
Measuring Science Inquiry Skills in Youth Development Programs: The Science Process Skills Inventory

Mary E. Arnold

School of Social and Behavioral Health Sciences
Oregon State University
Corvallis, OR
mary.arnold@oregonstate.edu

Virginia D. Bourdeau

School of Social and Behavioral Health Sciences
Oregon State University
Corvallis, OR

Brooke D. Nott

School of Social and Behavioral Health Sciences
Oregon State University
Corvallis, OR



Measuring Science Inquiry Skills in Youth Development Programs: The Science Process Skills Inventory

Mary E. Arnold, Virginia D. Bourdeau and Brooke D. Nott
Oregon State University

Abstract: In recent years there has been an increased emphasis on science learning in 4-H and other youth development programs. In an effort to increase science capacity in youth, it is easy to focus only on developing the concrete skills and knowledge that a trained scientist must possess. However, when science learning is presented in a youth-development setting, the context of the program also matters. This paper reports the development and testing of the Science Process Skills Inventory (SPSI) and its usefulness for measuring science inquiry skill development in youth development science programs. The results of the psychometric testing of the SPSI indicated the instrument is reliable and measures a cohesive construct called science process skills, as reflected in the 11 items that make up this group of skills. The 11 items themselves are based on the cycle of science inquiry, and represent the important steps of the complete inquiry process.

Outcomes of Science-Based Youth Development Programs: What to Measure?

In recent years there has been an increased emphasis on science learning in 4-H youth development programs. Science has been identified as one of the three "Mission Mandate" areas for the 4-H program nationally. This emphasis was highlighted by the call from National 4-H to have one million youth who have never been in 4-H before enroll in 4-H Science programs (Mielke, LaFleur, & Sanzone, 2010). While youth involved in 4-H projects have been engaged in science-related endeavors for years, the formal call to increase science programming has changed the face of 4-H programs across the country. Since 2006, 4-H has invested considerable resources in the advancement of science learning, and a recent report by external evaluators of the 4-H Science initiative indicated there is "encouraging growth and variety" of science programs across the 4-H program (Riley & Butler, 2012). Typically, outcomes for youth participants in community-based science programs fall into one of five categories: (1)

Awareness, knowledge, or understanding; (2) engagement or interest; (3) attitude; (4) behavior; and (5) skills (Dierking, 2008). Knowledge refers to what a youth understands and comprehends in relation to science, engagement considers the extent to which youth are excited and involved in science learning, while attitude considers the quality of one's long-term perspectives toward science. Behavior refers to actions that youth take as a result of participating in science programs, and skills reflect one's ability to conduct procedures related to science and science inquiry (Dierking, 2008). While these five categories provide a lucid way to conceptualize the potential outcomes of 4-H science programs, there is general recognition that the categories are not exclusive; programs most likely focus intentionality on more than one category at a time, and the development of outcomes across categories is a likely outcome for most youth science programs. Consider for example, a program that emphasizes increasing science skills. Done well, this program will also increase youth engagement, attitudes, and potential behavior. The key to this cross-category benefit lies in the phrase *done well*. What contributes to a youth science program that is done well?

Learning Science in a Youth Development Program Context

In an effort to increase science capacity in youth, it is easy to focus on the content that youth need to learn. That is, to focus on the concrete skills and knowledge that a trained scientist must possess. However, when science learning is presented in a youth-development setting, the context of the program also matters (Campbell, 2008; Horton, Gogolski, & Warkenstein, 2008). Positive youth development (PYD) programs such as 4-H embrace eight well-defined program setting elements that serve to distinguish PYD programs from other programs where the focus is not primarily on youth development (Eccles & Gootman, 2002). These elements are:

- 1) a positive relationship with a caring adult;
- 2) a physically and emotionally safe environment;
- 3) opportunities for mastery;
- 4) opportunities to value and practice service to others;
- 5) opportunities for self-determination;
- 6) an inclusive environment;
- 7) opportunities to see oneself as an active participant in the future; and
- 8) engagement in learning.

Subsequently, it is important to consider the context of the program being conducted when measuring science skills in youth and community programs.

