



Received: 19 October 2015
Accepted: 28 January 2016
First Published: 05 February 2016

*Corresponding author: Corliss Bean
School of Human Kinetics, University of
Ottawa, 125 University Private, Ottawa,
Ontario, Canada K1N 6N5
E-mail: cbean@uottawa.ca

Reviewing editor:
Stefan Elmer, University of Zurich,
Switzerland

Additional information is available at
the end of the article

DEVELOPMENTAL PSYCHOLOGY | RESEARCH ARTICLE

Re-examining the youth program quality survey as a tool to assess quality within youth programming

Corliss Bean^{1*} and Tanya Forneris¹

Abstract: The Youth Program Quality Survey is a 24-item measure of program quality designed to evaluate participant perceptions of experiences in short- and long-term youth programs. The Youth Program Quality Survey was developed based on the National Research Council and Institute of Medicine's eight program setting features that can contribute to the positive development of youth. This measurement tool is quite new within the field and as such little research has been conducted to determine its validity and reliability. The current study is designed to examine three previously proposed factor structures with a sample of 391 youth between the ages of 10 and 18 who participated in 38 different youth programs (e.g. sport, leadership, in-school-mentoring). Confirmatory factor analysis results indicate model fit issues with all three proposed factor structures. Therefore, an exploratory factor analysis is performed to improve model fit, and a revised 4-factor 19-item model is proposed. An analysis of invariance by age shows that the measurement model did not vary between younger and older youth participants. Practical implications and areas of future research are offered.

Subjects: Children and Youth; Community Sport Development; Evaluation/Program Evaluation; Sports Development; Testing, Measurement and Assessment; Youth Sport

Keywords: adolescence; program quality; youth; evaluation; exploratory factor analysis

ABOUT THE AUTHORS

Corliss is a PhD candidate in the School of Human Kinetics at the University of Ottawa. Her research focuses within the domain of positive youth development, and specifically, the development, implementation, and evaluation of youth programs. Examining the influence of program quality on basic needs support and psychosocial outcomes is at the core of her study interests. Recently, she has been working with community organizations to aid in the facilitation of high-quality programs that enable youth opportunities to develop important life skills cbean@uottawa.ca.

Tanya Forneris is an associate professor in the School of Human Kinetics at the University of Ottawa. She has a background in sport and counseling psychology. Her research interests lie in positive youth development, and specifically, the development, implementation, and evaluation of life skills-based sport and physical activity programs to enhance youth development, particularly with marginalized populations.

PUBLIC INTEREST STATEMENT

The Youth Program Quality Survey is a measure of program quality designed to evaluate youth perceptions of experiences in short- and long-term programs. The survey was developed based on the National Research Council and Institute of Medicine's eight proposed setting features of program quality that aim to facilitate positive developmental outcomes. The current study provided a revised survey with sound psychometric properties that can be used by both researchers and programmers with youth from 10 to 18 years of age in various types of youth programming. Practical implications and areas of future research are offered.

1. Introduction

Each year, youth development programs serve millions of youth across North America and internationally (Eccles & Gootman, 2002; Mahoney, Larson, Eccles, & Lord, 2005; National Collaboration for Youth, 2011). Such programs have proved to be important with youth across North America, as research has indicated that more than 86% of youth in Canada and 57% of youth in the United States participate in at least one organized out-of-school time activity (Guèvremont, Findlay, & Kohen, 2014; United States Census Bureau, 2014). National youth-serving organizations (e.g. Big Brothers Big Sister, Boys and Girls Club, Boys and Girls Clubs of Canada, YMCA, YWCA) consistently offer the most out-of-school time youth development programs (Boys & Girls Clubs of Canada, 2008; Quinn, 1999; The Bridgespan Group, 2005), with the context of sport identified as the most popular type of program for youth to spend their time (Guèvremont, Findlay, & Kohen, 2008; Martelaer & Theeboom, 2006; United States Census Bureau, 2014). Research has shown that youth experience more positive developmental outcomes (e.g. increased relationships with adults and peers, life skills development, levels of engagement) when participating in organized out-of-school time activities compared to passive leisure activities (e.g. spending time with friends, watching TV, reading, listening to music; Barber, Eccles, & Stone, 2001; Benson, Scales, & Syvertsen, 2011; Eccles & Gootman, 2002; Employment & Social Development Canada, 2014). Therefore, because of the potential impact such programs can have on youth participants, it is critical to deliver high-quality programs for these youth.

Researchers argue that program quality is the best predictor of not only risk prevention, but also positive developmental outcomes, which are two predominant foci of youth programs (Catalano, Hawkins, Berglund, Pollard, & Arthur, 2002; Durlak, Mahoney, Bohnert, & Parente, 2010; Roth & Brooks-Gunn, 2003; Yohalem & Wilson-Ahlstrom, 2010). The Committee on Community-Level Programs for Youth within the National Research Council and Institute of Medicine (NRCIM) responded to the identified gap in which very little research exists outlining what programs can do to facilitate development. As such, research was compiled from the developmental science literature over the course of two decades and eight contextual features were identified as the most likely to promote positive developmental outcomes within youth programming (Eccles & Gootman, 2002). These elements include: (1) physical and psychological safety; (2) appropriate structure; (3) supportive relationships; (4) opportunities to belong; (5) positive social norms; (6) support for efficacy and mattering; (7) opportunities for skill-building; (8) integration of family, school, and community efforts. Some setting features have been identified as necessary components within youth programs, such as safety and structure, while others are considered higher order elements that may not be a focus within all programs, such as integration of family, school, and community efforts. Although each feature can contribute to development individually, there is often overlap between the features in which features work together to foster greater psychosocial outcomes (Eccles & Gootman, 2002). For example, programs that outline clear and consistent rules and expectations for youth demonstrate appropriate structure, yet these rules may also help to foster a physically and psychologically safe environment. Several researchers argue the more of the eight setting features that a program has, the greater the influence it will have on the positive development of youth (McLaughlin, 2000; Merry, 2000).

