

Monitoring Structured Experiences during Youth Programs: Development of Brief Measures of Perceived Value and Engagement

Gary D. Ellis

Texas A&M University
Gellis1@tamu.edu

Allen S. Taggart

Texas A&M University
ASTaggart@tamu.edu

Jill Martz

Texas A&M University
Jill.Martz@tamu.edu

Toby Lepley

Texas A&M University System AgriLife Extension Service
Toby.Lepley@ag.tamu.edu

Tazim Jamal

Texas A&M University
tjamal@tamu.edu

Abstract: A brief questionnaire that can be used to routinely monitor the quality of structured experiences for youth is discussed in this paper. Structured experiences are discrete periods of time in which youth gather for activity under the supervision of adult or youth leaders. Four-item measures of perceived value and engagement were created. A questionnaire including these measures along with items from the 4-H Common Measures was administered to 219 youth from 11 4-H clubs. Data were analyzed for evidence of reliability and validity. Alpha reliability estimates were .82 and .71 for the two multiple-item monitoring instruments. Correlations (validity coefficients) ranged from .48 to .61. Multiple regression results were consistent with predictions. Significant relations were found between perceived value, engagement, meaningfulness and supportiveness of social climate and safety of social climate. Results thus suggest that these instruments may be appropriate for use in monitoring structured experiences for youth.

Introduction

Many of the world's most influential visionaries in management, among them Peter Drucker and W. Edwards Demming, have asserted that, "if you are not measuring it, you are not managing it." That adage is particularly notable to organizations that provide meetings, competitions, practice sessions, classes, and related "structured experiences" (Duerden, Ward, & Freeman, 2015) for youth. The extent to which structured experiences are engaging and valued by youth has substantial impact on attrition and on long-term developmental outcomes. Yet, instruments currently available to measure outcomes are suited for intensive and long-term summative evaluations of program processes, structures, and results (Yohalem & Wilson-Ahlstrom, 2010) rather than each meeting, competition, practice session, or class period. The availability of tools for monitoring structured experiences for youth would allow organizations to establish goals per session, plot scores over time, and evaluate strategies they create to enhance or sustain the situational quality of individual sessions. The purpose of this study, then, was to develop brief measures of youths' perceived value of structured experiences, their situational engagement during structured experiences, and the meaningfulness of structured experiences.

Background

The vast array of out-of-school time programs available to today's youth afford rich opportunity for youth to grow physically, emotionally, socially, and spiritually, and to develop skills that will help them be effective professionals, family members, and citizens. But, youth must sustain participation in those programs if positive effects are to occur (Mead, Rodriguez, Hirschl, & Goggin, 1999; Simpkins, Chaput, Little, & Weiss, 2004), and attrition from youth programs is often substantial. In youth sports programs, for example, approximately 35% of the estimated 45 million participants quit each year, leading to a drop-out rate of 70%-80% by age 15 (Merkel, 2013).

Attrition has also been a longstanding issue in 4-H. The attrition problem in 4-H was first documented over 30 years ago. In 1983, Hartley reported that approximately 45% of new 4-H members become drop-outs after their first year. Twenty years later, attrition continued to be a problem. In 2005, Harder, Lamm, Lamm, Rose, and Rask reported estimated dropout rates of almost 40% for 18-year old members, over 20% for 17 year-olds, and approximately 18% for 16 year olds. Anecdotal accounts from 4-H leaders indicate that attrition continues to be a significant concern.

Attrition can be attributable to many factors, some of which are beyond the ability of 4-H leaders to control. But, it is reasonable to assume that engaging and meaningful structured experiences will be valued by participants, leading to sustained participation over time. Structured experiences are discrete periods of time in which youth gather for activity or collaboration under the supervision of adult or youth leaders. Examples of structured experiences are club meetings, competitions, practice sessions, livestock shows, class sessions, and meetings in which youth collaborate on projects.

