, personal communication, September

10, 2003). The Coalition's revised curriculum guidelines (2001) appear to be the most widely recognized financial education standards and provide a foundation for measuring understandings. Additional documentation must support the surveys' construct.

Congruent validity. No established measure serves as a standard for the surveys. Before the surveys' development, the Joint Council of Economic Education's (1971) (the present National Council of Economic Education) *Test of Understanding in Personal Economics* (1971) standardized measures of high school students' financial understandings. Publication of the instrument ceased in 1980 (Rodriquez, personal communication, September 9, 2002). Since the instrument's cessation, the Consumer Federation of America and American Express (n.d.) assessed U.S. high school students' consumer understandings, and the Americans for Consumer Education and Competition (2001) surveyed teens' financial understandings. No reliability and validity information for these surveys have been publicized. Thus, research lacks readily available standards to interpret congruency.

Predictive validity. The relationship of financial literacy to financial practice represents an interesting research consideration. While pencil and paper tests prompt thoughtful processes, financial decisions often involve spontaneous, impulsive, and/or high-pressured decisions differing from academic settings. J. D. Laney's (1985, Unpublished doctoral dissertation) study found high-school students instructed in cost-benefit analysis demonstrated higher economic reasoning for more familiar problems, regardless of the dilemmas' perceived importance. Laney extended the findings of Kourilsky and Murray (1981) in suggesting different patterns and stages of economic reasoning. Payne's (1995) observations supported theories of contextually related economic reasoning. Research should relate survey responses to financial practice.

Social Bias

If familiarity affects reasoning processes, and students of higher socioeconomic status regularly experience complex financial topics, then students of these contexts respond better to items involving complex financial issues. It is possible higher scores of upper socioeconomic students (Mandell, 1998, 2000, 2002) indicated greater familiarity with survey content, but not necessarily sounder financial practice. The article now provides statistics depicting respondents'

agreement that students of diverse backgrounds would similarly interpret survey items. (Table 4)

Interpretation. The statistics reveal 15 (48%) items involve means of less than 3, indicating presence of social bias. The researcher calculated bias associated with the following percentages of subscale items: income, 29%; money management, 40%; savings and investments,

TABLE 4
Descriptive Statistics: Agreement with Common Item Interpretation
(1 – Low, 4 – High)

	<i>N</i>	μ	σ
<i>Income Items</i>			
1. College vs. High School Earnings	27	3.07	.997
2. Academic vs. Socialite: Higher Income	27	3.07	.958
3. Primary Income Sources: Ages 20–35	26	3.00	.894
5. Taxes if Income Doubles	27	2.67	.832
6. Loss if Remaining in High Business Tax State	27	2.78	.974
8. Paycheck deductions	27	3.37	.839
9. Facts about sales taxes	27	3.19	.921
<i>Money Management Items</i>			
4. Name for Retirement Income from Company	26	3.04	.720
7. Greatest problem with inflation	27	2.85	1.027
12. Greatest Need for Life Insurance	27	3.00	.920
13. Health Insurance Coverage	27	2.89	.892
14. Auto insurance covering injuries to others	27	3.30	.869
<i>Savings and Investment Items</i>			
10. Months to accumulate \$1,000	27	3.19	1.002
11. Interest on Bank Savings	27	3.15	.602
15. Safest place for college money	27	2.78	1.013
16. Highest growth over 18 years.	27	2.74	.859
17. Least use for emergencies.	27	2.59	.931
18. Unprotected by Federal Government	27	2.63	.927
19. Best Protection in Inflation	27	2.52	.935
20. Sarah saves faster than Ben, Who has more at the end?	27	3.00	1.040
<i>Spending and Credit Items</i>			
21. Instruments not associated with spending	27	3.04	.808
22. Greatest Finance Charges	27	2.85	1.027
23. Finance Charge on car versus vacation	27	2.96	.808
24. Not true about ATMs	27	3.00	.920
25. Not reduce cost of college loan	27	2.74	1.023
26. Credit company reduces risk with new borrower	27	2.81	1.002
27. Card loss if stolen	27	3.07	.874
28. Credit history	27	3.00	.832
29. You can check your credit history	27	3.15	.864
30. What can credit counseling do?	27	2.96	.567
31. Debt is beneficial	27	2.74	.984

62%, and spending and credit, 54%. The results suggest either many of the items do not address contexts of all students or different students receive different amounts of exposure to associated topics.

An examination of least agreed upon items by financial area provided the following findings. The two least agreed with income items involved tax rates and business tax effects. The least agreed with money management items concerned inflation and health insurance coverage. The savings and investment items prompting least agreement addressed college savings, growth investments, emergency funds, government protection, and inflation. Least agreement occurred with spending and credit items concerning finance charges, college loan costs, credit risks, credit reporting, and debt benefits. In general, it appears least agreed upon items involved topics middle and upper economic class students may recognize or experience, but lower economic class students may not.