A number of common features, similar to those proposed by the NRCIM, have been outlined within the literature that can contribute to program quality (e.g. Durlak, Weissberg, & Pachan, 2010; Lerner, 2004; Roth & Brooks-Gunn, 2003). However, the program quality setting features proposed by the NRCIM have been the most widely used and researched. The Youth Program Quality Survey (YPQS; Silliman & Schumm, 2013), the measure of focus in this research, was developed based on the NRCIM's eight features of program quality outlined above that have been proposed to facilitate positive outcomes in youth programming. Holt and Jones (2008) outlined that there are few valid and reliable measures for youth developmental outcomes within program evaluation and even fewer measures available to understand the processes of program implementation. To date, the YPQS is the only known measure to utilize the youth perspective in assessing program quality. Other measures do exist to examine program quality from an external evaluator or a program leader

perspective (e.g. Youth Program Quality Assessment by High/Scope Educational Research Foundation, 2005; Program Quality Observation Scale by Vandell & Pierce, 2006), yet studies that assess program quality tend to be based on adult rather than youth perceptions of program quality (Urban, 2008; Yohalem & Wilson-Ahlstrom, 2010). Eccles and Gootman (2002) argue that the eight setting features need to fit well with the youth participants, as the features exist as an interaction with the program setting and are not independent from one another. Therefore, it is also critical to obtain youth participants' perspectives as youth voice is critical in program evaluation (Camino & Zeldin, 2002; Hamilton, Hamilton, & Pittman, 2004; Powers & Tiffany, 2006; Walker, 2007). A limitation of the current field is that there are few measures that attain the youth perspective regarding the quality of experiences within programs (Vandell, Larson, Mahoney, & Watts, 2015). As such, it is necessary to develop a structurally sound measure of program quality that provides youth voice in the evaluation process.

Previously referred to as the Youth Program Climate Survey (Silliman, 2008a, 2008b, 2008c; Silliman & Shutt, 2010), the YPQS (Silliman & Schumm, 2013) is a measure that aims to examine youth perceptions of program contexts with the goal of understanding if a program provides a climate that promotes positive youth development (Silliman, 2008b). The survey was developed to provide a "reliable, research-based, user-friendly tool to help youth leaders incorporate youth voice in assessment of the features in the process of program development, monitoring, and improvement" (Silliman & Schumm, 2013, p. 648). As the YPQS was developed based on the NRCIM's eight program setting features, the original measure consisted of eight subscales including: Safety, Support, Social Norms, Social Inclusion, Skill-building Opportunities, Self-efficacy, Structure, and Synergy with Family and Community. Three items were developed for each of the eight contextual features that have been suggested to promote positive youth development (Eccles & Gootman, 2002), resulting in a 24-item measure. This measure was used for all youth regardless of age.

Prior studies using the YPQS have indicated moderate to high instrument reliability with Cronbach's alphas (Cronbach, 1951) ranging from .60 to .96 when examining the total scale and individual factors (Silliman, 2008a; Silliman & Schumm, 2013; Silliman & Shutt, 2010). Previous research has also found the measure reliable with both child and teen samples (Silliman, 2008b; Silliman & Shutt, 2010). Moreover, Silliman and Shutt (2010) indicated few differences when examining the measure by age, gender, and race. However, more recently, in study conducted by Silliman and Schumm (2013), the authors divided the sample of youth participants into two categories based on previous research outlining that youth of different ages are at different developmental levels (Lerner, 2009). These two categories included younger youth between the ages of 10 and 13 and older youth between the ages of 14 and 17. From this, the authors conducted two exploratory factor analyses that resulted in the construction of two factor models (one for each age group). The factor model for younger youth (ages 10–13) outlined five subscales: (1) Positive Emotional Climate, (2) Empowered Skill-building, (3) Expanding Horizons, (4) Structure, and (5) Negative Experiences. The factor model for the older youth (ages 14–17) revealed similar, yet slightly modified subscales: (1) Empowered Skill-building, (2) Positive Values, (3) Expanding Horizons, (4) Adult Support, and (5) Negative Experiences. Therefore, based on Silliman and Schumm's (2013) findings, researchers utilizing this tool should distribute and analyze two different measures based on the age of participants within a program. This infers that for youth of different ages, different elements of program quality may be of importance. From this outlined research conducted with the YPQS, there have been three proposed factor models using the items from the YPQS identified within the literature to date (an eight-factor model for youth of all ages, a five-factor model for youth ages 10–13, and a five-factor model for youth ages 14–17). However, no further research has been conducted to test these models beyond their initial proposal as a result of exploratory factor analyses. Confirmatory factor analyses have not been conducted on any of the three proposed factor models and as a result there has been little advancement in validating the YPQS in the literature.

The previously conducted exploratory factor analyses outlined above were tested with what one may argue to be a homogenous sample in terms of the type of program context. All youth within

these studies were part of 4-H camps and conferences that were of short duration, lasting between 3 and 7 days. As this survey was proposed to evaluate participant perceptions of experiences in a variety of youth programs including those of both short and long duration (Silliman & Schumm, 2013), more research is necessary to examine the YPQS with youth in various types of programs and durations. Moreover, the factor structures proposed by Silliman and Schumm were based on particular age ranges which may not be in fact ideal for two reasons. First, the NRCIM's eight setting features were proposed with the purpose of encompassing program quality for youth programs globally and not by age group (Eccles & Gootman, 2002). Previous research has found that features of high-quality programs serving school-age children are also similar program quality features to those for adolescents (Vandell, Shumow, & Posner, 2005), highlighting the importance of having one program quality measure across youth of all ages. The majority, if not all, of research-based measures that examine the processes and outcomes of youth development programs are not broken down by age (e.g. Positive Youth Development Inventory (PYDI) by Arnold, Nott, & Meinhold, 2012; Youth Experience Survey (YES) 2.0 by Hansen & Larson, 2005). Second, in practice, youth programs have various combinations of ages involved. For example, some programs may have youth from ages 11 to 16 whereas others may have youth from 12 to 14 years of age. Having factor structures based on age range would make using the YPQS, or any other program quality measure, difficult as the particular programs would have to align with age groups proposed. As such, it would be most appropriate to have a validated measurement tool with a strong factor structure that can be universally utilized across all youth program contexts, regardless of age, to evaluate program quality.

Therefore, the purpose of this study was to conduct confirmatory factor analyses on three factor structures previously proposed by (Silliman, 2008c; Silliman & Schumm, 2013; Silliman & Shutt, 2010) with a large sample of youth involved in programs with various contextual program features and varying degrees of duration. Validating the YPQS with such a population would provide an important evaluation tool for programmers and administrators to assess perceived program quality through the eyes of youth participants. A secondary purpose of this study was to examine potential invariance by age with respect to the emergent factor structure of the YPQS. If the analysis of invariance conducted on the measurement model does not vary by age, this would indicate that the model (elements of program quality) does not vary between younger and older youth. As such, the model would be appropriate for all youth between 10 and 18 years of age, which would also indicate the identified elements of program quality would be appropriate for all youth within this age range.

