Statistical Testing of a Measure of Youth's Perceived Improvement in Life Skills

Dr. Lisa A. Guion, Ed.D.
Associate Professor
Department of Family, Youth and Community Sciences
University of Florida
Gainesville, FL
laguion@ifas.ufl.edu

Blanca E. Rivera
Department of Education
College of Education
University of Florida

Statistical Testing of a Measure of Youth's Perceived Improvement in Life Skills

Dr. Lisa A. Guion and Blanca E. Rivera
University of Florida

Abstract: This article presents findings from the statistical test of an instrument designed to measure youth's perceptions of the life skills that were improved as a result of their participation in 4-H Clubs. The questionnaire was administered to 126 4-H club members in Florida. The 19-item self-rating Life Skills Improvement Scale was examined for face and content validity. The results were also submitted for exploratory factor analysis and internal consistency testing. The factor analysis yielded a four-factor solution to the 19-item scale, which accounted for 62.6% of the variance in the scale. The Cronbach's alpha reliability coefficient for the 19 items was 0.88. The article also discusses implications and future use of the instrument, as well as recommendations for further study.

Introduction

An important objective of 4-H Youth Development programs is to help young people develop life skills. Increasingly, 4-H Extension educators are being required to evaluate their programs to determine whether targeted life skills were developed, improved and/or enhanced. Consequently, it is critical that 4-H educators have evaluation tools/instruments that are both, valid and reliable.

"Validity refers to the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration." (Babbie, 2001, p.143) 4-H educators need instruments that are truly measuring what is intended to be measured. On the other hand, "reliability is defined as "an estimate of the stability, dependability, or predictability of a measure." (Thomas, 2005, p.370) According to Santos, "when you have a variable generated from a set of questions that return a stable response, then your variable is said to be reliable." (Santos, 1999, p.2) Reliability focuses on whether the instrument would yield consistent results if/when applied repeatedly with the same audience. Reliability and validity of an instrument increases the faith in and credibility of the results.

Severs, Dormody & Clason (1995) stress the importance of 4-H, FFA and other youth serving organizations having valid, reliable measurement instruments. Their work in testing leadership instruments represented a significant contribution to the field in that it produced a valid and reliable measure of youth leadership skills. However, 4-H focuses on the development and enhancement of many other types of life skills as well. In a search of the literature, the researcher could not identify an instrument that had been scientifically tested that measured a broader aspect of 4-H life skill development.

Therefore, the purpose of this study was to test the validity and reliability of a scale designed to measure youth's perceptions of their improvement in key life skill areas resulting from their involvement in 4-H Clubs.

Methods

Instrument

The Life Skills Improvement Instrument includes 19 indicators of life skills and abilities. Each indicator used a five point Likert Scale with 1 being (strongly disagree), 2 (disagree), 3 (neutral), 4 (agree) and 5 (strongly agree). The items included in the instrument were determined by conducting two strategic steps. First, the researcher surveyed the literature that conceptualized 4-H life skills. For example, life skills from the Targeting Life Skills model (Hendricks, 1998) were identified. Ultimately, life skills from the Texas 4-H evaluation instrument, which is based on the Hendrix model, were adapted for use in the Life Skills Improvement Scale. The Texas model was adapted because "the youth development skills section is a set of statements that are relevant to all project experiences and to youth of all ages and backgrounds." (Howard, Boleman, Alvey, Burkhum, Chilek, Stone, et.al., 2001, p.2).

Second, nine Extension 4-H Agents from different districts in the state of Florida were asked to select the life skills that their 4-H program targets. They were also encouraged to add to or refine the list of life skills. Those items that had the greatest level of consensus were chosen for inclusion in the Life Skills Improvement Scale. Attachment 1 provides a copy of the Life Skills Improvement Scale.

Participants

Participants of the study were 126 youth members of 4-H Clubs in Florida, of which 36% (n=45) were male and 64% (n=79) female. The average age was 13.8 years, ranging from 7 to 18 years old. Participants have been members of 4-H an average of 4.7 years ranging from 2 months to 12 years. More than half (66%, n=83) of the youth in this study described themselves as Caucasian/White, 22% as African-American (n=28), 7% as Hispanic/Latino (n=9), and 5% described themselves as Other (n=6).