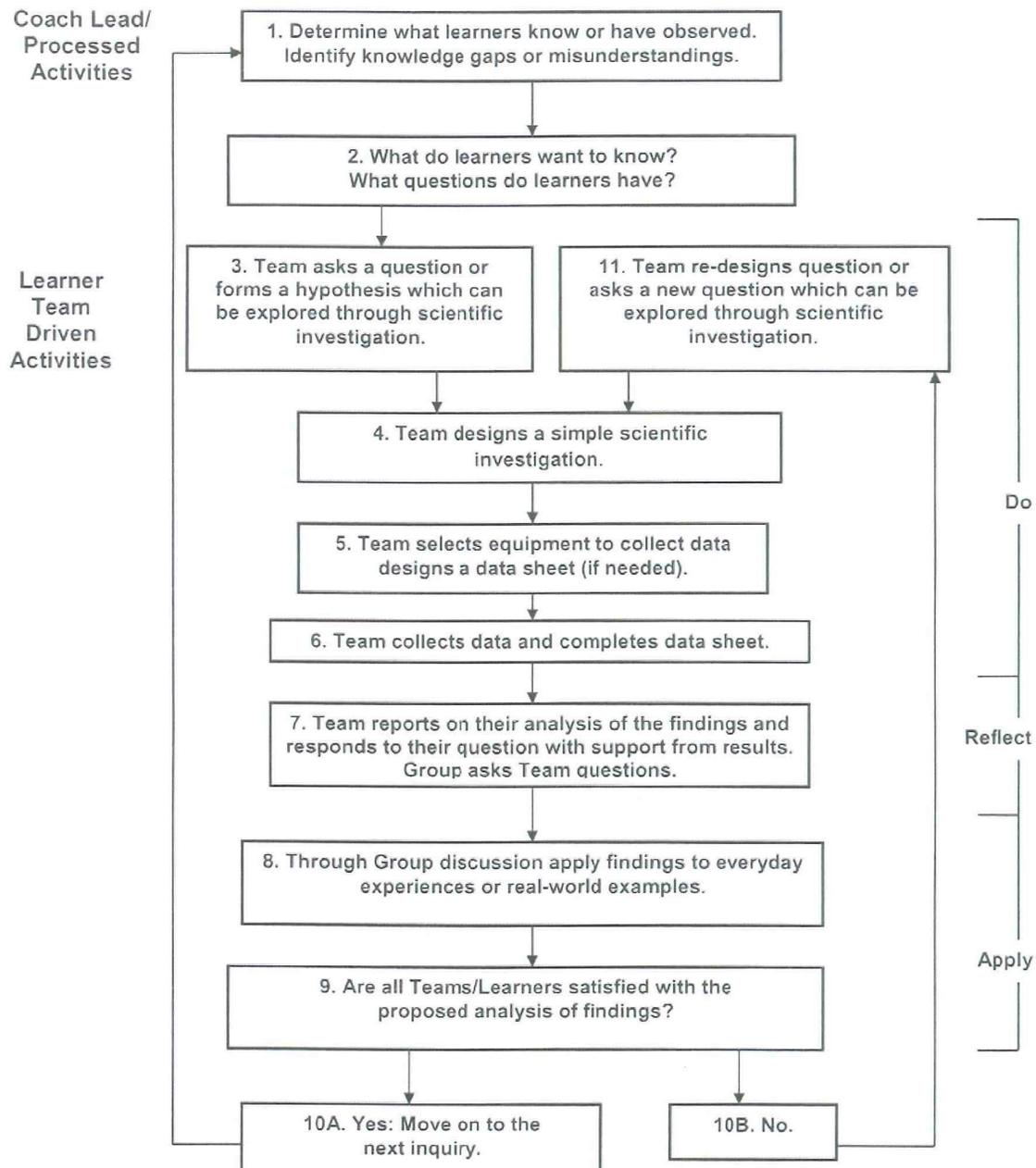
Available resources, skills of the facilitator, the atmosphere of the program, for example, are all important program contexts that influence ultimate program outcomes (Campbell, 2008).

Learning by Doing: A Natural Partnership with Science Inquiry

One of the long-standing program contexts for 4-H is that youth learn "by doing." This is evidenced by the program's embrace of the experiential learning model as a foundational principle for constructing youth learning experiences. This model, often referred to as "Do-Reflect-Apply," has guided the pedagogical approach of 4-H educators and volunteers for many decades. The experiential learning model is based in part on the work of Kolb (1984) who argued that learning is a process, and that one's ideas and thoughts are not fixed, but rather are "formed and re-formed" based on experience. Furthermore, Kolb claimed that learning cannot be defined by an "outcome" only. Kolb went on to highlight Bruner's (1966) claim that the goal of education should be in the *development of skills* that are useful in gaining knowledge and understanding.

