Journal of Human Sciences and Extension

Manuscript 1349

Youth Sparks via 4-H: Relations with Program Quality, Dosage, and Involvement

Alexandra Skrocki

Gary Ellis

Andrea Ettekal

Darlene Locke

Youth Sparks via 4-H: Relations with Program Quality, Dosage, and Involvement

Alexandra Skrocki
Gary Ellis
Andrea Ettekal
Darlene Locke
Texas A&M University

Using a statewide sample of participants in 4-H programs, we tested associations among youth sparks, program quality, program dosage, and four types of involvement (breadth of participation in out-of-school-time beyond 4-H, scope of learning strategies, degree of specialization, and cumulative program immersion). Participants were 180 Texas 4-H alumni who graduated from high school in 2013 and 2014 and had two or more years of involvement in Texas 4-H programs. Data were collected via an electronic questionnaire based on a database of former Texas 4-H members. Twelve hypotheses were tested, linking program quality to sparks through indirect and direct relations. Eleven of the twelve hypotheses were supported. Program quality had a strong direct relation to sparks (b = .41). Direct relations between sparks and the facets of immersion were also found (b = .31 for the degree of specialization in one of Texas 4-H's program areas and b = .27 for cumulative program immersion). Indirect associations were also significant. Dosage (years in Texas 4-H) was not significantly related to sparks.

Keywords: 4-H, sparks, program quality, dosage, involvement

Introduction

A primary goal of research and practice in youth work is to describe, explain, and optimize youth development (Lerner, 2012), thereby helping youth thrive. Many impactful ideas about optimizing youth development stem from the field of positive youth development (PYD), which conceptualizes thriving as a process occurring through the alignment of youth strengths and ecological assets (Lerner et al., 2015). Out-of-school-time (OST) programs are a major focus in PYD scholarship because of extensive empirical support indicating that OST programs are exemplary ecological assets to support youth thriving (Lerner et al., 2015; Vandell et al., 2015). However, empirical studies of program involvement primarily focus on tests of associations between involvement and developmental outcomes, such as academic success, career success, health, and quality of life (Roth & Brooks-Gunn, 2016). Few empirical studies examine how program involvement is associated with specific youth strengths that may facilitate those outcomes, such as academic self-efficacy, persistence, leadership skills, and citizenship skills.

In this study, we focus on a specific youth strength "sparks" (Benson & Scales, 2009). Sparks are youths' internally derived passions. Sparks motivate youth to actively seek opportunities to engage their self-identified interests and purposes. Although the concept of sparks has received considerable attention (e.g., Arnold, 2018; Arnold & Gagnon, 2019; Benson & Scales, 2011; Scales et al., 2011), few empirical studies have been directed at identifying developmental results of sparks and factors that may facilitate youths' discovery and development of sparks. Sparks have been found to be positively associated with developmental outcomes (e.g., Ben-Eliyah et al., 2014; Scales et al., 2011). Correlational research also suggests that involvement in high-quality OST programs may be a determinant of sparks (e.g., Lewis et al., 2021).

Yet, involvement in OST programs embraces many potential vectors of influence in addition to program quality. Program dosage (Roth et al., 2010), degree of engagement (Lerner et al., 2005), and specialization vs. breadth of participation (Linver et al., 2009) also seem to be influential. Specialization vs. breadth of participation has been a topic of keen interest among recreation and leisure studies scholars. Bryan (2000) refers to specialization as the placement of recreation participants "on a continuum from general interest and low involvement to specialized interest and high involvement. Each level of specialization carries distinctive behaviors and orientations" (p. 18). Both sparks and specialization involve the continued pursuit of skill development in an activity; thus, the two concepts are closely aligned. Yet, sparks must be ignited; interests are discovered through experiences. Many programs and caregivers strive to expose youth to a myriad of experiences to help them discover their sparks. Both specialization and breadth of participation may contribute significantly to youth development. Research on the relationship between sparks and these types of involvement is limited. We thus examined associations between sparks, program quality, and types of involvement in a statewide OST program of youth experience opportunities.

Positive Youth Development: Individuals and Their Contexts

PYD is a strengths-based approach, in contrast to a deficit-reduction approach, to support youth thriving (Lerner et al., 2015). Most PYD models derive from the metatheory of relational developmental systems, which explains development as involving exchanges between individuals and their contexts, termed individual(context relations) (Overton & Lerner, 2014). Positive development occurs when individual⇔context relations are mutually beneficial, such that the individual acts to benefit the context and the context supports the individual (Lerner et al., 2015). Such mutually beneficial individual⇔context relations are termed adaptive developmental regulations (Lerner et al., 2015). Youth thriving is the process and outcome of adaptive developmental regulations and can be operationalized to include myriad youth strengths or positive features of contexts.

Thriving, Self-System, and Youth Sparks

Thriving is an active process involving individual agency (Benson & Scales, 2009). Individual agency is a primary function of the self (e.g., Csikszentmihalyi, 1993; Harter, 2012). Thus, attributes of the self are critical to the thriving process. The self-system serves three related but distinct roles in youth development (Jacobs et al., 2003). First, the self-system is a filter through which information is processed, such that youth make sense of their worlds and organize information based on self-representations (organizational function). Second, youth rely on aspects of the self-system, such as self-images, to make important decisions and plan future actions and initiatives (goal function). Third, self-related beliefs and values (e.g., self-concepts) motivate future behavior (motivation function). In brief, youth drive the thriving process through an engine fueled by their self-system. Specific attributes of the self likely align with the contextual assets provided in youth programs.