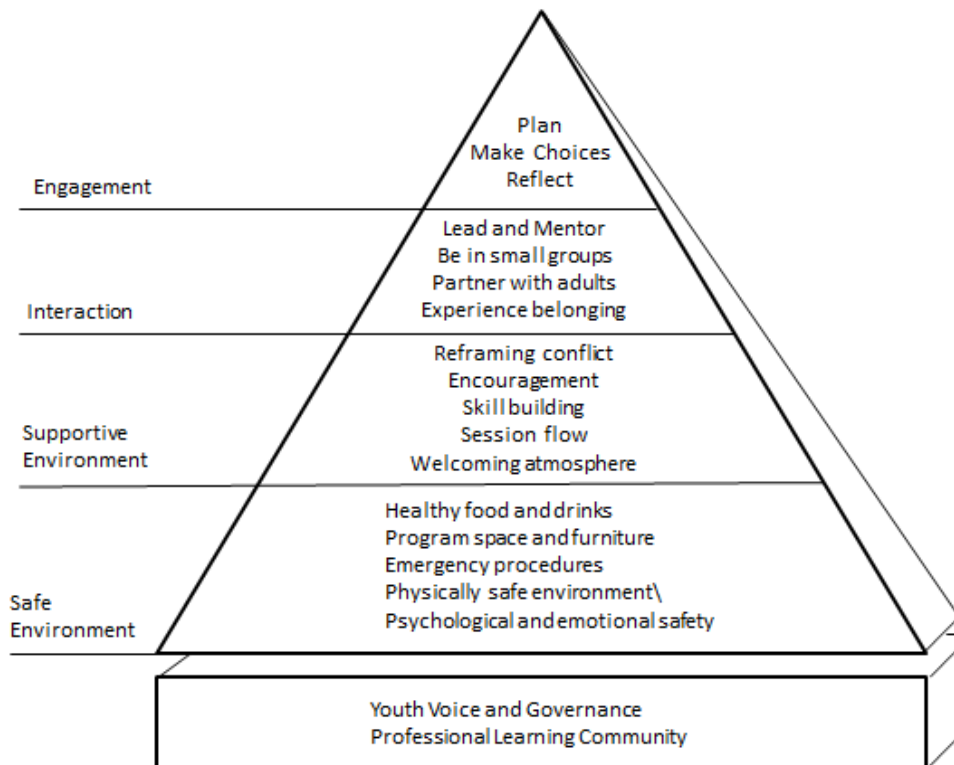
Programs designed to increase youths' perceived value and engagement have been shown to reduce attrition rates. Harder, et al. (2005) reported a 21% increase in retention as a result of a program in which 4-H leaders first surveyed members to determine topics valued by youth and then designed structured experiences to address those. The contextual elements of

structured experience are also very important. Hartley (1983) found that 4-H members value opportunities to “to be with friends,” “to have fun,” “to make or do something,” and “receive praise.” Thus, designing structured experiences that address youths’ values and that highlight opportunities for engagement --through fun, praise, creation, and being with friends-- can be expected to decrease attrition.

Youth development scholars have proposed additional features of structured experiences they consider to be essential to quality. The National Research Council and Institute of Medicine (Eccles & Gootman, 2002), for example, proposed that all structured experiences for youth should be characterized by safety, structure, supportive relationships, opportunities to belong, positive social norms, support for efficacy and mattering, opportunities for skill building, and integration of family, community, and school. Many of the same structured experience features are highlighted in the HighScope Educational Research Foundation’s (2005; see also Smith, Akiva, Arrieux, & Jones, 2006) “Pyramid of Program Quality:” professional learning community, youth voice and governance, safe environment, supportive environment, social interaction, and engagement (Figure 1). The National 4-H Council identifies four youth development outcomes that signal quality: communication, connections, contributions, and positive choices (Lewis, Horillo, Widaman, Worker, & Trzesniewski, 2015). Other approaches (Granger, et al., 2007; Durlak & Weissberg, 2007) show that programs can best achieve targeted developmental outcomes by designing structured experiences that are explicit, sequenced, focused, and targeted.

Figure 1

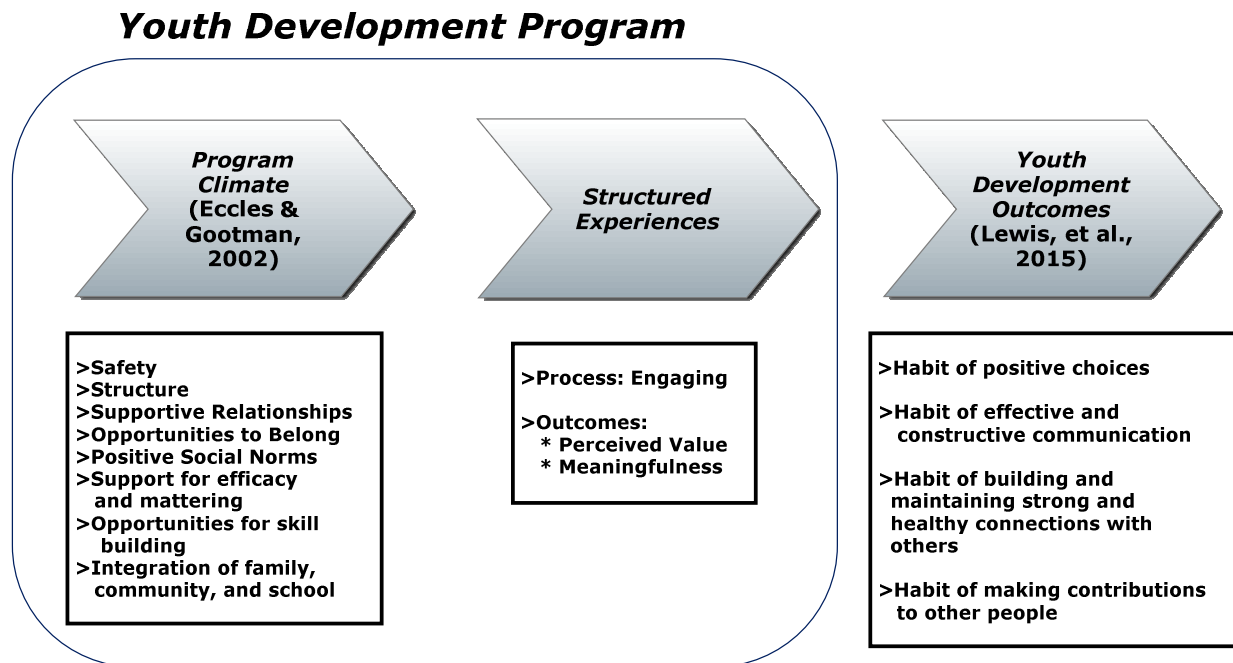
HighScope Educational Research Foundation’s
Pyramid of Program Quality