Participants and their parents signed informed consent forms and no compensation was provided for participation in the study. The instrument was administered during a regular 4-H club meeting.

Instrument Testing

Validity. Face validity and content validity were used to determine the measure's validity. Face validity refers to an agreed upon meaning of concepts (Babbie, 2001). The measure is determined to be valid "on its face " (Babbie, 2001). Content validity refers to how much a measure covers the meanings included in the construct to be researched/evaluated (Babbie,

2000). Face and content validity were assessed using a panel of experts. The six-member expert judge panel included three 4-H Extension Specialists, two faculty members in Schools of Education, and one Extension Evaluation Specialist. A structured process for the evaluation of face and content validity was given to each expert. Each expert independently rated the relevance of each item to the identified objective using a 4-point rating scale: 1= not relevant, 2 = somewhat relevant, 3 = quite relevant, 4 = extremely relevant. Finally, content validity index was calculated for the measure. The overall content validity index for the instrument was 0.95, which is the proportion of items rated as content valid (a rating of 3 or 4) by the six experts.

Reliability. Cronbach's alpha, a numerical coefficient of reliability, was used to test the reliability of the Life Skills Improvement Scale. Cronbach's alpha was chosen because it "can be computed from data on a single administration of a test and does not require parallel forms, a test-re-test scenario, or multiple judges for which an intra-class correlation coefficient can be used." (Zumbo & Rupp, 2004, p.79).

Alpha coefficients range from 0 to 1. The higher the score, the more reliable the generated scale is. A computed alpha coefficient of 1 denotes perfect internal reliability, whereas 0 indicates no internal reliability (Bryman, 2001). An alpha of 0.80 is typically employed as a rule of thumb as an acceptable level of internal reliability (Bryman, 2001). Therefore, 0.80 was set as the threshold for this study.

Factor Analysis. Exploratory factor analyses were conducted for the Life Skills Improvement Scale using Principal Component extraction and Varimax rotation with an eigenvalue > 1 to explore the factor structure of the instrument. "The purpose of the principal component analysis is to explain as much of the total variation in the data as possible with as few factors as possible" (Kleinbaum, Kupper, & Muller, 1988, p.615). The Kaiser-Meyer-Olkin (KMO) measures of sampling adequacy and Bartlett's test of sphericity were used to determine the suitability of the matrix for factor analytic procedures. The KMO serves as an index of the strength of relations among variables. "This index yields an assessment of whether the variables belong together psychometrically and thus, whether the correlation matrix is appropriate for factor analysis" (Dziuban & Shirkey, 1974, p. 359). KMO correlation magnitudes of .80 and .90 indicate highly acceptable relations in the matrix, whereas results of .60 and below suggest relations of inferior or unacceptable quality not justifying further data analysis. The Bartlett's test of sphericity is a chi-square test of the significance of a correlation matrix. According to Pedhazur and Schmelkin (1991), the null hypothesis is that the matrix is an identity matrix, that is, all the correlations in the matrix are equal to zero. The Bartlett's test of sphericity determines whether the hypothesis that all the correlations in the matrix are not statistically different from zero can be rejected (Pedhazur & Schmelkin, 1991). When this hypothesis cannot be rejected, the matrix should not be factor analyzed (Tinsley & Tinsley, 1987).

Findings/Results

Factor Analysis
Results from the KMO (.81) and Bartlett's test (χ^2=1038.80, df= 171, p<.001) indicated highly acceptable and statistically significant relationships among variables in the matrix. The factor analysis yielded a four-factor solution to the 19-item scale, which accounted for 62.6% of the variance in the scale. Eigenvalues were 6.44 for leadership, 2.20 for basic life skills, 1.96 for 4-H Animal Projects, and 1.30 for workforce preparation. All individual items had loadings above .50 except item 17, "leading a healthy lifestyle" which had a loading of .43 in factor 1,

.46 in factor 2, and .43 in factor 3. One item from the basic life skills factor (#11 "write more clearly) also loaded in the leadership factor. And one item from the workforce preparedness factor (#10 speak publicly) loaded in the leadership factor. These two items had loading below .50. The items and their loadings are presented in Table 1.