2. Method

2.1. Context and participants

Three hundred and ninety-one youth (162 male, 229 female) who participated in 38 different programs within six youth organizations across south-eastern Ontario, Canada were involved in this study. Programs ranged from 6 weeks to 9 months in duration and were run between one and three times per week throughout its duration. Programs ran between 60 and 180 min in length ($M = 110$). Youth ranged in age from 10 to 18 years ($M_{\text{age}} = 14.09$, $SD = 2.16$). The sample was distributed as follows: Caucasian (68.7%), black (6.9%), Aboriginal (2.1%), Arabic (6.6%), Asian (6.9%), multiracial (6.4%) and 2.4% of the sample did not disclose their ethnicity. Participation within a given program ranged from 6 to 29 youth, with the average program having 11.5 ($SD = 4.7$) youth involved. Youth who completed the questionnaire were identified by program leaders as regular participants (i.e. participating in the program 75% of the time or more). Table 1 outlines additional characteristics of programs involved within the study sample, including the organization and program types, as well as the type of participants involved in each program and within the six organizations.

2.2. Measures

The YPQS is a 24-item measure that examines participants' ratings of their experiences within youth programs (Silliman, 2008c; Silliman & Schumm, 2013; Silliman & Shutt, 2010). As mentioned above, this measure focuses on examining program quality from the perspective of youth using the eight contextual setting features of youth programs shown to promote positive youth development

Table 1. Program characteristics of study sample

Programs characteristics		# of programs (sites)	# of youth
Type of site	For-profit	21 (2)	194
Type of program	Not-for-profit	17 (7)	197
	Leadership	6 (2)	72
	Mentoring	2 (1)	24
	Sport	30 (5)	295
	Competitive	20 (2)	194
	Recreational	10 (3)	101
Type of participants	All girls	16 (3)	159
	All boys	6 (3)	59
	Mixed gender (boys and girls)	16 (3)	173

identified by the NRCIM as a guiding framework (Eccles & Gootman, 2002). From these eight setting features, the YPQS was originally designed with eight subscales including: Safety, Support, Social Norms, Social Inclusion, Skill-building Opportunities, Self-efficacy, Structure, and Synergy with Family and Community (Silliman, 2008c). Each subscale contained three items. Based on previous use of the scale, all items were answered on a four-point scale from 1 (*strongly disagree*) to 4 (*strongly agree*). With the current sample, Cronbach's alpha was conducted on the initial 8-factor, 24-item measure proposed by Silliman (2008c). Factors ranged from .11 to .74 indicating unacceptable to acceptable internal consistency, with the majority of factors falling within the poor to unacceptable range (below .70; George & Mallery, 2003).

2.3. Procedures

Following ethical approval from the affiliated institution's Office of Research Ethics and Integrity, the lead researcher contacted various community youth organizations in south-eastern Ontario. As this was part of a larger study exploring program quality in youth programming, information about this study was communicated to interested community programmers and directors. Six organizations agreed to participate in the study and a total of 38 programs were identified based on interest. The researcher met with each program leader, and parents if available, to explain the study. The researcher then provided the parental consent forms and assured confidentiality prior to data collection. Youth who attained parental consent completed a paper version of the questionnaire at the end of the program (within the last two weeks of the program) with supervision from the first author, which provided opportunities for youth to ask questions if needed.

2.4. Data analysis

Using the structural equation modeling AMOS 23 software program (Arbuckle, 2014), three confirmatory factor analyses were conducted with a Maximum Likelihood estimation method on the data to assess fit for the three proposed models previously outlined: (1) the 8 factor, 24-item measurement model for youth of all ages, (2) the 5-factor, 24-item measurement model proposed by Silliman and Schumm (2013) for youth ages 10–13, and (3) the 5-factor, 24-item measurement model proposed by Silliman and Schumm (2013) for youth ages 14–17. However, as the main goal of this study was to create one factor structure that fit well across youth of all ages, all three of the models outlined above were tested using our entire sample (youth 10–18 years old). A ratio of 15:1 was exceeded for subjects-to-items ratio indicating that recommended requirements were attained (Everitt, 1975; Pedhazur, 1997). Latent factors were allowed to correlate, uniquenesses were not free to correlate, and a path from each latent variable to one of its indicator variables was constrained by assigning the value of 1.0. According to Awang (2012), in order to ensure uni-dimensionality of a measurement model, any item with a loading factor lower than .50 should be deleted. Model fit was assessed using a combination of indices: Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Standardized Root Mean Square Residual (SRMR), Root Mean Square Error of Approximation (RMSEA),

and the χ^2 statistic (Byrne, 2010; Hu & Bentler, 1998). Fit indices were deemed to indicate good model fit if: CFI and TLI > 0.90, SRMR < 0.08, and RMSEA < 0.05 (Byrne, 2010; Tabachnick & Fidell, 2013).

3. Results

Prior to analyses, data were screened for missing data, which indicated less than one percent of the data were missing. When less than 5% of data are missing, influences of missing data are negligible (Tabachnick & Fidell, 2013). Missing data were replaced with multiple imputations (Yuan, 2010). Data were then screened for outliers. No outliers were identified within the data-set. All variables were normally distributed and there were no instances of multi-collinearity ($r > 0.90$) throughout the data-set. Statistics for Variance Inflation Factor (VIF) and tolerance ranged from 1.074 to 2.008 which fell within the acceptable range (Hair, Black, Babin, & Anderson, 2010) for all variables.

As mentioned, three different models have been proposed within the literature: (1) one eight-factor model for youth of all ages, (2) one five-factor model for youth ages 10–13, and (3) one five-factor model for youth ages 14–17. Confirmatory factor analyses were conducted on all three of these models to test for model fit.

3.1. Confirmatory factor analysis for eight-factor model

Summary statistics for the confirmatory factor analysis indicated some issues with the model fit: CFI = .901, TLI = .878, SRMR = .0545, RMSEA = .056 (90% CI = .049 – .062), $\chi^2/df = 2.206$, and $\chi^2 = 494.032$, $p > .0001$. Although the model approached fitting the data adequately, there were several issues of collinearity between factors (e.g. .99) and low loading items to various factors (e.g. 8 items showed low loading <.50; Awang, 2012). Finally, diagnostics indicated it was not possible to retain a factor because it contained less than three items, as it has been identified that a factor with fewer than three items is generally weak and unstable (Costello & Osborne, 2011). Therefore, the factor and the remaining items were removed. Based on the outlined issues, it was deemed that the eight-factor structure was not suitable.