It is helpful to distinguish between the *process* features of youth development programs and *outcomes* of those programs and processes. The features identified by the National Research Council and Institute of Medicine (Eccles & Gootman, 2002) focus on process; they reflect the *climate* (e.g., Ehrhart, Schneider, & Macey, 2014) that the Research Council believes ought to be created and maintained throughout all structured experiences that comprise a youth program. Specifically, from the perspective of the Research Council, a quality youth program is one that consistently provides safety, structure, supportive relationships, opportunities to belong, positive social norms, support for efficacy and mattering, opportunities for skill building, and integration of family, community, and school. The presence and effect of such a climate should be manifest in each of the individual structured experiences of the organization, whether those are meetings, competitions, collaborative work, classes, or practice sessions. Given a quality climate and effective “staging” (e.g., Ellis & Rossman, 2008; Rossman & Schlatter, 2011), each structured experience should be engaging and should yield perceived value and meaning among the youth who participate. Exposure to the program “climate” and the series of structured experiences over time may be expected to yield “youth development outcomes” (Lewis, et al., 2015), which, from the perspective of the National 4-H Council are the following habits: effective communication, building and maintaining positive connections with others, making positive choices, and contributing to the well-being of other people. Figure 2 provides a graphic illustration of this process.

Figure 2

Relations among Program Climate, Structured Experiences, and 4-H Youth Development Outcomes



If adult leaders are to use such strategies to elevate the quality of structured experiences, measurement tools are needed to monitor the structured experience outcomes and assure that techniques are effective. Researchers have developed a variety of tools for assessing quality of

programs, but those measures focus on the program climate or outcomes and are also too lengthy to use to routinely monitor the quality of individual structured experiences within those programs. The Youth Program Quality Survey (Silliman & Schumm, 2013), for example, includes 24 items measuring five dimensions: positive emotional climate, empowered skill-building, expanding horizons, structure, and negative experiences. Another example is a 20-item measure developed by Tiffany, Exner-Cortens, and Eckenrode (2012), which measures four dimensions: personal development, voice and influence, safety and support, and community engagement. Administering a questionnaire of 20 items or more following every structured experience would in itself be an intrusive deterrent to experience quality.

Further, most of the instruments available require trained observers or are designed to be completed by program staff (Yohalem & Wilson-Ahlstrom, 2010). Observers and staff may not be the best option for evaluating structured experiences. Use of staff and external evaluators raises two concerns. One of these is the labor cost. Observers must be trained and data collection through participant observation can be a very tedious and time-consuming process. Perhaps more importantly, it is not clear that observers can accurately measure subjective states (e.g., engagement, perceived value, meaningfulness) of youth they observe. As Silliman and Schumm (2013, p. 7) noted, "...youth are the most authentic evaluators of their own experience." A brief, minimally intrusive questionnaire that can be administered at the conclusion of each structured experience is needed. Thus, the purpose of this study was to construct and evaluate a minimally intrusive questionnaire that could be used to routinely monitor three indicators of quality of 4-H meetings: perceived value, engagement, and meaningfulness.

Method

Population and Sample

Following Institutional Review Board approval of our protocol, we randomly selected 13 4-H clubs in central Texas and invited the club managers to participate in the study. As an incentive, three service-learning grants of \$500 were offered to each club in which 75% of members participated. Of the 13 clubs selected, 11 accepted the invitation to participate.

Two-hundred nineteen youth from these 11 clubs participated in the study. The number of participants per club ranged from 7 to 40 ($M=19$, $SD=10.44$). Their average age was 12.40 years ($SD=2.72$). Sex was measured by the researchers counting boys and girls present at each meeting. Logistical matters and unexpected complexities limited the counts to 173 participants. Eighty-seven of those counted were males and 86 were females.

The number of adult leaders present (apart from the research team) ranged from one (one club) to five (one club). Four clubs had two leaders present, three clubs had three leaders present, and one club had four leaders present. Meetings averaged 34.30 minutes in duration ($SD=19.35$). The longest meeting was 65 minutes long and the shortest meeting was 10 minutes in duration. All meetings began with the traditional 4-H opening ritual.

Measures

Perceived value

In her seminal paper on perceived value, Zeithaml (1988) defined that concept as "the consumer's judgment about the superiority or excellence of a product." Building on her

definition, we constructed a four-item scale that can be administered at the conclusion of structured experiences. We defined perceived value as the individual's degree of contentment with her or his decision to participate in the structured experience. High scores indicate that youth consider their choice to participate in the structured experience to be superior to other options that could have been pursued or, more specifically, an excellent choice for their investment of time. Four items were included:

- This was an excellent use of my time.
- I am glad I chose to do this.
- I made a good choice when I decided to do this.
- I wish I had spent my time doing something else.