Table 1

Summary of factor loadings for orthogonal four-factor solution for the
Life Skills Improvement Scale

Activity	Factor Loadings			
	Leadership	Basic Life Skills	4-H Animal Projects	Workforce Preparedness
1. keep accurate records	.65			
2. plan/organize	.80			
3. set goals	.75			
4. solve problems	.70			
5. make decisions	.73			
6. serve my community or volunteer	.56			
7. lead a group				.67
8. get ready for a job				.56
9. plan my career				.65
10. speak publicly	.45			.52
11. write more clearly	.40	.55		
12. solve conflicts	.56			
13. sew		.82		
14. cook		.82		
15. groom an animal/pet			.86	
16. feed and care for animal/pet			.91	
17. lead a healthier lifestyle	.43	.46	.43	
18. use a computer or other technology		.63		
19. learn photography or other media		.69		

Reliability Analyses
The Cronbach alpha reliability coefficient for the 19-item Life Skills Improvement Scale was 0.88. There are four subscales. The Leadership Subscale is comprised of questions 1, 2, 3, 4, 5, 6, and 12. The Workforce Preparation Subscale consists of questions 7, 8, 9, and 10. The Basic Life Skills Subscale is comprised of questions 11, 13, 14, 17, 18, and 19. The fourth and final subscale is 4-H Animal Project Skills, which consists of questions 15 and 16.

Table 2 shows the alpha for each sub-scale. Three of the four sub-scales were found to be highly reliable based on the predetermined criteria of alpha greater than or equal to 0.80. These include: 1) Leadership Skills (.86), 2) Basic Life Skills (.81), and 3) 4-H Animal Project Skills (.90). Therefore, those three subscales can be used independently to measure leadership skills, basic life skills or 4-H animal project skills respectively. To a lesser extent, the Workforce Preparation Subscale was moderately reliable (.70).

Table 2
Scale structure and Cronbach alpha reliability coefficient for the
sub-scales of the Life Skills Improvement Scale

Factor	Items from Table 1	Alpha
Leadership skills	1, 2, 3, 4, 5, 6, 12	.86
Workforce preparation	7, 8, 9, 10	.70
Basic life skills	11, 13, 14, 17, 18, 19	.81
4-H animal project skills	15, 16	.90

Implications and Recommendations

The results of this analysis indicate that the Life Skills Improvement Scale is a valid and reliable measure of youth's perceptions of their improvement in key life skill areas resulting from their involvement in 4-H. This scale can be used, with confidence, in both formative and summative evaluation. Formatively, Extension 4-H educators can use this tool to earmark life skills that are not perceived by the youth in their program to be improved. Armed with this information the educators can make future program adjustments to address the issue. In relation to summative evaluation, the instrument provides one way that Extension 4-H educators can demonstrate the effectiveness of their 4-H Club Program in improving key life skills among 4-Hers.

However, in the interest of scholarship and refining knowledge in the 4-H field, the instrument should continue to be tested. Further psychometric testing could focus on the criterion validity and/or construct validity of the instrument. The instrument could be tested with youth who have other types of 4-H involvement such as after-school, camping, school enrichment, etc. The instrument could be tested with 4-H Programs in other states. Also, while the sample size was sufficient for statistical analysis, further studies could be conducted with larger sample sizes that have even greater age, gender and/or ethnic diversity. Comparatively, the instrument can be used with 4-H youth and youth in other youth-serving organizations to determine differences in perceptions of life skill improvement resulting from participation in their respective youth organization.