Benson and Scales (2009) theorized that specific aspects of the self, namely self-identified passions, are central to the process of youth thriving. These passions are termed sparks in Benson's and Scales' conceptualization of thriving. More specifically, sparks are the "passion for a self-identified interest, skill, or capacity that metaphorically lights a fire in an adolescent's life, providing energy, joy, purpose, and direction" (Scales et al., 2011, p. 264). Empirical research supports the importance of sparks. Youth who had a spark (i.e., were able to identify a core passion) fared better than youth who did not have a spark on a set of more than 15 developmental outcomes (Benson & Scales, 2009). The links between sparks and developmental outcomes are reflected in other empirical studies (e.g., Arnold, 2018; Ben-Eliyahu et al., 2014; Lewis et al., 2021). Sparks appear to be influential in promoting thriving. A subsequent question, then, is how may youth sparks be fostered to promote youth thriving? We now turn our attention to the contextual assets involved in sparks and thriving.

Contextual Assets to Support Thriving: Youth Programs

A primary contextual asset featured in many PYD models is youth participation in voluntary OST programs (Lerner et al., 2015). Youth programs build skills, are led by adults, and engage youth in increasingly complex activities over time (Roth & Brooks-Gunn, 2016). Youth programs have features that make them effective in promoting thriving: the voluntary nature, interest-based content, and skill-building activities. Youth program participation is ordinarily non-obligatory; thus, program participants tend to have at least some level of intrinsic motivation to engage in the activities (Larson, 2000). Different types of programs are available to support a diversity of interests (e.g., sports, arts, agriculture, leadership development, and many others). Youth programs promote the development of skills by offering challenging activities requiring concentration (Roth & Brooks-Gunn, 2016). The confluence of free choice, interest, and challenge in high-quality youth programs produces an exemplary opportunity to support youth

thriving. Indeed, substantial evidence supports the conclusion that participating in youth programs is associated with increased thriving (Vandell et al., 2015).

Capturing Variation in Program Involvement

An expansive body of literature is devoted to understanding the role of youth programs for thriving (Vandell et al., 2015). Since its inception in the 1980s, research on youth programs has evolved in two important ways. The first wave of research on youth programs, conducted through the 1990s, involved categorizing youth as either participants or non-participants and comparing them to a set of developmental outcomes (e.g., Feldman & Matjasko, 2005). The second wave commenced around the turn of the century and involved refining understanding of what it meant to participate in a program. Research in the 2000s focused on quantifying "how much" participation is needed and qualifying "how good" the program was, with a general finding that "more is better" on quantity and quality indicators (e.g., Farb & Matjasko, 2012). The current iteration of youth program research involves further specification of the program dimensions involved in thriving (Lerner et al., 2021; Simpkins, 2015).

Quantity of Programs. Participants vary in the extent of involvement within programs (Roth et al., 2010). Thus, youth who participate in programs receive different *dosages* of the program. Dosage is a term borrowed from the medical profession, representing the amount of a program received (Roth et al., 2010). Immediate involvement in programs is operationalized as intensity (amount of time spent per program meeting, such as 1–2 hours) and frequency (number of program meetings per a given time interval, such as one session per week), and long-term involvement is operationalized as duration (e.g., number of years involved). In the context of OST programs, meeting attendance could be considered a measure of program involvement. Another set of indicators, termed *engagement*, captures the extent to which youth become behaviorally and psychologically involved in the program when they attend. Engagement captures behavioral, emotional, agentic, and psychological dimensions of participation (e.g., Lerner et al., 2005; Reeve & Tseng, 2011). Increases in dosage on any indicators have been shown to be associated with increases in youth thriving (e.g., Roth et al., 2010). With this context in mind, we intend to measure levels of engagement in parallel with meeting attendance to create a measure of cumulative program immersion (CPI). To our knowledge, no studies have tested associations between indicators of dosage and sparks. However, if programs are indeed a context to support thriving, existing literature suggests positive associations between dosage and sparks.

Quality of Programs. Programs vary in quality, such that high-quality programs promote thriving, and low-quality programs can do more harm to youth than good (Yohalem & Wilson-Ahlstrom, 2010). The predominant conceptualization of program quality derives from a landmark publication by Eccles and Gootman (2002) summarizing a large research project supported by the National Academies of Sciences. In their review of the nearly two decades of research on youth programs, Eccles and Gootman identified eight features of programs that had

sufficient empirical support for their association with youth thriving, including safety, appropriate structure, supportive relationships, opportunities to belong, positive social norms, support for efficacy and mattering, skill-building activities, and integration across settings. Substantial empirical evidence supports the conclusion that program quality increases are associated with increased developmental outcomes (e.g., Durlak & Weissberg, 2012; Yohalem & Wilson-Ahlstrom, 2010). Preliminary empirical support indicates that specific features of program quality, such as relationships and general operationalizations of quality, are positively associated with youth sparks (Ben-Eliyahu et al., 2014; Lewis et al., 2021). The program quality features address many aspects of the self-system described earlier, such as support for efficacy (e.g., self-concept of ability) and mattering (e.g., self-worth). Thus, it is reasonable to expect that increases in program quality should be associated with increases in sparks.