Response options were "true," "mostly true," "neither true or false," "mostly false," and "false." A perceived value score was calculated by assigning a value to each response (5 for true, 4 for mostly true, 3 for neither, 2 for mostly false, and 1 for false) and summing across the four items. The fourth item listed above was reverse coded.

Engagement

Student engagement has been referred to as "...the glue, or mediator, that links important contexts—home, school, peers, and community—to students and, in turn, to outcomes of interest" (Reschly & Christenson, 2013, p. 3). This definition embraces a complex, broad-based system involving parents, teachers, and other community elements that focus students on learning outcomes. But, engagement can also be applied to immediate, "task-specific," structured experiences. Reeve (2013, p. 150), for example, defined task-specific engagement as "a student's active involvement in a learning experience" and pointed out that situational engagement has behavioral, emotional, cognitive, and agentic elements. Ridley, McWilliam, and Oats (2000, p. 134) defined task-specific engagement as "the amount of time children spend interacting with the environment (with adults, peers, or materials) in a developmentally and contextually appropriate manner." Our construction of engagement follows from these task-specific approaches. It yields a self-report of the fraction of time youth consider themselves to have been engaged during the structured experience. Specifically, defined situational engagement as the extent to which participants report active motivational involvement in a structured experience. Four items are included, each following the stem, "During this meeting...":

- I felt excited about things we were doing
- I felt curious about something.
- I felt I was being useful.
- I felt important.

Participants used a "slider scale" to respond to these items; they placed a mark on a line between two anchor points to represent their status relative to each item. Anchor points were "none of the time" and "all of the time." Scores were derived by measuring the distance between that mark and the "none of the time" response. That distance was divided by the total length of the line, resulting in an estimate of the percentage of time during the structured experience that the research participant felt engaged. A total score was derived by summing across the four item scores and dividing by four (the number of items).

Meaningfulness

Many structured meetings target learning or long-term developmental outcomes. Participants may be expected to learn new knowledge or a new skill, or the structured experience may be intended to invite youth to reexamine their values or strengthen commitments. Evidence of a structured experience having created such an effect would be found in an individual leaving the meeting thinking about something that she or he learned. Meaningfulness was thus defined as the extent to which the structured experience yielded active thought about something learned during that experience. A single item was used: "I am still thinking about something I learned." A five-point response format was used "True," "Mostly true," "Neither True or False," "Mostly False," and "False."

Items from the 4-H "Common Measures"

The monitoring tool we set out to develop thus included a total of nine items: four items each for perceived value and engagement and one item for meaningfulness. If these instruments are appropriate for monitoring quality of structured experiences, their scores must change when the elements of program climate change. In other words, criterion-related evidence of validity of the monitoring tools should be reflected in correlations between those measures and measures of program climate. As such, for the current study, we included seven items representing various facets of the program climate:

- All youths were friendly to me
- The adult leader(s) cared about me
- I know who to go to if I had a problem
- I felt free to share my ideas
- I felt bullied
- Other people made fun of me
- I felt left out

A slider scale was used for responses. Anchor points were "None of the Time" and "All of the time." Item scores were derived by measuring the distance from the "None of the Time" anchor to the point at which participants' marks crossed the scale and dividing by the total length of the line.

Factor analysis was used to construct measures of dimensions among the items. Factors were rotated to simple and orthogonal structure using Kaiser's varimax criterion. Results are presented in Table 1. Two factors explained 61% of the variance. The two-factor solution corresponds well to the "safe environment" and "supportive environment" dimensions of the HighScope Educational Research Foundation's "Pyramid of Program Quality" (HighScope Educational Research Foundation, 2005; Smith, Akiva, Arrieux, & Jones, 2006), so the factors were named accordingly. Factor 1 was named "supportive environment," and included "the adult leaders cared about me" (.76), I knew who to go to if I had a problem (.74), "all youths were friendly to me" (.71), and "I felt free to share my ideas" (.68). Factor 2, "unsafe environment," included "I felt bullied" (.89), "other people made fun of me" (.87), and "I felt left out" (.64). Scores for each participant on each factor were calculated for evaluation of validity of the measures of perceived value, engagement, and meaningfulness.

Table 1Factor Analysis of Items Based on the 4-H Common Measures Project, *N*=219

Item	Factor 1: Supportive Environment	Factor 2: Safe Environment	Communality	
The adult leaders cared about me.	.76	-.08	.58	
I knew how to go do if I had a problem.	.74	.01	.55	
All youths were friendly to me.	.71	-.31	.60	
I felt free to share my ideas.	.68	-.05	.47	
I felt bullied.	-.01	.89	.77	
Other people made fun of me.	<-.01	.87	.79	
I felt left out	-.33	.64	.51	
	Eigenvalue	2.20	2.03	4.24
	Percent Variance Explained	32%	29%	61%

Procedure

Data were collected on-site by members of the research team. Parents and youth reviewed informed consent and assent forms upon arrival at the meeting site, and those who choose to do so signed the forms. All participants were asked to complete a questionnaire immediately following conclusion of the meeting. Two forms of the questionnaire were used. One was an electronic format, accessed through smart phones and the other was a paper questionnaire. Although use of a common medium would be ideal, logistical challenges did not permit that approach. It is notable that Silliman and Schumm (2013) used both electronic and paper and pencil approaches to data collection in construction of their Youth Assessment of Program Quality questionnaire.