Learning then, is a process accomplished through multiple experiences. Likewise, science inquiry is a process of discovery, many times through multiple experiences, for even when an answer is reached, the answer itself leads to a new question to be asked. Science inquiry facilitates a learning process of establishing ideas, testing their merit, revising as needed, communicating results, and developing new ideas. As such, the science inquiry process and experiential learning are quite similar, as demonstrated by Bourdeau (2003) in the *4-H Inquiry in Action* model. This model overlays the experiential learning process and science inquiry and provides a clear picture of the natural fit of science and 4-H youth development (see Figure 1).

Figure 1
4-H Inquiry in Action



Evaluating 4-H Science Programs: What to Measure?

As mentioned earlier, there are five domains of science outcomes that are typically measured, of which skills related to science and science inquiry are one. While the other domains are all important in their own right and for their own goals, we argue that the process of doing science inquiry is a critical outcome for science learning conducted in the context of 4-H and other positive youth development settings. While building skills in science inquiry, we are building skills of learning through experience, and creating an atmosphere of learning that is consistent with the principles of positive youth development. To this end, we have developed and tested the Science Process Skills Inventory (SPSI) which has been requested for use in programs around the world to measure the development of science process skills. This paper presents the results of the psychometric testing of the SPSI with data collected between 2007 and 2011 from youth participants in a residential summer science camp.

The Science Process Skills Inventory

The Science Process Skills Inventory is an 11-item scale that mirrors the steps of the science inquiry process. Youth are prompted to respond to each statement using a 4-point Likert scale indicating how often they practice each of the items when doing science: Never (1), sometimes (2), usually (3), and always (4). Recommended scoring of the SPSI is the *calculation of a composite* science process skills score. This is calculated by summing the individual ratings for each item. The score range for the composite score is 11-44.

SPSI Testing Participants

The SPSI was used to collect data from 252 youth in sixth (80), seventh (86) and eighth (86) grades. Fifty percent were male. The ethnicity/race distribution of the participants was: Caucasian (35%); Hispanic (27%); Asian (11%); African- American (7%); Native American (7%); Pacific Islander (2%); Sub-Continent Indian (1%), mixed (8%), and other (1%). One youth did not report ethnicity. The youth participated in one of five science-focused residential camps held in the summers of 2007 through 2011. Fifty-nine youth participated in 2007, 48 youth participated in 2008, 47 youth participated in 2009, 50 youth participated in 2010, and 48 youth participated in 2011. These youth completed the SPSI pre and post-camp.

Data Analysis Strategy

Factor analysis using principal component analysis (PCA) was used to assess the latent structure of the SPSI pre and post-test scales. In PCA an extraction of the factors occurs and this method was used to determine if the set of items were measuring a single construct made of discreet science process skills. Eigenvalues (sum of the squared factor loadings) greater than one were used as the extraction method, with orthogonal (varimax) rotation, and scree plot tests to determine the factor solutions (i.e. the number of factors to be retained). Items loading on one factor above .40 are considered efficient factor loadings; thus, we used the .40 threshold (Kline, 2005). Identical PCAs were performed on both the pre and post-test scales and the results were compared.

The SPSI was also assessed in terms of internal-consistency reliability. Cronbach's alpha measures the consistency of responses on the items and the correlations among the scale's items should remain consistent, showing the SPSI items used are reliably measuring the science skills concept. To test for differences between possible groups within the sample, ANOVA analyses were also conducted.

Results

The factor analysis of the pre-test items revealed two eigenvalues above one (1.16 and 4.27); however, we also considered the scree plot of eigenvalues that showed a significant drop-off after the first component. After orthogonal (varimax) rotation eleven items loaded on two components (item 10 did not load on either factor above .40 but rather loaded on both factors at .38 and .34). Items 5,6, & 7 had factor loadings above .68 on the second factor. These three items asked students questions specifically about their experience with data and students on average had higher scores on these three items compared to the other eight items in the scale. However, the analysis on the post-test items suggested the retention of one factor with one eigenvalue above 6 and the rest below .93. The principal component analysis after rotation on the post-test items also yielded one factor; all the items loaded on one factor above .67. The scree-plot also confirmed a one-factor solution.

These results have interesting implications for the measurement of science processing skills. The post-test data were collected at the end of a two-week residential camp that focused heavily on developing science process skills in the context of a positive youth development program. As such, each step of the inquiry process was taught, utilized, and emphasized during the two-week camps. By the end of the camps, the SPSI appears to be measuring a more unified construct of science processing skills better than it did at the beginning of camp. Which is to say, that the better the program teaches the individual skills as part of a complete cycle of science inquiry, the better the SPSI will serve as a measurement of that construct. The correlations among the scales' items are presented in Table 1 and 2. The results of the factor analyses are presented in Table 3 and 4.