Types of OST Programs. A wide range of OST youth program types exist, including those involving sports, arts, faith-based activities, leadership development, academic development, and nature-based experiences. Research suggests that the type of program matters. Different types of programs provide different kinds of experiences, which, in turn, promote different outcomes (Larson et al., 2006). Breadth is an indicator of program involvement that captures the number of different types of programs in which youth participate (Linver et al., 2009). Increased breadth has been found to be associated with increased thriving (Roth et al., 2010). Participating in multiple types of programs is beneficial because doing so provides youth the opportunity to practice skills in various settings, which promotes mastery as well as the transfer of skills from one setting to another (e.g., Pierce et al., 2017). It is unclear whether breadth should be positively associated with sparks. Sparks involve a depth of personal interest that may be better fostered through specialization than breadth. Thus, for exploratory purposes, we examine associations between breadth and sparks.

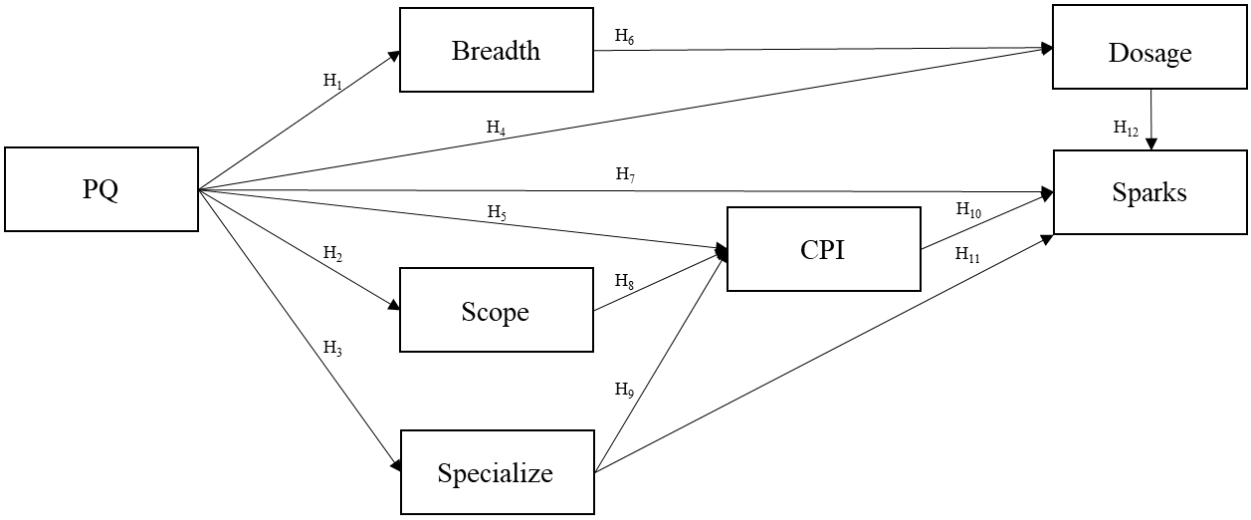
Aligning Individual and Contextual Assets to Support Sparks: The Present Study

We tested associations between sparks and types of involvement in youth programs. Our data are drawn from a statewide program, Texas 4-H, which offers a rich diversity of types of activities in which youth may selectively participate. A substantial empirical base exists for 4-H, suggesting that increased involvement in 4-H programs, operationalized to include indicators of program quantity, quality, and types, is associated with increases in a range of developmental outcomes (e.g., Bowers et al., 2015; Lerner et al., 2005, 2015, 2019, 2021; Geldhof et al., 2013). Few studies have examined associations between involvement in 4-H and youth sparks. Such research has linked assets (whether individual or contextual) with developmental outcomes. However, more research is needed that links individual and contextual assets together, as such alignment is the precondition for sparks and thriving.

Hypotheses

The path model in Figure 1 depicts the hypotheses we tested. They are presented as research hypotheses in Table 1. Sparks was hypothesized to be associated with program quality, dosage, and four types of involvement: breadth of participation in OST beyond 4-H, scope of learning strategies, degree of specialization in one of the Texas 4-H "Big 5" program areas, and cumulative program immersion (CPI). We reasoned that program quality promotes involvement and dosage (i.e., youth tend to become more involved if the programs are of high quality), and that program quality also has a direct effect on sparks. Both direct and indirect effects are likely, considering that elements of program quality (i.e., safety, structure, supportive relationships, belonging, positive social norms, support for efficacy and mattering, and opportunities for skill building) are consistent with fundamental psychological needs of youth (and adults) for autonomy, competence, and relatedness (Ryan & Deci, 2002).

Figure 1. Hypotheses



Note. PQ = Program Quality; Breadth – Breadth Other OST; Scope = Scope of Learning Strategies; Specialize = Specialization Big 5; CPI = Cumulative Program Immersion

Table 1. Research Hypotheses

Number	Research Hypotheses
H_1	As program quality increases, breadth in other OST increases.
H_2	As program quality increases, scope of learning strategies increases.
H_3	As program quality increases, specialization increases.
H_4	As program quality increases, dosage increases.
H_5	As program quality increases, cumulative program immersion increases.
H_6	As breadth in other OST increases, dosage increases.
H_7	As program quality increases, sparks increase.
H_8	As scope of learning strategies increases, cumulative program immersion increases.
H_9	As specialization increases, cumulative program immersion increases.