Method of Data Analysis

Data were analyzed in terms of reliability and validity. Cronbach's alpha was calculated as the measure of reliability. Criterion-related evidence of validity was assessed in three phases. In the first phase, Pearson correlation coefficients were calculated among the three monitoring instruments (perceived value, engagement, and meaningfulness). Significant, positive correlations were hypothesized for each relation.

The second phase involved examination of relations between the factor scores representing the two dimensions of the Pyramid of Program Quality (Smith, et al., 2006) and each of the monitoring tools. Three multiple regression models were constructed. Each of the monitoring tools (engagement, perceived value, and meaningfulness) was regressed on the two factor scores (i.e., supportive environment and safe environment) and their interaction effect (a product vector). Validity is indicated by the significance, strength, and direction of the relation between each of the three monitoring scales and the two predictors. It is reasonable to expect that participants will report higher perceived value and greater engagement when meetings are perceived as being both safe and supportive (e. g., Eccles & Gootman, 2002; Smith, Akiva, Arrieux, & Jones, 2006).

The third phase of the analysis involved development of a classification of meeting types. It is reasonable to assume that engaging structured experiences that are valued by participants

have potential for more substantial impact on learning and development, and thus meaningfulness, than structured experiences that yield lower levels of perceived value and engagement. Given this assumption, four structured experience types were defined, using median-splits of engagement and perceived value:

- Positive Youth Development Structured Experience (high perceived value and high engagement)
- Instrumental Structured Experience (high perceived value and low engagement)
- Hedonic Structured Experience (low perceived value and high engagement)
- Failed Structured Experience (low perceived value and low engagement)

Meaningfulness means for each meeting were then plotted across the identified structured experience types. The positive youth development structured experience type was expected to yield the highest mean, and the failed structured experience type was expected to have the lowest mean.

Results

Reliability

The Cronbach alpha estimate of reliability of the four-item measure of perceived value was .82. Descriptive statistics are presented in Table 2. The mean was 4.41 on the five-point scale and the distribution was negatively skewed and leptokurtic (i.e., more peaked than the normal curve). Examination of a histogram revealed that the extreme kurtosis and the negative skewness were the result of large numbers of participants choosing the highest possible score for perceived value (i.e., 5). Eighty-three participants (37.9%) chose the highest possible score on all four perceived value items.

Table 2 also includes reliability and descriptive statistics for the engagement scale. The reliability coefficient was .71. The mean shows that participants reported being engaged during approximately 69% of the meeting. The distribution was negatively skewed and platykurtic ("flat"), but skewness and kurtosis of the distribution of engagement scores was much smaller in magnitude than the distribution of perceived value scores.

The average meaningfulness score was 3.82 on the five-point scale. The distribution had negative skewness and was leptokurtic. Respective item scores ranged from 5 to 1. The departure from normality was much less dramatic for meaningfulness than perceived value.

Table 2

Central Tendency, Dispersion, and Shape of the Distributions of Perceived Value, Engagement, and Meaningfulness

Scale	<i>N</i>	<i>N</i> items	Mean	Std Dvn	Skewness	Kurtosis	Alpha
Perceived Value	219	4	4.41	.77	-2.01	4.76	.82
Engagement	219	4	.69	.22	-.70	-.10	.71
Meaningfulness	219	1	3.82	1.18	-.95	.14	----

Validity

Bivariate correlations

All correlation coefficients supported criterion-related evidence of validity. All coefficients were significant ($p < .001$), positive, and moderate in strength:

- Perceived Value and Engagement: $r = .52$
- Perceived Value and Meaningfulness: $r = .61$
- Engagement and Meaningfulness: $r = .48$

Multiple Regression Models

Table 3 and Figure 3 provide a summary of the results of the three multiple regression analyses. Results provide criterion-related evidence of validity. As predicted, the measures of supportive environment and unsafe environment are significant ($p < .01$) predictors of each of the three monitoring scales. The direction of the relations is also consistent with predictions. The positive regression (beta) coefficients associated with supportive environment show that as supportive environment increases, scores on the monitoring scales increase. The negative regression coefficients for unsafe environment show, as predicted, as scores on unsafe environment increase, scores on the monitoring scales decrease. R^2 values are .18, .30, and .09 for perceived value, engagement, and meaningfulness, respectively.

The interaction effect was significant in the model predicting perceived value. A plot was constructed to facilitate interpretation of that effect. As Figure 3 shows, the slope of the regression of perceived value on supportive environment is steeper for individuals at higher levels of unsafe environment, as compared to the slope for individuals with low levels of unsafe environment. In more simple terms, in the presence of an unsafe environment, the presence of a supportive environment becomes much more important in determining perceived value.

Meaningfulness per Structured Experience Type

The result of the analysis of meeting types is presented in Figure 4. The four quadrants represent each of the four meeting types: positive youth development, instrumental, hedonic, and unproductive. The means in each cell are the average meaningfulness score, per club. As predicted, the positive youth development structured experience yielded the highest mean and the unproductive structured experiences produced the lowest meaningfulness mean.