Table 1*Correlations between Pre-test SPSI Items*

Variables	1	2	3	4	5	6	7	8	9	10	11
1. Pre Use scientific knowledge to form a question	-										
2. Pre Ask question to be answered by collecting data	.37***	-									
3. Pre Design a scientific procedure to answer question	.48***	.42***	-								
4. Pre Communicate a scientific procedure	.34***	.35***	.47***	-							
5. Pre Record data accurately	.18***	.25***	.26***	.31***	-						
6. Pre Use data to create graph for presentation	.17*	.28***	.25***	.23***	.40***	-					
7. Pre Create display of data	.20**	.34***	.35***	.30***	.40***	.43***	-				
8. Pre Analyze results	.42***	.30***	.45***	.32***	.43***	.17*	.38***	-			
9. Pre Use science terms to share results	.39***	.34***	.36***	.41***	.40***	.18***	.28***	.53***	-		
10. Pre Use models to explain results	.26***	.26***	.28***	.27***	.24***	.27***	.21**	.29***	.31***	-	
11. Pre Use results to answer questions	.37***	.35***	.38***	.22***	.31***	.13	.26***	.41***	.35***	.22**	-

*p<.05; **p<.01; ***p<.001.

Table 2*Correlations between Post-test SPSI Items*

Variables	1	2	3	4	5	6	7	8	9	10	11
1. Post Use scientific knowledge to form a question	-										
2. Post Ask question to be answered by collecting data	.57***	-									
3. Post Design a scientific procedure to answer question	.65***	.67***	-								
4. Post Communicate a scientific procedure	.64***	.66***	.67***	-							
5. Post Record data accurately	.38***	.39***	.42***	.45***	-						
6. Post Use data to create graph for presentation	.44***	.55***	.52***	.54***	.56***	-					
7. Post Create display of data	.47***	.60***	.49***	.51***	.53***	.66***	-				
8. Post Analyze results	.59***	.65***	.63***	.70***	.51***	.49***	.54***	-			
9. Post Use science terms to share results	.58***	.62***	.55***	.68***	.44***	.41***	.47***	.65***	-		
10. Post Use models to explain results	.48***	.50***	.53***	.59***	.43***	.48***	.47***	.53***	.53***	-	
11. Post Use results to answer questions	.57***	.57***	.57***	.55***	.50***	.49***	.50***	.57***	.58***	.56***	-

*p<.05; **p<.01; ***p<.001.

Table 3

Summary of Exploratory Factor Analysis Results for Pre-test SPSI Following Orthogonal Rotation (n=204)

Item	Factor Loadings	
	SPSI	DATA
Pre Use scientific knowledge to form a question	.76	.28
Pre Ask question to be answered by collecting data	.54	.17
Pre Design a scientific procedure to answer question	.71	.20
Pre Communicate a scientific procedure	.55	.15
Pre Record data accurately	.28	.67
Pre Use data to create graph for presentation	.02	.85
Pre Create display of data	.23	.74
Pre Analyze results	.69	.25
Pre Use science terms to share results	.67	.24
Pre Use models to explain results	.40	.32
Pre Use results to answer questions	.63	.13
Eigenvalues	4.25	1.22
% of variance	38.68	11.11

Note: Factor loadings over .40 appear in bold.

Table 4

Summary of Exploratory Factor Analysis Results for Post-test SPSI Following Orthogonal Rotation (n=204)

Factor Loadings	
SPSI Item	
Post Use scientific knowledge to form a question	.76
Post Ask question to be answered by collecting data	.82
Post Design a scientific procedure to answer question	.79
Post Communicate a scientific procedure	.83
Post Record data accurately	.67
Post Use data to create graph for presentation	.71
Post Create display of data	.73
Post Analyze results	.82
Post Use science terms to share results	.77
Post Use models to explain results	.72
Post Use results to answer questions	.77
Eigenvalues	6.49
% of variance	58.98

Note: Factor loadings over .40 appear in bold.

Tests for internal reliability revealed a Cronbach's alpha coefficient of .84 for the pre-test scale and .93 for the post-test scale. Alpha coefficients by camp year were also investigated (see Table 5 for the alpha coefficients by camp year).