Number	Research Hypotheses
H_{10}	As cumulative program immersion increases, sparks increase.
H_{11}	As specialization increases, sparks increase.
H_{12}	As dosage increases, sparks increase.

Method

Participants

We analyzed existing data from a previous study that was directed at evaluating the academic and career success of young adults who had been involved in 4-H (Skrocki, 2021; Skrocki et al., 2023). Participants were 180 Texas 4-H alumni who graduated from high school in 2013 and 2014 and had two or more years of involvement in Texas 4-H. The sampling frame was 4HOnline, a database managed by Texas 4-H. That database includes enrollment history, demographic profile, and primary 4-H project for all Texas 4-H members. The sample included 67.8% females and 31.7% males. Among the respondents, 87.9% were White, non-Hispanic Latino, and 12.1% were Non-White or Hispanic Latino. The average age of participants was 25 (SD = .79).

Invitations to participate were sent electronically to the 1,636 alumni, and 180 provided usable questionnaires, a response rate of 11%. Missing data for some variables reduced the sample size for some analyses to 166 (response rate of 10%). Low response rates are common when 4HOnline is used for sampling. Culp et al. (2022), for example, reported a response rate of 1.17%. Our low response rate would be a catastrophic limitation if our intent were to estimate the magnitudes of population parameters from a representative sample. Not only is the response rate unacceptably low, but other characteristics of the sample suggest substantial departures from the population. The percentage of the sample reporting having received major 4-H scholarships, for example, is substantially greater in the sample than the corresponding ratio in the population (Skrocki, 2021). The purpose of our study, though, was not to estimate the magnitude of parameters of individual variables (e.g., What is the average level of sparks in the population of Texas 4-H alumni?). Rather, our goal was to test theory about relations among and between variables. In Privitera's (2017, p. 251) words, our goal was "theoretical generalization," not "empirical generalization." It is reasonable to assume that overrepresentation of alumni who were academically successful (i.e., scholarship recipients) would tend to constrain variance on measures of sparks, program quality, and involvement. As a result, coefficients in our study are likely conservative estimates of corresponding and unknown coefficients in the population of 4-H alumni.

Measures

Our questionnaire included measures of seven variables: sparks, program quality, dosage, breadth of involvement in other OST programs, scope of learning strategies, specialization in

Texas 4-H Big 5 program areas, and cumulative program immersion (CPI). A full copy of our questionnaire may be provided upon request from the lead author. Descriptions of these measures follow.

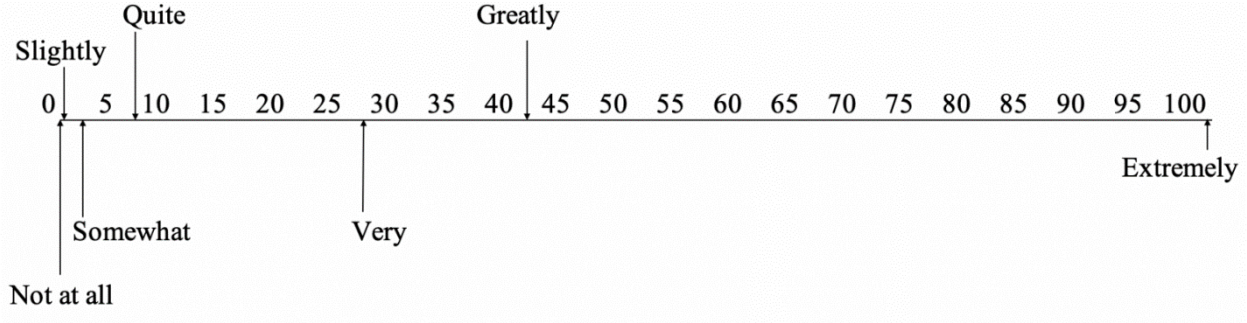
Sparks

Sparks is "a passion for a self-identified interest, skill, or capacity that metaphorically lights a fire in an adolescent's life, providing energy, joy, purpose, and direction" (Scales et al., 2011, p. 264). Six items were utilized to operationalize our measurement of sparks. Five items were modeled after Arnold and Gagnon (2018, 2019). We slightly adapted these items to fit our study's context. Arnold and Gagnon's (2018) item, "I am passionate about the things I do in this program," appeared as "In general, how passionate were you about the things you did in Texas 4-H?" in our questionnaire. We added a sixth item to represent the impact dimension of sparks, which is included in the sparks definition used by Arnold and Gagnon (2019). The item we added was "How impactful was Texas 4-H in helping you discover what you have to offer the world?" Content-related evidence of validity was enhanced through the addition of this item.

Arnold and Gagnon (2018) used a five-point response scale, ranging from not *true at all* to *very true*. To increase variation and reduce the probability of a halo effect (i.e., substantial negative skewness and constrained variance; Thorndike, 1920), we used a labeled magnitude scale (LMS; Schutz & Cardello, 2001). Participants responded along a 101-point continuum of intensity (0-100). The scale included seven adverbial modifiers placed at empirically identified intervals (Figure 2). The adverbial modifiers were *not at all*, *slightly*, *somewhat*, *quite*, *very*, *greatly*, and *extremely*. The modifiers were previously calibrated to their respective ratio-scaled distances in intensity through magnitude scaling procedures (e.g., Lodge, 1981; Meilgaard et al., 2016; Stevens, 1959). Ratings of intensity were provided by 158 judges who were university students and thus of similar ages as our research participants. Consistent with magnitude scaling procedures, two modalities were used for calibrating intensity judgments—line drawing and numeric estimation.