3.2. Confirmatory factor analysis for five-factor model based on youth sample (10–13 years old)

The proposed five-factor model identified by Silliman and Schumm (2013) for younger (10–13) youth was tested using our total sample to test the fit of the model. Summary statistics for the confirmatory factor analysis indicated some issues with model fit: CFI = .906, TLI = .893, SRMR = .0803, RMSEA = .052 (90% CI = .046 – .059), $\chi^2/df = 2.059$, and $\chi^2 = 498.184$, $p > .0001$. Although there were no issues with collinearity between latent factors, there were four low loading items (<.50). After removing these items one at a time starting with the lowest loading item, the model fit improved (CFI = .925, TLI = .911, SRMR = .0466, RMSEA = .054 (90% CI = .046 – .062), $\chi^2/df = 2.132$, and $\chi^2 = 341.168$, $p > .0001$); however, the model resulted in two factors with only two items each; thereby highlighting an unstable factor structure (Costello & Osborne, 2011). Therefore, it was deemed that this five-factor structure was not ideally suited for youth between the ages of 10 and 18 years because of the unstable structure.

3.3. Confirmatory factor analysis for five-factor model based on youth sample (14–17 years old)

The proposed five-factor model identified by Silliman and Schumm (2013) for older (14–17) youth participants was tested using our total sample to test the fit of the model. Prior to analysis, it was believed there may be issues with the previously outlined factor structure, as one factor within the model was comprised of only two items, violating the recommendation to have a minimum of three items loaded to one factor (Costello & Osborne, 2011). However, despite this, the model fit was tested. Results indicated that there were issues with the model fit: CFI = .887, TLI = .869, SRMR = .0678, RMSEA = .060 (90% CI = .053 – .067), $\chi^2/df = 2.388$, and $\chi^2 = 475.282$, $p > .0001$. Moreover, there were multiple correlations between latent factors as three exceeded .90 (Kline, 2010), suggesting issues with multicollinearity. Merging of these three factors would have resulted in a three-factor structure

comprised of a 15-item factor, a 2-item factor, and a 5-item factor. Moreover, there were several low loading items (e.g. .12) to various factors. As a result of the combination of issues identified, it was deemed that this structure was not a good fit to our sample of youth between the ages of 10 and 18 years.

3.4. Exploratory factor analysis

Based on the lack of support from the initial factor analyses of the three previously proposed structural models of the YPQS, an exploratory analysis was justified to modify the measurement model. The exploratory factor analysis, using an eigenvalue = 1.0 criterion with maximum likelihood factor analysis resulted in a five-factor model. However, this was minimized to a four-factor model as only one item had loaded to the fifth factor and did not load to any of the four other factors. Therefore, this item was removed. Attempts to extract more than five factors yielded minor factors with loadings on only one item. Extracted communalities ranged between .37 and .72. All but two items loaded on to at least one factor at $<.50$.

Factor analysis was conducted to test the results of the exploratory factor analysis for model fit. An iterative process guided by modification indices, parameter change estimates, and theoretical justifications was used to re-specify the 23-item measurement model (Garson, 2010; Kline, 2010). Each modification occurred one iteration at a time and parameter estimates were recalculated after each modification (Garson, 2010; Kline, 2010). We did not free cross-loadings; diagnostics indicating cross-loadings resulted in the subtraction of problematic items from the model one at a time. Items were considered for removal when standardized estimates fell below .50 (Awang, 2012). All scales in the re-specified model had factor loadings ranging from .51 to .72. A total of four items were removed during the re-specification process in addition to the one initial item removed in the previous identified step, resulting in a four-factor model with a total of 19 items. Results revealed a good model fit: CFI = .932, TLI = .920, SRMR = .0456, RMSEA = .037 (90% CI = .032 – .042), $\chi^2/df = 1.861$, and $\chi^2 = 543.42$, $p > .0001$. Table 2 outlines the descriptive statistics for the individual items of the re-specified model. There was only one instance of kurtosis, which was for an item identified as a negative perceived experience and therefore, was often rated a 1 on the four-point scale. Table 3 displays the descriptive statistics of the four final dimensions of the re-specified model and sample specific correlations between subscale scores for each pair of latent factors in the final retained measurement model; correlations were below .66. Building off Silliman and Schumm's (2013) work, four subscales were identified: Appropriate Adult Support and Structure, Empowered Skill-building, Expanding Horizons, and Negative Experiences. Subscale names were used or slightly modified based on these authors' previous classifications (Silliman & Schumm, 2013).

All eight of the NRCIM's program setting features proposed to foster positive youth development outcomes were represented in the revised 19-item, 4-factor model of the YPQS (see Table 4 for the breakdown by item corresponding to the NRCIM features). The first factor loaded heavily on five items related to Appropriate Adult Support and Structure and included items from three setting features outlined by the NRCIM, including Physical and Psychological Safety, Appropriate Structure, and Supportive Relationships (Eccles & Gootman, 2002). The second factor loaded primarily on eight items and was labeled Empowered Skill-building and contained items related to Opportunities for Skill-building, Positive Social Norms, Support for Efficacy and Mattering, and Integration of Family, School, and Community Efforts. The third factor, labeled Expanding Horizons, loaded on four items, including Opportunities to Belong, Positive Social Norms, and Integration of Family, School, and Community Efforts because the items referred to broadening views of and the world in context of family and community. Lastly, the fourth factor loaded heavily on three items that were all negatively worded. These items were consistent with NRCIM's three features including Physical and Psychological Safety, Opportunities to Belong, and Positive Social Norms.

Because previous research utilizing the YPQS proposed two models based on age of participants (Silliman & Schumm, 2013), an analysis of invariance was conducted to see if the four-factor model was a good fit for both age groups of youth (10–13 and 14–17). Analysis of invariance is a set of

Table 2. Descriptive statistics for individual items of the YPQS

Item	M	SD	Skew.	Kurt.
Adult Support and Structure				
1. Activities promoted healthy habits	3.32	.73	-.10	.92
2. Adults listened to what I had to say	3.30	.64	-.60	.56
3. I felt comfortable going to adults for advice	3.21	.73	-.63	.07
4. Rules and expectations were clear	3.36	.72	-1.04	1.04
5. Activities were just right for my age	3.31	.72	-.92	.80
Empowered Skill-building				
6. I learned to work with others as a team	3.50	.65	-1.23	1.60
7. Activities taught me to develop a plan to reach my goals	3.32	.72	-.86	.47
8. I was challenged to think and build skills	3.31	.68	-.79	.67
9. There were opportunities to learn new subjects	3.22	.78	-.75	.06
10. I felt that I could make a difference	3.14	.77	-.69	.20
11. I was encouraged to take responsibility	3.34	.72	-.91	.58
12. I gained a broader view of [name of program]	3.35	.71	-.94	.66
Expanding Horizons				
13. Serving others and volunteering was important	3.25	.69	-.77	.85
14. I learned to accept differences in others	3.47	.62	-.88	.38
15. I gained a broader view of the world beyond my community	3.13	.75	-.44	-.42
16. Activities were related to issues in my club, my family and my community	2.74	.93	-.24	-.83
Negative Experiences				
17. I was embarrassed or put down	1.64	.79	1.06	.39
18. Conflicts between people were a problem	1.881	.93	.78	-.35
19. I felt like I didn't belong	1.49	.75	1.65	2.43

Notes: YPQS variables were measured on a scale of 1–4. For further suggestions regarding the use of this measure, see Recommendations within the Discussion.