Conclusion

An essential part of 4-H Youth Development program planning is the coordination of life skills to be taught with the indicators to be used in the evaluation process (Loeser, Bailey, Benson, & Deen, 2004). Once indicators of program outcomes are selected, then extension educators must identify or develop evaluation tools (surveys, scales, tests, etc.) to measure those indicators. These tools must be tested for validity and reliability, at a minimum, if we are to place faith in program evaluation results. Also, continued research to refine and test the evaluation tools must also occur if we are to truly advance scholarship in our 4-H Youth Development program evaluation work.

EVALUATION: Florida 4-H Club Member Survey

We want to know your opinion about your 4-H club experience this past year. Your answers to following questions are very important in helping us learn not only what is working well in 4-H, but also what can be approved upon. **Please answer all questions based on this past year only (200X Club year). Circle the number that corresponds with your level of agreement with each. Circle only one response for each question.**

As a result my 4-H club involvement this past year, I "improved" my ability to:

		Strongly Disagree	Disagree	*Neutral*	Agree	*Strongly Agree*
1	…keep accurate records	1	2	3	4	5
2	…plan/organize	1	2	3	4	5
3	…set goals	1	2	3	4	5
4	…solve problems	1	2	3	4	5
5	…make decisions	1	2	3	4	5
6	…serve my community or volunteer	1	2	3	4	5
7	…lead a group	1	2	3	4	5
8	…get ready for a job	1	2	3	4	5
9	…plan my career	1	2	3	4	5
10	…speak publicly	1	2	3	4	5
11	…write more clearly	1	2	3	4	5
12	…resolve conflicts	1	2	3	4	5
13	…sew	1	2	3	4	5
14	…cook	1	2	3	4	5
15	…groom an animal/pet	1	2	3	4	5
16	…feed & care for animal/pet	1	2	3	4	5
17	…lead a healthier lifestyle	1	2	3	4	5
18	…use a computer or other technology	1	2	3	4	5
19	…learn photography or other media	1	2	3	4	5

References

Babbie, E. (2001). *The practice of social research* (9ᵗʰ ed.). Belmont, CA: Wadsworth/Thomas Learning.

Bryman, A. (2001). *Social science research*. New York: Oxford University Press.

Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, 81(6), 358-361.

Hendricks, P. (1998). *Developing youth curriculum using the targeting life skills model: Incorporating developmentally appropriate learning opportunities to assess impact of life skill development* (Tech. Rep. No. 4H-137A). Ames: Iowa State University Extension.

Howard, J.W., Boleman, C.T., Alvey, A., Burkhum, A.B. Chilek, K.D., Stone, C.C. et. al. (2001). Developing a program evaluation instrument for Texas 4-H: A work in progress. *Journal of Extension*, 39(4), Retrieved on January 7, 2005, from http://www.joe.org/joe/2001august/iw4.html.

Kleinbaum, D. G., Kupper, L. L., & Muller, K. E. (1988). *Applied regression analysis and other multivariable methods* (2ⁿᵈ ed.). Boston, MA: PWS-Kent Publishing.

Loeser, D.M., Bailey, S.J., Benson, R.L. & Deen, M.Y. (2004). Measuring impacts with young audiences: Adapting a life-skills instrument for use with third to fifth grade youth. *Journal of Extension*, 42(4), Retrieved June 24, 2005, from http://www.joe.org/joe/2004august/rb1.shtml

Perdhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Santos, J.R. (1999). Conbach's Alpha: A tool for assess the reliability of scales. *Journal of Extension*, 37(2)/. Retrieved June 23, 2005, from http://www.joe.org/joe/1999april/tt3.html.

Seevers, B.S., Dormody, T.J. & Clason, D.L. (1995). Developing a scale to research and evaluate youth leadership life skills development. *Journal of Agricultural Education*, 36(2)

Thomas, C.L.. (2005). Reliability. In S. Mathison (Ed.), *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage Publications.

Tinsley, H. E. A., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, 34(4), 414-424.

Zumbo, B.D. & Rupp, A.A. (2004). Responsible modeling of measurement data for appropriate inferences: Important advances in reliability and validity theory. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences*. Thousand Oaks, CA: Sage Publications.