The composite reliability for our sparks measure was $\boldsymbol{\omega}$ = .96 and Average Variance Extracted (AVE) = .79, exceeding the standard of >.50 to indicate convergent validity within this data set (Collier, 2020). Heterotrait-Monotrait Ratio of Correlations (HTMT) provided evidence of discriminant validity between sparks and program quality (HTMT = .70) and sparks and cumulative program immersion (HTMT = .49). Both of these were below the HTMT standard, indicating discriminant validity for this data set (i.e., HTMT < .85).

Figure 2. Labeled Magnitude Scale (LMS)



Program Quality

Our measure of program quality (see Table 2) included six items representing the program quality indicators identified by Eccles and Gootman (2002) and Smith and Hohmann (2005). Arnold and Gagnon (2019) created a measure of program quality based on the program quality indicators identified by these scholars.

Our program quality scale used the same LMS response scale as our sparks measure. Our program quality scale yielded a composite reliability coefficient of $\omega = 0.94$. Arnold and Gagnon (2019) reported a composite reliability (ω) coefficient of .89. The AVE was .71, thus exceeding the standard of > .50 to indicate convergent validity within our data set (Collier, 2020).

Dosage: Years in Texas 4-H

Eligibility for participation in Texas 4-H begins upon entry into third grade and continues through twelfth grade. A single questionnaire item measured the years of membership in Texas 4-H. The scale of measurement ranged from zero to twelve. We included the 4-H Clover Kid program (kindergarten-2nd grade), thus giving us a 0–12 scale.

Breadth of Involvement in Other OST Programs

We measured breadth of participation in formal out-of-school time (OST) activities through a single questionnaire item asking participants to list all forms of formal OST involvement other than Texas 4-H:

> During your youth, were you involved in formal out-of-school time programs in addition to 4-H? Examples include Young Life, Scouts, sports teams, Boys and Girls Clubs of America, science camps, theater and arts camps, and others. If so, please list those other than 4-H in the space below.

Participants' qualitative responses were sorted into the six developmental opportunities as classified by Larson and colleagues (2006). Categories included (1) sports, (2) performance and

fine arts, (3) academic clubs and organizations, (4) community-oriented, (5) service, and (6) faith-based youth groups. Breadth scores ranged from 0 to 6, a count of additional developmental opportunities in which participants were involved.

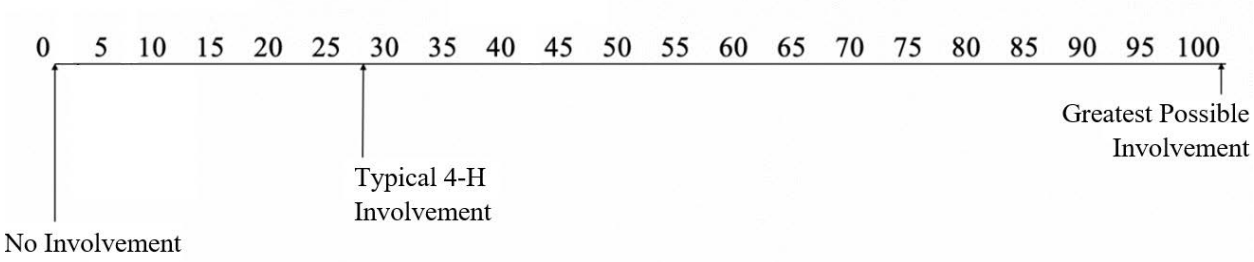
Scope of Learning Strategies

All Texas 4-H members designate one or more primary projects, examples of which are livestock, nutrition, computer science, shooting sports, and outdoor education. Six options are available as strategies for pursuing learning and development within these projects: (1) project meetings, (2) project workshops, (3) project-related tours, (4) virtual learning experiences, (5) educational demonstrations, and (6) career shadowing. We defined as a count of the number of strategies used. Thus, scores ranged from 0 to 6.

Specialization in Texas 4-H "Big 5" Program Areas

Primary projects within Texas 4-H are grouped into five overarching program areas known as the Big 5 and include (1) Agriculture and Livestock, (2) Family and Community Health, (3) Leadership and Citizenship, (4) Natural Resources, and (5) STEM. Participants were asked to indicate their depth of involvement in each of the Big 5 program areas. We used "direct magnitude scaling" (Lodge, 1981, p. 61) to measure specialization, yielding a ratio-level measure. To establish the reference point required for the direct magnitude scale, the 28th unit on the continuum indicated "typical 4-H involvement." It is notable that this measure may have produced imprecise data because different respondents likely had different perspectives on what typical 4-H involvement meant. Nonetheless, the measure served as a reasonably adequate reference point for involvement within our data set. Our measure of specialization was the Big 5 program area for which participants reported the highest score. An image of the direct magnitude scale response format is presented in Figure 3.