The researcher also compared means against the dividing point associated with the response format. The instrument contained a 4-point Likert-style response structure. By considering 2.5 as the lower limit for agreement, a positive image of the survey emerges. None of the response items prompted mean agreements less than 2.5. These findings indicated no social bias occurred.

Discussion

Analyses indicate the 1997 and 2000 Jump\$tart Surveys possess moderately high internal consistency overall; however, the subscales' lower internal consistencies indicate subscales require reconsideration. Research should both examine whether 30 or 31 survey items provide a complete measure of financial literacy and consider if the surveys adequately differentiate items among the four financial areas.

Improving the surveys requires reconciling social bias in measurements. Linklater (2002) observed struggles of late 17th and early 18th century surveyors, tradesmen and their traditional, subjectively founded standards against the government's objective systems. Likewise, the Jump\$tart surveys should include financial topics applicable to all societal participants by adding items to cover these issues. For example: income items could address career choices and employment pursuits; money management items could include shopping comparisons; savings and investments could address minimum balance requirements; and spending and credit items could ask about credit insurance products.

Conclusions

The Jumpstart surveys possess moderately high inter-correlation consistency overall and some degree of face and content validity. However, limited evidence supports their construct, congruent, and predictive validity. Prudence necessitates further research into the surveys and measures of financial understandings.

A complete conceptualization of survey bias requires larger and more diverse populations. A complete picture of financial education priorities requires examination of financial issues concerning populations of all socioeconomic contexts.

References

- American Express Corporation (n.d.). *Consumer knowledge quiz: High school*. Retrieved July 22, 2003, from <https://www.americanexpress.com/corp/consumerinfo/collegequiz.shtml>.
- Americans for Consumer Education and Competition (2001). *America's Money Skills Report Card*. Retrieved April 23, 2003, from <http://www.acecusa.org/reportcard/>.
- Bryant-Quinn, J. (1999). Money management is a life skill students need to know. *Jumpstart Update*, 3(2), 6.
- Danes, S. M. (1994). Parental perceptions of children's financial socialization. *Financial Counseling and Planning*, 5, 127–145.
- Fraenkel, J. R., & Wallen, N. E. (1990). *How to design and evaluate research in education*. New York: McGraw Hill.
- Joint Council of Economic Education. (1971). *Test of understanding in personal economics: Interpretive manual and discussion guide, and test form* (ERIC Document Reproduction Service No. ED 054030).
- Jumpstart Coalition. (1997, May 22). High school seniors lack financial smarts shows survey. Retrieved March 9, 2002, from <http://www.Jumpstart.org/upload/news.cfm?recordid=37>.
- Jumpstart Coalition (1999a). Personal finance education guidelines and benchmarks. *Jumpstart Update*, 3(1), 2.
- Jumpstart Coalition (1999b). Personal financial management guidelines. Retrieved July 30, 1999, from: <http://www.jumpstartcoalition.org>.
- Jumpstart Coalition (2000, April 9). Financial literacy declining among 12th graders, coalition urges states to include personal finance in curriculum standards. Retrieved March 9, 2002, from <http://www.Jumpstart.org/upload/news.cfm?recordid=60>.
- Jumpstart Coalition (2001). National Standards in Personal Finance. Retrieved August 2, 2003, from the Jumpstart Coalition Web site: <http://www.jumpstartcoalition.org/Standards & Benchmarks.pdf>.
- Jumpstart Coalition (2002, April 23). From bad to worse: Financial literacy drops among 12th graders. Retrieved May 25, 2002, from <http://www.Jumpstart.org/upload/news.cfm?recordid=99>.
- Kourilsky, M. L., & Murray, T. (1981). The use of economic reasoning to increase satisfaction with family decision making. *The Journal of Consumer Research*, 8, 183–188.
- Linklater, A. (2002). *Measuring America: How an untamed wilderness shaped the United States and fulfilled the promise of democracy*. New York: Walker and Company.

- Mandell, L. (1998). Our vulnerable youth: The financial literacy of American 12th graders. Washington: Jump\$tart Coalition for Financial Literacy.
- Mandell, L. (2000). Improving financial literacy: What schools and parents can and cannot do. Washington: Jump\$tart Coalition for Financial Literacy.
- Mandell, L. (2002). *Financial literacy: The growing problem*. Washington: Jump\$tart Coalition for Financial Literacy.
- Payne, R. K. (1995). A framework: Understanding working with students and adults from poverty. Baytown, TX: RFT Publishing.
- SPSS (2002). SPSS 11.5 for Windows [Computer software]. Chicago: SPSS.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.