Table 3. Descriptive statistics of the four dimensions and inter-factor correlations of the re-specified model

Scale	α	M	SD	Skew.	Kurt	1	2	3	4
Adult support and structure	.798	3.30	.52	-.54	.17	–	.66**	.53**	-.34**
Empowered skill-building	.838	3.31	.51	-.74	.963		–	.61**	-.26**
Expanding horizons	.761	3.15	.54	-.12	-.40			–	-.19**
Negative experiences	.633	1.67	.63	1.05	1.36				–

** $p \leq .01$.

hierarchically ordered and increasingly restrictive tests of equality between groups. This analysis was comprised of three steps: (1) establishing the model for age groups individually, (2) testing the model with both groups simultaneously (configural invariance), and (3) examining the invariance of factor loadings between groups (metric invariance) using a chi-squared difference test. The measurement model revealed good fit as a baseline model for the younger (10–13 years old) sample: CFI = .896, TLI = .886, SRMR = .0714, RMSEA = .060 (90% CI = .043 – .076), $\chi^2/df = 1.513$, and $\chi^2 = 220.947$. Additionally, the measurement model also revealed good fit for the older (14–17 years old) sample: CFI = .933, TLI = .922, SRMR = .0459, RMSEA = .052 (90% CI = .043 – .060), $\chi^2/df = 2.007$, and $\chi^2 = 293.029$. All factor loadings were significant in both models. A test for configural invariance through a simultaneous multi-group model also showed good fit of the model to the data: CFI = .910,

Table 4. Items of the YPQS organized by subscales and corresponding NRCIM features

Subscale	Item numbers	NRCIM feature
Adult support and structure	1	Physical and psychological safety
	2, 3	Supportive relationships
	4, 5	Appropriate structure
Empowered skill-building	6	Positive social norms
	7, 8, 9	Skill-building opportunities
	10, 11	Support for efficacy and mattering
	12	Integration of family, school and community efforts
Expanding horizons	13	Positive social norms
	14	Opportunities to belong
	15, 16	Integration of family, school and community efforts
Negative experiences*	17	Physical and psychological safety
	18	Positive social norms
	19	Opportunities to belong

Note. *Items under the subscale negative experiences are negatively worded.

TLI = .895, SRMR = .0711, RMSEA = .042 (90% CI = .036 – .049), χ^2/df = 1.696, and χ^2 = 495.216. The third step in the analysis of invariance, which restrained variance of factor loadings between the two samples, showed the following goodness of fit indices: CFI = .907, TLI = .897, SRMR = .0765, RMSEA = .042 (90% CI = .035 – .048), χ^2/df = 1.680, and χ^2 = 522.413. From these analyses, it appears that the factorial structure of the YPQS is well represented by the 19-item, 4-factor model and is invariant (i.e. equivalent) across age groups for youth participants.

4. Discussion

The current study was conducted to explore the factor structure of the previously established YPQS by Silliman and Schumm (2013). Three confirmatory factor analyses were conducted and did not confirm previously identified models of the YPQS. As such, an exploratory factor analysis was conducted outlining a four-factor, 19 item measure that showed good model fit, resulting in a revised form of the YPQS. McCune (1989) cautioned that factor analysis studies of the same instrument could yield different results depending on the samples used, therefore it was not surprising that the exploratory factor analysis conducted within this study revealed slightly different factors than results identified by Silliman and Schumm (2013). However, the exploratory factor analysis conducted in this study provides good support for the model as there was a great deal of consistency between the current work and Silliman and Schumm's (2013) work. The results revealed the elimination of some of the same items that were also removed by Silliman and Schumm (2013), highlighting consistently problematic items. From this, the current study was able to build on the previous work conducted by these authors. Lastly, the proposed 4-factor 19-item resulted in being invariant across age groups for youth participants within this sample, outlining the appropriateness of using this tool as a measure of perceived program quality for youth between the ages of 10 and 18.

Eccles and Gootman (2002) have argued “the boundaries between features are often quite blurred” (p. 88), indicating an overlap between program setting features does occur within youth programming. Findings from previous research that utilized the eight program setting features as a guiding framework in measurement development, have outlined a similar notion, proposing other categorizations or subscales that encompass these features (e.g. H/SERF, 2005; Silliman & Schumm, 2013). For example, the Youth Program Quality Assessment, an observational assessment tool that can be used internally by program leaders and administrators or externally by researchers to examine quality within youth programming, has outlined four domains to assess (Safe Environment, Supportive Environment, Interaction, and Engagement; H/SERF, 2005). Based on previous work

examining program quality, it can be argued that although the YPQS and Youth Program Quality Assessment differ slightly from one another, there is alignment as they follow a similar structure whereby physical and psychological safety and program structure are seen as essential and foundational elements, and once satisfied, higher ordered elements of program quality can then be facilitated within program (Akiva, 2005; Eccles & Gootman, 2002).

Within this study, Negative Experiences fell under the acceptable reliability coefficient level of .70, yet Cronbach's alpha is often expected to be lower when there are fewer items, as in the current study as this subscale is comprised of three items. Moreover, this subscale measures a wide range of constructs, including perceptions of negative feelings pertaining to belonging, conflicts, and embarrassment, which may also contribute to the low reliability of this subscale. However, it is important to recognize that this study, as well as previous research on program quality (e.g. Silliman & Schumm, 2013), have integrated a subscale that measures youth perceptions of negative experiences within programs. Although this has not been identified as a program setting feature by Eccles and Gootman (2002), it is important to take into consideration the potential for negative experiences within youth program contexts. The inclusion of a subscale that measures negative experiences is common among other measures of youth development (e.g. YES 2.0 by Hansen & Larson, 2005; YES for Sport by Sullivan, LaForge-Mackenzie, & Marini, 2015; Summer Program Experiences Survey by Vandell, Hall, O'Cadiz, & Karsch, 2012).