Figure 3. Direct Magnitude Scale



Cumulative Program Immersion

We measured CPI as the percentage of meetings attended and the level of engagement at each of two levels: club and primary project. For the attendance in meetings measure, participants were provided a 101-point slider scale ranging from 0% to 100% and were asked, "What percentage of

meetings do you attend…". We asked participants to use a slider scale to rate their overall degree of engagement ("Please indicate how engaged you typically are…"). The following statement was provided to assist participants in reporting their overall degree of engagement:

> *Engagement* is when you focus your attention on new information (for example, a presentation, a display, or a story) and...
>
> - you feel extremely interested, alert, and excited
> - your attention is fully focused on what is happening
> - you feel intense emotions
> - you are thinking about what might happen or solutions to a problem in the story

We used the same LMS response scale as the sparks and program quality measures (Figure 2). Four scores were thus obtained for our CPI measure: (1) attendance of club meetings, (2) attendance of primary project meetings, (3) engagement in club meeting experiences, and (4) engagement in primary project experiences. Factor analysis was used to create a CPI score. Factor loadings ranged from .47 to .92, and the single factor explained 62% of the total variance. The composite reliability coefficient was .80 and AVE = .52, exceeding the standard of >.50 to indicate convergent validity within this data set (Collier, 2020).

Procedure

Inclusion criteria were identical to that of the original study (Skrocki, 2021). An initial email was sent to Texas 4-H alumni meeting the study criteria, requesting their participation in the completion of an online questionnaire. The initial inquiry described the purpose of the study and notified alumni of incentives for completing the questionnaire. The incentive was a fifteen-dollar gift card distributed to the first one hundred participants. Following Dillman's survey methodology (Dillman et al., 2014), we sent additional emails once per week for three consecutive weeks. Our reminder emails included a shorter summary of the study's purpose, a report of the number of responses received, and the average questionnaire completion time.

Data Analysis

We cleaned and evaluated data using pairwise deletion. We removed cases for which box and whisker plots indicated outliers or *z*-scores were greater than 3.0. We calculated descriptive statistics using SPSS to summarize central tendency, shape, and dispersion of distributions of our variables and tested twelve hypotheses (see Figure 1) using path analysis (AMOS software version 28). Path analysis was originally a set of regression analyses among a nomological network of variables. The classic method of path analysis involves calculating separate least squares regression equations for each dependent variable in the network (Wright, 1934). Contemporary methods use maximum likelihood methods to calculate all coefficients simultaneously, resulting in more precise estimates and less accumulation of experimental error

across multiple tests. Path analysis implies a specific causal structure, but because the data are correlational and not experimental, it does not confirm that observed relations are causal. In the path analysis, full information maximum likelihood was used to account for missing data. Program quality was an exogenous variable, and the indicators of program involvement were presumed to have direct and indirect effects on sparks and dosage (years in Texas 4-H).

Results

Descriptive Statistics

Table 2 contains the descriptive statistics for distributions of scores on each measure. Among the variables measured on a scale of 0–100, the lowest means were "To what extent did you feel supported by other youth in Texas 4-H?" ($\bar{x}$ = 69.53) and "Rate your overall degree of engagement during community club experiences" ($\bar{x}$ = 69.39). The highest mean was "To what extent did you feel safe in Texas 4-H?" ($\bar{x}$ = 88.98). Another notable finding among the descriptive statistics is the mean of 1.52 for Breadth Other OST. The range of possible response options for that variable was 0–5. Thus, our sample of Texas 4-H participants tended to have limited involvement in other OST programs. The two distributions with the largest skewness and kurtosis values were "To what extent did you feel safe in Texas 4-H?" (skewness = -2.30, kurtosis = 4.53) and "What percentage of community club meetings were you able to attend?" (skewness = -2.44, kurtosis = 5.18). Different authorities provide different guidelines concerning what acceptable levels of skewness and kurtosis are (e.g., Collier, 2020; Kline, 2011; Tabachnick & Fidell, 2013). If we accept all standards proposed by these authors, we conclude no notable departure from normality exists in our data.

Table 2. Descriptive Statistics

	n	**λ**	**Min**	**Max**	**Mean**	**SD**	**Skew**	**Kur**
***PQ* (ω = .94, AVE = .71)**								
To what extent did you feel…								
welcome in Texas 4-H	166	.901	0	100	75.70	29.87	-1.23	.30
safe in Texas 4-H	166	.771	5	100	88.98	21.19	-2.30	4.53
supported by adults in…	166	.830	3	100	80.95	27.29	-1.52	1.15
supported by other youth in…	166	.857	0	100	69.53	30.39	-.85	-.49
Texas 4-H enforced rules	166	.825	0	100	73.60	29.77	-1.00	-.188
like you mattered in 4-H	166	.877	0	100	72.10	31.44	-.98	-.32
***CPI* (ω = .80, AVE = .52)**								
% of club meetings attended	170	.417	0	100	83.95	24.82	-2.44	5.18
Club meeting engagement	170	.854	3	100	69.39	30.55	-.80	-.66
% of proj meetings attended	169	.538	0	100	74.28	34.35	-1.3	.22
Project meeting engagement	167	.942	0	100	72.22	32.60	-.96	-.52

	n	λ	Min	Max	Mean	SD	Skew	Kur
Sparks* (ω = *0.96, AVE = .79)								
Passion for things you did	180	.906	1	100	74.90	28.52	-1.09	.03
Desire to learn all you could	180	.892	0	100	73.78	27.40	-1.09	.18
4-H important who you were	180	.848	0	100	70.41	29.87	-.910	-.36
Enthusiastic about projects	180	.938	1	100	77.00	27.22	-1.37	.92
Enthusiastic everything 4-H	180	.934	0	100	74.08	28.08	-1.18	.34
What you have to offer	180	.805	0	100	70.99	28.91	-.90	-.24
Breadth	180	--	0	5	1.52	1.14	.32	-.56
Specialize	171	--	12	100	79.07	25.02	-1.12	.03
Scope	180	--	0	6	4.12	1.98	-.73	-.78
Dosage	176	--	1	12	8.26	2.91	-.57	-.83

Note. PQ = Program Quality; CPI = Cumulative Program Immersion; Breadth = Breadth Other OST; Specialize = Specialization Big 5; Scope = Scope of Learning Strategies

Table 3 presents the correlations among variables in the path model. Notably, there was a strong positive correlation of .63 between program quality and sparks, indicating that sparks increase as program quality increases. Sparks were also found to have strong relations with specialization (r = .62) and cumulative program immersion (r = .62).