Table 5

Summary of Alpha Coefficients for Pre and Post-test SPSI by Camp Year (n=252)

Alpha Coefficient		
Camp Year	Pre-test	Post-test
2007	.83	.94
2008	.85	.88
2009	.81	.92
2010	.85	.95
2011	.85	.91
All Years	.84	.93

ANOVA analyses were conducted to determine possible differences in the SPSI by gender, ethnicity, and grade level. These analyses were performed for the total scores on both the pre and post-tests. All tests for significance were deemed insignificant except one. A significant difference was found for grade levels on the post-test at the $p < .05$ level [$F(2, 241) = 3.60, p = .03$]. Post hoc comparisons using ANOVA contrasts indicated that the averages of grades 6 and 7 were significantly different compared to grade 8 scores. A significant difference was also found between grade 6 versus grade 8 (grade 6, $M = 35.74, SD = 5.93$; grade 8, $M = 38.12, SD = 5.37$). The pre and post-test mean scores by are presented in Table 6.

Table 6
Means of SPSI Scores by Grade Level

Grade Level			
	Grade Six	Grade Seven	Grade Eight
SPSI Score	n = 64	n = 70	n = 70
Pre-test	34.12	33.25	34.20
Post-test**	35.55	37.10	38.36

** post-hoc contrasts revealed significant differences between grades ($p < .05$)

Discussion and Conclusion

Overall, the results of the psychometric testing of the SPSI indicated the instrument is reliable and measures a cohesive construct called science process skills, as reflected in the 11 items that make up this group of skills. The 11 items themselves are based on the cycle of science inquiry, and represent the important steps of the complete inquiry process.

In addition to providing support for the overall soundness of the SPSI, the psychometric testing revealed a couple other qualities of the SPSI that have important implications for using this measure . First, was the finding that the post-test cohesiveness of the SPSI was stronger than the pre-test. As noted, the post-test was given after a two-week intensive residential science camp that focused heavily on inquiry in a positive youth development setting. While we recognize that not all programs will have this level of intensity and dosage, the fact that the SPSI factored more strongly at the end of such an experience supports its use as an effective way to measure the development of science inquiry skills. In short, the more emphasis a program places on developing science inquiry skills, the better the SPSI will measure the presence of those skills.

The second important finding was the differences in scores between youth in the 6th, 7th, and 8th grades. One would expect that 8th grade youth will possess more science inquiry skills than younger youth, if for no other reason that the science curriculum of an 8th grader is usually more advanced than that for younger students. These grade-related differences were found at the pre-test time. However, all three grade groups reported stronger use of science processing skills at the end of camp, again with those in 8th grade having the strongest scores. Such findings support the ability of the SPSI to measure increases in skill level regardless of the student's pre-program skill level.

Overall, the SPSP appears to perform well for measuring the development of science process skills in youth who participate in 4-H and other youth development programs that emphasize science inquiry.

References

Bourdeau, V.D. (2003). The 4-H inquiry in action model. Corvallis, OR: Oregon State University. Available at: http://oregon.4h.oregonstate.edu/sites/default/files/inquiry_in_action_model.pdf

Bruner, J. (1966). *Toward a theory of instruction*. Cambridge, MA: Belknap Press of Harvard University Press.

Campbell, P.B. (2008). Evaluating Youth and Community Programs. In Friedman, A. (Ed.). *Framework for evaluating impacts of informal science education projects*. Online available at: http://informalscience.org/researches/EvalGuide_work.pdf

Dierking, L.D. (2008). Evidence and categories of ISE impacts. In Friedman, A. (Ed.). *Framework for evaluating impacts of informal science education projects*. Online available at: http://informalscience.org/researches/EvalGuide_work.pdf

Eccles, J., & Gootman, J. (Eds.). (2002). *Community Programs to Promote Youth Development*. Washington, DC: National Academy Press.

Horton, R.L., Gogolski, J., & Warkentein, C. (2007). Science, engineering, and technology (SET) programming in the context of 4-H youth development. The Ohio State University 4-H Youth Development Program: Columbus, OH. On-line Available at: <http://www.ohio4h.org/publications/experience.html>

Kline, R.B. (2005). *Principles and Practice of Structural Equation Modeling* (2nd Edition). New York: The Guilford Press.

Kolb, D.A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.

Mielke, M., LaFleur, J., & Sanzone, J. (2010). *4-H science, engineering, and technology initiative: Youth engagement, attitudes, and knowledge study*. Washington, DC: Policy Studies Associates, Inc.

Riley, D. & Butler, A. (2012). *Priming the pipeline: Lessons from promising 4-H science programs*. Washington DC: Policy Studies Associates, Inc.

© Copyright of Journal of Youth Development ~ Bridging Research and Practice. Content may not be copied or emailed to multiple sites or posted to a listserv without copyright holder's express written permission. Contact Editor at: patricia.dawson@oregonstate.edu for details. However, users may print, download or email articles for individual use.

ISSN 2325-4009 (Print); ISSN 2325-4017 (Online)