The current results extended previous work in several ways. The sample used within this study would arguably be identified as a more heterogeneous sample than what was used in previous measurement testing with the YPQS. Previous research conducted with the YPQS included a relatively homogenous sample of youth from 4-H conferences and camps. In contrast, this study recruited youth from 38 different programs across multiple program contexts; therefore, allowing from a more heterogeneous sample that more accurately reflects the general context of youth programming, which can include a variety of different characteristics (e.g. program length, program context, age and gender of participants, type of organization). The current study is the first to provide support that the factor structure of the YPQS holds true across programs of differing lengths, as previous samples used to test the YPQS ranged from 3 to 7 days, whereas this sample recruited participants between 6 weeks and 9 months in length.

The factor structures previously proposed by Silliman and Schumm (2013) were based on particular age ranges which may not be in fact ideal. Previously identified program quality characteristics, and more specifically the eight identified setting features proposed by the NRCIM, have provided such recommendations that encompass program quality for all youth programs not by age, outlining similar program features within both school-age and adolescent programs (Vandell et al., 2005). Researchers that have proposed best practices for program quality have consistently identified global features as opposed to features-based age. Although the structure of the program may be adapted based on the age of participants (e.g. more emphasis on the structure for younger youth, yet more focus on expanding horizons for older youth), all elements are still critical to facilitate a high-quality program. Findings from this research supported the need for one measurement tool with a strong factor structure that takes into account all eight setting features of program quality proposed by the NRCIM.

5. Recommendations

Although Silliman and Schumm (2013) identified two different measures based on the classification of youth aged 10 and 17 by age and classified both measures as the YPQS, our analyses took into consideration one large sample of youth between 10 to 18 years. The World Health Organization (1989) defines 'youth' as individuals between 15 and 24 years of age, which is not consistent with the samples utilized within previous work with Silliman (2008c), Silliman and Shutt (2010), Silliman and Schumm (2013) or within the current study. Therefore, it would be suggested to change the name of the measure from the YPQS to the Adolescent Program Quality Survey (APQS) to fall in line with the World Health Organization's (1989) definition of 'adolescence', categorizing individuals

between 10 and 19 years. Moreover, it would be recommended for future testing and use of the new proposed APQS to utilize a five-point Likert scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*) instead of the four-point scale that has been previously used (Silliman, 2008c; Silliman & Schumm, 2013; Silliman & Shutt, 2010) to not only guard against kurtosis and ensure more normal distributions, but also allow for the neutral response and has been recommended for use with youth (Intelligent Measurement, 2007; Kline, 2010; Malhotra, 2006). These recommendations are suggested to be taken into consideration with the measure's items outlined in Table 2.

6. Limitations and future directions

The new proposed APQS has the potential to become a valuable tool for assessing program quality across a variety of youth programming contexts. It is important note that the factor structure presented in this study was determined through an exploratory process using data acquired from a single sample of 391 youth involved in various programming contexts. However, youth within this sample were from several different programming contexts (e.g. sport, leadership, in-school mentoring), yet sport programming was prominently represented in this sample. Future research should continue to test the new proposed APQS using confirmatory techniques with new samples of varying types of program contexts before claims of structural or external validity can be concretely made. Participants from this study were comprised of a diverse sample of youth; however, it is suggested that future research continues to validate the new proposed APQS with a variety of samples, including various youth program types (e.g. sport, leadership, arts), contexts (e.g. competitive vs. non-competitive, team/group vs. individual, mixed-gender vs. single gender), and organization types (not-for-profit, for-profit), as well as programs that target youth of various ages, ethnicities, and from differing socioeconomic backgrounds. Moreover, it is suggested to integrate a measure of youth engagement (e.g. Snapshot Survey of Engagement tool; Busseri, Rose-Krasnor, & Centre of Excellence for Youth Engagement, 2009) or a measure of youth psychosocial outcomes (e.g. YES 2.0; Hansen & Larson, 2005; PYDI; Arnold et al., 2012) as dependent variables to understand if indicators of quality are related to program engagement and/or outcomes.

At this time, the YPQS should be considered as an assessment tool for program quality specifically related to Appropriate Adult Support and Structure, Empowered Skill-building, Expanding Horizons, and Negative Experiences, which integrates the eight setting features proposed by the NRCIM (Eccles & Gootman, 2002). The current study proposes one factor model for all youth between the ages of 10 and 18, while Silliman and Schumm proposed two factor structures based on youth between the ages of 10 to 13 and youth ages 14 to 17. As no previous research or quantitative process or outcome measures within youth programming has created separate measures based on age, it can be argued that utilizing a program quality measure for that can be delivered across all youth programming contexts regardless of age (Vandell et al., 2005). However, future research is needed to further examine this. It should be noted that pilot work is underway that is continuing to examine the APQS within another large and diverse sample of youth involved in various youth programs. From an applied perspective, it is recommended that the proposed APQS may be used in conjunction with other measures (e.g. Youth Program Quality Assessment) to further triangulate program quality within youth programs. Lastly, it is possible that there are additional program quality elements that are not presented in the modified version of the APQS as research in this field is continually advancing. Although the eight setting features from NRCIM are represented, future research is warranted to examine if items need to be added to the APQS to broaden the scope of program quality assessment.

7. Conclusion

In conclusion, the new proposed APQS offers good psychometric properties and retains the conceptualization of program quality outlined by NRCIM's eight program setting features (Eccles & Gootman, 2002). As such, even with the deletion of items with lower factor loadings, the four-factor structure provides support for the eight program setting features. At this time, it also appears that APQS offers a potentially viable and psychometrically sound measure of program quality for youth between the

ages of 10 to 18 that are involved in a variety of youth programs with varying contexts and characteristics. Future research should continue to test the psychometric properties of the new proposed APQS based on the above recommendations and to ensure the NRCIM eight setting features are comprehensively encompassed within the measure.

Funding

Support for the writing of this manuscript was given through a doctoral research grant from the Social Sciences and Humanities Research Council of Canada.

Competing Interests

The authors declare no competing interest.

Author details

Corliss Bean¹

E-mail: cbean@uottawa.ca

Tanya Forneris¹

E-mail: tanya.forneris@uottawa.ca

ORCID ID: <http://orcid.org/0000-0002-8262-5412>

¹ School of Human Kinetics, University of Ottawa, 125 University Private, Ottawa, Ontario, Canada K1N 6N5.