Table 3

Regression of Monitoring Scales on Supportive Environment and Unsafe Environment

Scale	Supportive Environment (S)		Unsafe Environment (U)		S by U Interaction		R^2
	<i>beta</i>	<i>t</i>	<i>beta</i>	<i>t</i>	<i>beta</i>	<i>t</i>	
Perceived Value	.25	4.01*	-.28	-4.39*	.14	3.20*	.18
Engagement	.53	9.11*	-.16	-2.76*	<.01	-.02	.30
Meaningfulness	.21	3.11*	-.19	-2.82*	.08	1.75	.09

* $p < .01$

Figure 3

Interaction Effect: Regression of Perceived Value on Supportive Environment at High (+1 Standard Deviation) and Low (-1 Standard Deviation) Levels of Unsafe Environment

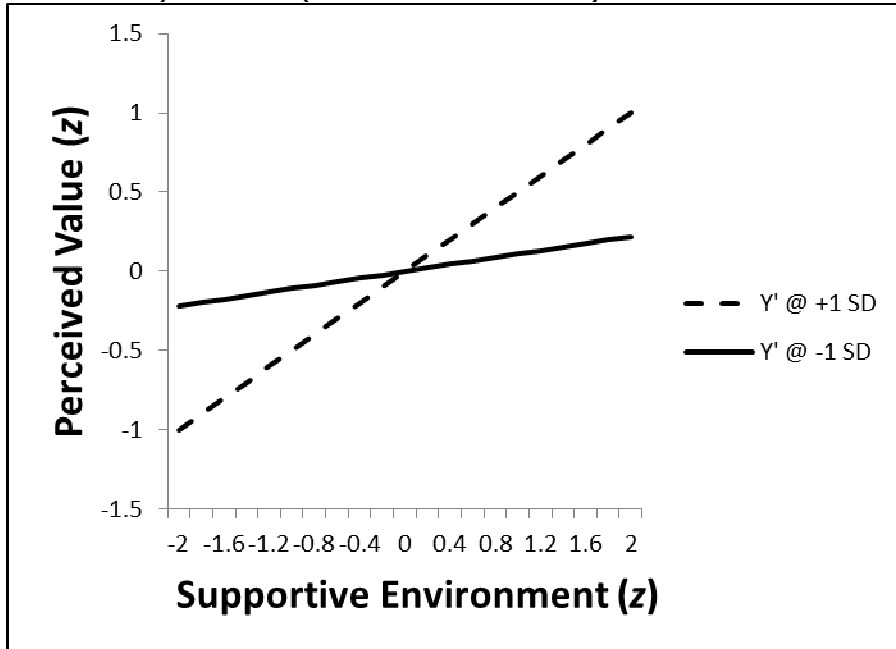
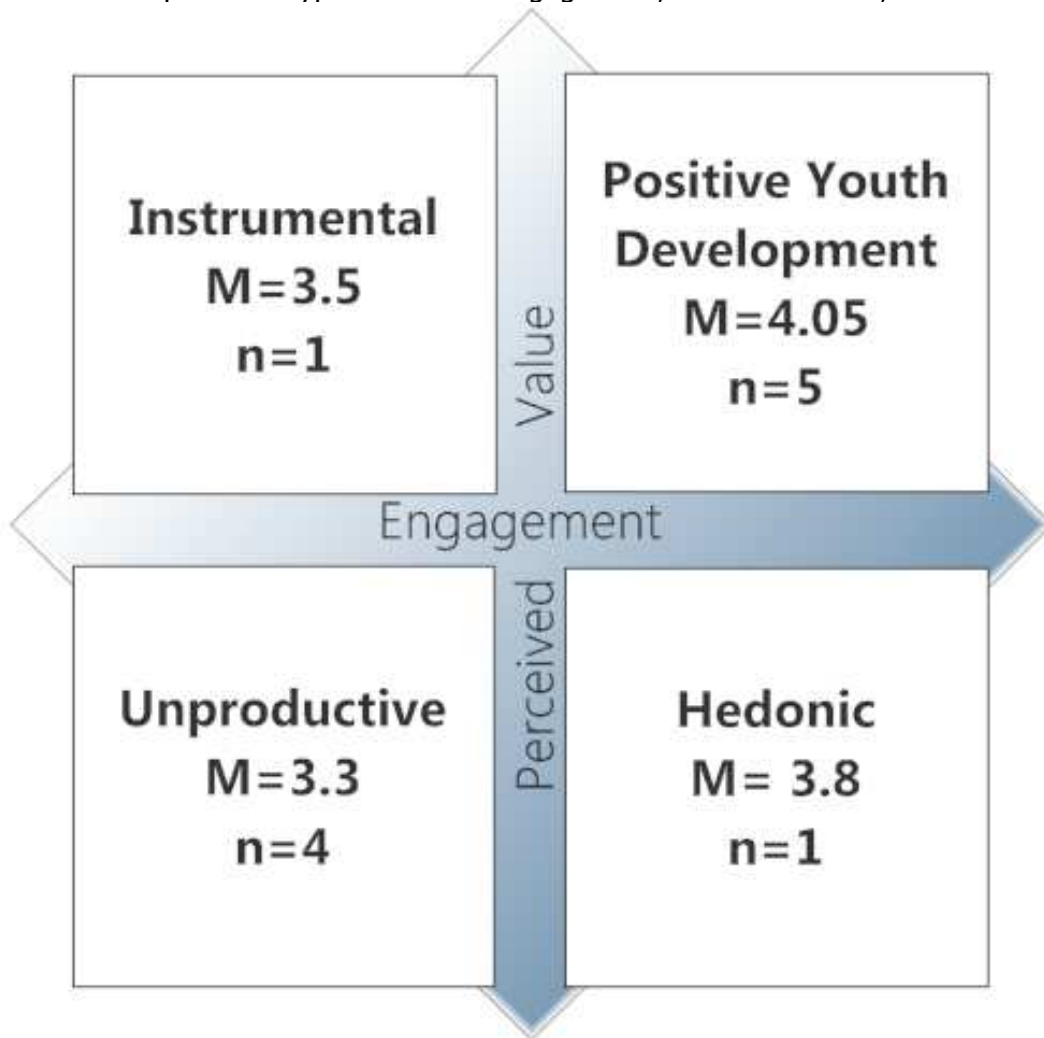