Table 3. Correlation Matrix

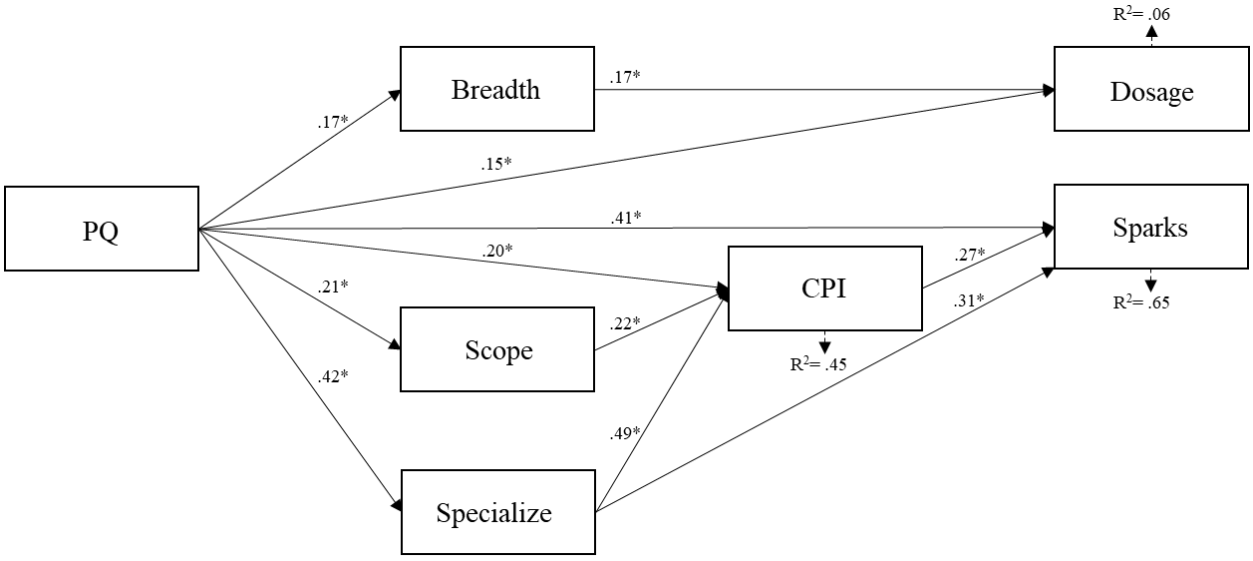
	PQ	Breadth	Scope	Specialize	CPI	Dosage	Sparks
PQ	1.00						
Breadth	.18*	1.00					
Scope	.18*	0.08	1.00				
Specialize	.42**	.16*	.11	1.00			
CPI	.45**	.18*	.29**	.60**	1.00		
Dosage	.17*	.20**	.16*	.27**	.17*	1.00	
Sparks	.63**	.19*	.25**	.62**	.62**	.13	1.00

Note. PQ = Program Quality; Breadth = Breadth Other OST; Scope = Scope of Learning Strategies; Specialize = Specialization Big 5; CPI = Cumulative Program Immersion
* $p < .05$; ** $p < .01$

Figure 4 shows significant coefficients in our path model, and Table 4 provides detailed results of the path analysis. Eleven of the twelve hypotheses were supported. The hypothesis not supported (H_{12}) proposed that sparks increase with increased dosage. No indirect pathways were found to be significant following additional analysis. The standardized coefficients associated with the significant hypothesis tests ranged from b = .15 (program quality and dosage) to b = .49 (specialization and cumulative program immersion). The multiple R^2 values were .65 (sparks), .45 (cumulative program immersion), and .06 (dosage). Program quality had the largest standardized total effect on sparks (.67). Specialization and cumulative program immersion also had notable standardized total effects on sparks, .44 and .27, respectively.

Results suggest that the application of PYD principles within some Texas 4-H programs may have cultivated youth sparks. This conclusion is supported by the large standardized total effect of program quality on sparks, .67. Our data also allowed us to address the question of whether sparks are more effectively facilitated by broad exposure to different types of programs (breadth) or by focused specialization in specific programs. The path from specialization to sparks was significant (b = .31). An exploratory analysis of a model that included a path from breadth to sparks yielded a nonsignificant path coefficient (b = .05, p = .30). Yet, from Table 1, we note that the zero-order correlation between breadth and sparks is significant (r = .19, p < .05). Thus, our results suggest that specialization may be somewhat more influential in promoting sparks than breadth of involvement in many types of programs.