Citation information

Cite this article as: Re-examining the youth program quality survey as a tool to assess quality within youth programming, Corliss Bean & Tanya Forneris, *Cogent Psychology* (2016), 3: 1149265.

Corrigendum

This article was originally published with errors. This version has been corrected. Please see Corrigendum (<http://dx.doi.org/10.1080/23311908.2016.1164500>).

References

- Akiva, T. (2005). Turning training into results: The new youth program quality assessment. *High/Scope Resource*, 24, 21–24.
- Arbuckle, J. (2014). *AMOS 23 reference guide*. Meadville, PA: AMOS Development Corporation.
- Arnold, M. E., Nott, B. D., & Meinhold, J. L. (2012). *The positive youth development inventory*. Corvallis, OR: Oregon State University 4-H Youth Development Program.
- Awang, Z. (2012). *Structural equation modeling using AMOS graphic*. Shah Alam: UiTM Press.
- Barber, B. L., Eccles, J. S., & Stone, M. R. (2001). Whatever happened to the jock, the brain, and the princess?: Young adult pathways linked to adolescent activity involvement and social identity. *Journal of Adolescent Research*, 16, 429–455.
<http://dx.doi.org/10.1177/0743558401165002>
- Benson, P. L., Scales, P. C., & Syvertsen, A. K. (2011). The contribution of the developmental assets framework to positive youth development theory and practice. In R. M. Lerner, J. V. Lerner & J. B. Benson (Eds.), *Advances in child development and behavior: Positive youth development research and applications for promoting thriving in adolescence* (pp. 195–228). London: Elsevier.
- Boys and Girls Clubs of Canada. (2008). Transforming out-of-school hours to achieve better health, educational attainment and lifelong productivity for young Canadians. Retrieved October 3, 2015, from: https://www.bgccan.com/en/AboutUs/PublicPolicy/Documents/bgccsubmission_finance_aug08_final.pdf
- Busseri, M. A., Rose-Krasnor, L., & Centre of Excellence for Youth Engagement. (2009). The engagement landscape: A snapshot of your activity. Retrieved September 28, 2015, from: http://www.excellenceforchildandadolescent.ca/sites/default/files/meas_attach/Snapshot_Youth_Engagement_Survey_-_The_Engagement_Landscape.pdf
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge.
- Camino, L., & Zeldin, S. (2002). From periphery to center: Pathways for youth civic engagement in the day-to-day life of communities. *Applied Developmental Science*, 6, 213–220.
http://dx.doi.org/10.1207/S1532480XADS0604_8
- Catalano, R. F., Hawkins, J. D., Berglund, M. L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: competitive or cooperative frameworks? *Journal of Adolescent Health*, 31, 230–239.
[http://dx.doi.org/10.1016/S1054-139X\(02\)00496-2](http://dx.doi.org/10.1016/S1054-139X(02)00496-2)
- Costello, A., & Osborne, J. (2011). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10, 1–9.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
<http://dx.doi.org/10.1007/BF02310555>
- Durlak, J. A., Mahoney, J., Bohnert, A., & Parente, M. (2010). Developing and improving after-school programs to enhance youth's personal growth and adjustment: A special issue of ajcp. *American Journal of Community Psychology*, 45, 285–293.
<http://dx.doi.org/10.1007/s10464-010-9298-9>
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, 45, 294–309.
<http://dx.doi.org/10.1007/s10464-010-9300-6>
- Eccles, J., & Gootman, J. A. (Eds.). (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.
- Employment and Social Development Canada. (2014). Leisure—Active leisure time. Retrieved June 19, 2014, from: <http://www4.hrsdc.gc.ca/3ndic.1t4r@-eng.jsp?id=51>
- Everitt, B. (1975). Multivariate Analysis: The need for data, and other problems. *The British Journal of Psychiatry*, 126, 237–240. <http://dx.doi.org/10.1192/bjp.126.3.237>
- Garson, G. D. (2010). *Structural equation modeling*. Retrieved November 18, 2015, from: <http://faculty.chass.ncsu.edu/garson/PA765/structur.htm>
- George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference*. 11.0 update (4th ed.). Boston, MA: Allyn & Bacon.
- Guèvremont, A., Findlay, L., & Kohen, D. (2008). Organized extracurricular activities of Canadian children and youth. *Health Reports*, 19, 65–69.
- Guèvremont, A., Findlay, L., & Kohen, D. (2014). Organized extracurricular activities: Are in-school and out-of-school activities associated with different outcomes for Canadian youth? *Journal of School Health*, 84, 317–325.
<http://dx.doi.org/10.1111/josh.2014.84.issue-5>
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Hamilton, S. F., Hamilton, M. A., & Pittman, K. (2004). Principles for youth development. In S. F. Hamilton & M. A. Hamilton (Eds.), *The youth development handbook: Coming of age in American communities* (pp. 3–22). Thousand Oaks: Sage Publications.
<http://dx.doi.org/10.4135/9781452232560>
- Hansen, D. M., & Larson, R. W. (2005). *The youth experience survey 2.0: Instrument revisions and validity testing*. Unpublished manuscript.