Figure 4

Structured Experience Types based on Engagement, Perceived Value, and Meaningfulness



Discussion

This study was designed to produce tools that can be used to routinely monitor structured experiences of youth (Figure 5). Two, four-item questionnaires were created to measure perceived value and engagement. A single item was constructed to measure meaningfulness. Data were collected from 219 4-H members from 11 different clubs in Texas on these instruments and seven items from the 4-H "Common Measures" project. The perceived value questionnaire and the engagement questionnaire were both found to be reliable. Criterion-related evidence of validity was found in relations among scores on the monitoring tools and relations between those tools and two factors corresponding to the Pyramid of Program Quality (Smith, et al., 2006). Thus, results support the validity of the measures and also provide evidence of validity of the models of quality youth development programs that include components of safety and support.

Figure 5
Monitoring Instrument

My experience at today's meeting!

Please draw an X on the line that best shows your experience. For example, if you were happy for about half of the time you were here, you would put an X at the middle of the line, like this:

During this meeting...

I was happy.
 None of the time -----X----- All of the time

I felt excited about what we were doing
 None of the time ----- All of the time

I felt curious about something
 None of the time ----- All of the time

I felt I was being useful
 None of the time ----- All of the time

I felt Important
 None of the time ----- All of the time

Please mark the circle showing how you feel about each of the following.

	True	Mostly True	Neither True or False	Mostly False	False
<i>I am still thinking about something I learned.</i>	0	0	0	0	0
<i>I wish I had spent my time doing something else.</i>	0	0	0	0	0
<i>I am glad I chose to do this.</i>	0	0	0	0	0
<i>I made a good choice when I decided to do this.</i>	0	0	0	0	0
<i>This was an excellent use of my time.</i>	0	0	0	0	0

The uniqueness of this study is its focus on the immediate, structured experience of participating in a point-of-service (Smith, Peck, Denault, Blaxevski, & Akiva, 2010) encounter as part of a youth program. A fundamental assumption in proposing use of the measures of engagement, perceived value, and meaningfulness as monitoring tools is that their scores will change as key features of the program climate change. Results of this study suggest that this covariation occurs for two dimensions of climate: supportiveness of the environment and safety of the environment. But, several additional features of the climate exist that were not measured in the current study. Among these are opportunities for skill building; integration of family, community, and school; and support for building self-efficacy (Eccles & Gootman, 2002; Smith et al., 2006). Also, it is not clear how the monitoring measures fit within paradigms for directing structured experiences to accomplish targeted outcomes (e.g., Granger, et al., 2007; Durlak & Weissberg, 2007). Future research is needed to determine how scores on the monitoring tools are related to additional dimensions of climate and approaches to designing experiences for targeted outcomes.

The data supported the assumption that perceived value, engagement, and meaningfulness would decrease as safety decreases. But, it is very important to note that our measure of safety was limited to three items focusing on social or interpersonal risk: bullying, being “left out,” and “being made fun of.” Several important dimensions of risk were not assessed. Among these are such risks as physical injury, developing poor health habits, and establishing relations with peers who are not positive role models or influences. Intrapersonal risks were also not measured. Examples include threats to key beliefs about self and encountering conflicts that involve deep-seated values. Future research is needed to examine relations among additional dimensions of safety and the three monitoring measures.

A number of additional issues remain. Our structured experience monitoring approach rests heavily on the assumption that perceived value, engagement, and meaningfulness are significant agents in determining whether youth will sustain their involvement in programs. Although that assumption is fully intuitive – people continue to do activities that are rewarding and that meet their needs—it certainly warrants empirical investigation. Social validity (Wolf, 1978) is also pivotal. The usefulness of the instruments to program leaders should be tested. Efficient processes for collecting, analyzing, reporting data should be developed, and safeguards should be put in place for inappropriate use of scores. As an example, while it is very appropriate to investigate causes of a pattern of low scores or highly divergent scores from session to session, it is not appropriate to indiscriminately use those scores as a condemnation of the efficacy or commitment of program leaders. If scores are low, investigation of the root causes of those scores is appropriate. It is not appropriate to simply blame program leaders. This latter issue is the pivotal “consequential” dimension of validity of an instrument; instruments are only valid when they are used for the purpose for which they were intended (Messick, 1989).

Thus, this paper provides a point of departure toward routinely measuring the quality of structured experience. Additional research is needed to evaluate validity, usefulness, and scope of applicability. Perhaps research in this trajectory will lead to improved quality of programs, decreased attrition, and more powerful impacts for the millions of youth who participate in youth programs nationwide and worldwide.

References

- Duerden, M.D., Ward, P.J., & Freeman, P.A. (2015). Conceptualizing structured experiences. *Journal of Leisure Research, 47*(5), 601–620.
- Durlak, J.A., & Weissberg, R.P. (2007). The impact of after-school programs that promote personal and social skills. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.
- Eccles, J.S., & Gootman, J.A. (2002). *Community programs to promote youth development. Committee on Community-Level Programs for Youth; Jacquelynne Eccles and Jennifer Appleton Gootman, editors; Board on Children, Youth, and Families, Division of Behavioral and Social Sciences and Education, National Research Council and Institute of Medicine.* Washington, DC : National Academy Press.
- Ehrhart, M.G., Schneider, B., & Macey, W.H. (2014). Organizational climate and culture: An introduction to theory, research, and practice. East Sussex: Routledge.
- Ellis, G.D., & Rossman, J.R. (2008). Creating value for participants through experience staging: Parks, recreation, and tourism in the experience industry. *Journal of Park and Recreation Administration, 26*(4), 1–20.
- Granger, R., Durlak, J.A., Yohalem, N., & Reisner, E. (April, 2007). *Improving after-school program quality.* New York, N.Y.: William T. Grant Foundation.
- Harder, A., Lamm, A., Lamm, D., Rose, H., & Rask, G. (2005) An In-Depth Look at 4-H Enrollment and Retention. *Journal of Extension, 43* (5).
- Hartley, R, (1983). Keeping 4-H Members. *Journal of Extension, 21* (4).
- HighScope Educational Research Foundation (2005). *The Youth Program Quality Assessment: A research-validated instrument and comprehensive system for accountability, evaluation, program improvement.* Retrieved from <http://etools.highscope.org/pdf/YouthPQA.pdf>
- Lewis, K.M., Horrillo, S., Widaman, K., Worker, S.M., & Trzesniewski, K. (2015). National 4-H Common Measures: Initial Evaluation from California 4-H. *Journal of Extension, 53*(2), 1-36.
- Mead, J.P., Rodriguez, E., Hirschl, T.A., & Goggin, S.E. (1999, August). Understanding the difference 4-H clubs make in the Lives of New York youth: How 4-H contributes to positive youth development. *Cornell Cooperative Extension.* Ithaca, New York.
- Merkel, D.L. (2013). Youth sport: positive and negative impact on young athletes. *Open Access Journal of Sport Medicine. 4*, 151-160.
- Messick, S. (1989). Validity. In R. Linn (Ed.). Educational Measurement. (pp. 13-103). New York: Macmillan.

Reeve, J. (2013). A self-determination theory on student engagement. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 149–72). New York, NY: Springer.

Reschly, A. & Christenson, S.L. (2013). Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the Engagement Construct. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 3-19). New York: Springer.

Ridely, S., McWilliam, R.A., & Oats, C.S. (2000). Observed engagement as an indicator of child care program quality. *Early Education and Development*, 11 (2), 133-146.

Rossman, J.R., & Schlatter, B.E. (2011). *Recreation programming: Designing and staging leisure experiences* (6th ed.). Urbana, IL: Sagamore.

Silliman, B., & Schumm, W.R. (2013). Youth Program Quality Survey: Youth Assessment of Program Quality. *Marriage & Family Review*, 49(7), 647-670.
doi:10.1080/01494929.2013.803010

Simpkins Chaput, S., Little, P.M.D., & Weiss, H. (2004) Understanding and measuring attendance in out-of-school time programs. *Issues and Opportunities in Out-of-School Time Evaluation*, 7. Retrieved from www.hfrp.org/content/download/1099/48604/

Smith, C., Akiva, T., Arrieux, D., & Jones, M M. (2006). Improving quality at the point of service. *New Directions for Youth Development*, 2006(112), 93-108. doi:10.1002/yd.195

Smith, C., Peck, S.C., Denault, A., Blazeovski, J., & Akiva, T. (2010). Quality at the point of service: Profiles in After-School Settings. *American Journal of Community Psychology*, 45, 358-369. doi:10.1007/s10464-010-9315-z

Tiffany, J.S., Exner-Cortens, D., & Eckenrode, J. (2012). A new measure for assessing youth program participation. *Journal of Community Psychology*, 40, 277-291.

Wolf, M.M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11, 203-214.
doi:10.1901/jaba.1978.11-203

Yohalem, N. & Wilson-Ahlstrom, A. (2010). Inside the black box: Assessing and Improving quality in youth programs. *American Journal of Community Psychology*, 45: 350-357.

Zeithaml, V.A. (1988). Consumer perceptions of price, quality, and value: A means-ends model and synthesis of evidence. *Journal of Marketing*, 53(2), 2–22.

© 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)