- Retrieved from [http://web.aces.uiuc.edu/youthdev/UnpublishedManuscriptonYES2111%20\(2\).doc](http://web.aces.uiuc.edu/youthdev/UnpublishedManuscriptonYES2111%20(2).doc)
- High/Scope Educational Research Foundation (2005). *Youth program quality assessment*. Ypsilanti, MI: High/Scope Press.
- Holt, N. L., & Jones, M. I. (2008). Future directions for positive youth development and sport research. In N. L. Holt (Ed.), *Positive youth development through sport* (pp. 122–132). New York, NY: Routledge.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424–453.
- Intelligent Measurement. (2007). *Likert scale & surveys—Best practices*. Retrieved October 20, 2015, from: <http://intelligentmeasurement.net/2007/11/20/likert-scale-surveys-best-practices/>
- Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: The Guilford Press.
- Lerner, R. M. (2004). *Liberty: Thriving and civic engagement among American youth*. Thousand Oaks, CA: Sage Publications.
- Lerner, R. M. (2005). *Promoting positive youth development: Theoretical and empirical bases*. Paper presented at the workshop on the science of adolescent health and development, Washington, DC.
- Lerner, R. M. (2009). The positive youth development perspective: Theoretical and empirical bases of a strength-based approach to adolescent development. In C. R. Snyder & S. J. Lopez (Eds.), *Oxford handbook of positive psychology* (2nd ed., pp. 149–163). Oxford: Oxford University Press.
- Mahoney, J. L., Larson, R. W., Eccles, J. S., & Lord, H. (2005). Organized activities as developmental contexts for children and adolescents. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after-school and community programs* (pp. 3–22). Mahwah, NJ: Erlbaum.
- Malhotra, N. K. (2006). Questionnaire design and scale development. In R. Grover & M. Vriens (Eds.), *The handbook of marketing research: Uses, misuses, and future advances* (pp. 176–202). Thousand Oaks, CA: Sage Publications.
- Martelaer, K., & Theeboom, M. (2006). *Handbook of Physical Education*. Thousand Oaks, CA: Sage Publications. <http://dx.doi.org/10.4135/9781848680809>
- McCune, S. K. (1989). Reactions to field's investigation into the SDLRS: A statistical critique of field's investigation. *Adult Education Quarterly*, 39, 243–245. <http://dx.doi.org/10.1177/0001848189039004007>
- McLaughlin, M. (2000). *Community counts: How youth organizations matter for youth development*. Washington, DC: Public Education Network.
- Merry, S. (2000). *Beyond home and school: The role of primary supports in youth development*. Chicago, IL: Chapin Hall Center for Children.
- National Collaboration for Youth. (2011). The impact of youth development programs on student academic achievement. Retrieved from <http://nationalassembly.org/Uploads2/Resources/SchoolSuccessBrief.pdf>
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fort Worth, TX: Harcourt Brace College.
- Powers, J. L., & Tiffany, J. S. (2006). Engaging youth in participatory research and evaluation. *Journal of Public Health Management and Practice*, 12, S79–S87. <http://dx.doi.org/10.1097/00124784-200611001-00015>
- Quinn, J. (1999). Where need meets opportunity: Youth development programs for early teens. *The Future of Children*, 9, 96–116. <http://dx.doi.org/10.2307/1602709>
- Roth, J. L., & Brooks-Gunn, J. (2003). What is a youth development program? Identification and defining principles. In F. Jacobs, D. Wertlieb, & R. M. Lerner (Eds.), *Enhancing the life chances of youth and families: Public service systems and public policy perspectives: Handbook of applied developmental science: Promoting positive child, adolescent, and family development through research, policies, and programs*, vol. 2 (pp. 197–223). Thousand Oaks, CA: Sage Publications.
- Silliman, B. (2008a). 4-H junior camp at Millstone, 2008—Tradition and innovation. North Carolina State University. Report retrieved on June 21, 2015, from <http://www.nc4h.org/uploads/documents/professionals/greenlight/PDF/JuniorCampImpact2008.pdf>
- Silliman, B. (2008b). Taking congress home: Effects of NC 4-H congress on youth behavior and interaction. *Journal of Youth Development*, 3(Article 0803PA001).
- Silliman, B. (2008c). *Youth program climate survey*. Raleigh, NC: North Carolina Cooperative Extension Service.
- Silliman, B., & Schumm, W. R. (2013). Youth program quality survey: Youth assessment of program quality. *Marriage & Family Review*, 49, 647–670.
- Silliman, B., & Shutt, R. E. (2010). Weaving evaluation into the fabric of youth development. *Journal of Youth Development*, 5 (Article 100503FA003).
- Sullivan, P. J., LaForge-Mackenzie, K., & Marini, M. (2015). Confirmatory factor analysis of the youth experiences survey for sport (YES-S). *Open Journal of Statistics*, 5, 421–429. doi:10.4236/ojs.2015.55044.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (5th ed., vol. 7). Upper Saddle River, NJ: Pearson.
- The Bridgespan Group. (2005). *Growth of youth-serving organizations: Appendices*. Retrieved September 12, 2015, from: <http://www.bridgespan.org/Publications-and-Tools/Youth-Development/Growth-of-Youth-Serving-Organizations/Growth-of-Youth-Serving-Organizations-Appendices.aspx#VhZ-dvlVikp>
- United States Census Bureau. (2014). Nearly 6 out of 10 children participate in extracurricular activities, Census Bureau reports. Retrieved from: <http://www.census.gov/newsroom/press-releases/2014/cb14-224.html>
- Urban, J. B. (2008). Components and characteristics of youth development programs: The voices of youth-serving policymakers, practitioners, researchers, and adolescents. *Applied Developmental Science*, 12, 128–139. <http://dx.doi.org/10.1080/10888690802199400>
- Vandell, D. L., & Pierce, K. (2006). *Study of after-school care: Program quality observation*. Retrieved November 16, 2015, from: www.wcer.wisc.edu/childcare/pdf/asc/program_quality_observation_manual.pdf.
- Vandell, D. L., Shumow, L., & Posner, J. (2005). After-school programs for low-income children: Differences in program quality. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after-school and community programs* (pp. 437–456). Mahwah, NJ: Lawrence Erlbaum Associates.
- Vandell, D. L., Hall, V., O'Cadiz, P., & Karsch, A. (2012). *Power of discovery: STEM 2 Year 1 (2012–2013)*. Report findings, California: Department of Education.
- Vandell, D. L., Larson, R. W., Mahoney, J. L., & Watts, T. W. (2015). After-school programs for low-income children: Differences in program quality. In M. H. Bornstein, T. Leventhal, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Ecological settings and processes*, vol. 4 (pp. 305–344). Hoboken, NJ: John Wiley & Sons.
- Vandell, D. L., Shumow, L., & Posner, J. (2005). After-school programs for low-income children: Differences in program quality. In J. L. Mahoney, R. W. Larson, & J. S. Eccles

- (Eds.), *Organized activities as contexts of development: Extracurricular activities, after-school and community programs* (pp. 437–456). Mahwah, NJ: Lawrence Erlbaum Associates.
- Walker, K. (2007). Youth empowerment evaluation: Learning voice. *American Journal of Evaluation*, 28, 321–326. <http://dx.doi.org/10.1177/1098214007306377>
- World Health Organization. (1989). *The health of youth*. Document A42/Technical Discussions/2. Geneva: World Health Organization. Retrieved from: http://apps.who.int/iris/bitstream/10665/172241/1/WHO42_TD-1_eng.pdf
- 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. <http://dx.doi.org/10.1007/s10464-010-9311-3>
- Yuan, Y. C. (2010). *Multiple imputation for missing data: Concepts and new development (Version 9.0)*. Rockville, MD: SAS Institute.



© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Psychology (ISSN: 2331-1908) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