Figure 4. Path Model, Standardized Path Coefficients (beta)



Note. PQ = Program Quality; Breadth – Breadth Other OST; Scope = Scope of Learning Strategies; Specialize = Specialization Big 5; CPI = Cumulative Program Immersion
* = significant at p < .001

Table 4. Path Analysis Results

Path	*beta*	*B*	*SE*	*t*	*p*
Scope ← PQ	.21	.02	.01	2.84	< .01
Specialize ← PQ	.42	.43	.07	6.01	< .01
Breadth ← PQ	.17	.01	.00	2.17	.03
CPI ← PQ	.20	.01	.00	3.08	< .01
CPI ← Scope	.22	.11	.03	3.79	< .01
CPI ← Specialize	.49	.02	.00	7.62	< .01
Dosage ← PQ	.15	.02	.01	2.03	.04
Dosage ← Breadth	.17	.43	.19	2.27	.02
Sparks ← Specialize	.31	.33	.06	5.30	< .01

Path	*beta*	*B*	*SE*	*t*	*p*
Sparks ← CPI	.27	7.70	1.68	4.57	< .01
Sparks ← PQ	.41	.44	.06	7.72	< .01
		Standardized Total Effects			
	PQ	Scope	Specialize	Breadth	CPI
Scope	.22	.00	.00	.00	.00
Specialize	.42	.00	.00	.00	.00
Breadth	.17	.00	.00	.00	.00
CPI	.45	.22	.49	.00	.00
Sparks	.67	.06	.44	.00	.27
Dosage	.18	.00	.00	.17	.00
		Standardized Direct Effects			
Scope	.22	.00	.00	.00	.00
Specialize	.42	.00	.00	.00	.00
Breadth	.17	.00	.00	.00	.00
CPI	.20	.22	.49	.00	.00
Sparks	.41	.00	.31	.00	.27
Dosage	.16	.00	.00	.17	.00
		Standardized Indirect Effects			
Scope	.00	.00	.00	.00	.00
Specialize	.00	.00	.00	.00	.00
Breadth	.00	.00	.00	.00	.00
CPI	.25	.00	.00	.00	.00
Sparks	.26	.06	.13	.00	.00
Dosage	.03	.00	.00	.00	.00

Note. Scope = Scope of Learning Strategies; PQ = Program Quality; Specialize = Specialization Big 5; Breadth = Breadth Other OST; CPI = Cumulative Program Immersion; the beta coefficients are calculated from standardized scores, and B coefficients are calculated from raw scores.

Discussion

We tested relations between sparks and six variables: (a) program quality, (b) breadth of involvement in other OST, (c) scope of learning strategies, (d) degree of specialization, (e) cumulative program immersion (CPI), and (f) dosage. All hypotheses were supported except for the relation between dosage and sparks. Our findings are thus largely consistent with theory and research on sparks, thriving, PYD, and the 4-H Thriving Model (Arnold, 2018; Arnold & Gagnon, 2019).

Our findings parallel results produced by Vallerand and colleagues (2003). They found, “Youth with deep interests [sparks] are likely to be more engaged and invigorated, not just in their sparks activities, but in their schools and communities, resulting in higher grades, attendance, and a mastery focus in school” (Vallerand et al., 2003, p. 78). While thriving is not explicitly

mentioned within the discussion of their results, in piecing together their findings with our own and aligning it with Arnold's 4-H Thriving Model, thriving is clearly occurring.

Results related to program quality also support previous research. The regression coefficient predicting sparks from program quality was greater than the other predictors. As noted in our review of relevant literature, varying levels of quality exist across OST programs (Yohalem & Wilson-Ahlstrom, 2010). Participants in our study used nearly the entire range of our 101-point scale to show levels of feeling welcome, supported by youth, rule enforcement, and mattering. For safety and support by adults, they used more than 95% and 97% of the response scale (see Table 2). This variation may reflect the varied backgrounds of 4-H youth leaders. As a community-based program, 4-H relies heavily on adult volunteers to manage 4-H clubs at the community level. In utilizing this community-based structure, 4-H members have the autonomy to determine the degree to which they are involved, thereby impacting the extent to which PYD principles are fostered. Programs planned, managed, and structured by professionals trained in PYD would likely yield greater sparks.

While there is no known hierarchal nature to the eight features noted within Eccles and Gootman's (2002) work on program quality, our results suggest the need to evaluate reasons youth used the entire range of response options (0–100) to indicate their perception of some dimensions of program quality. Some youth thus reported that they felt unwelcome, unsafe, unsupported, rules were not being followed, and that their presence did not matter. Future research is needed to identify and evaluate the effectiveness and efficiency of specific techniques leaders can use to promote youth feeling welcome, safe, supported, that rules were followed, and that they matter during meetings and events. These features and relations among them may impact the development of sparks. Perhaps, for example, a threshold of safety must be met or exceeded for mattering and belonging to facilitate sparks.

Our results align well with current work being conducted by the Search Institute. Their work suggests youth who are able to cultivate and maintain supportive relationships with their peers, family members, and community are more likely to develop resiliency and thrive (Search Institute, 2020). Core concepts included within Search Institute's developmental relationships framework (2020) are express care, challenge growth, provide support, share power, and expand possibilities. Unlike other studies that have addressed sparks as discovery in the present (Arnold & Gagnon, 2018; Ben-Eliyahu et al., 2014), our study design was retrospective. Yet, our findings also confirm the ability of 4-H members to discover and develop sparks during their time of involvement. One could connect the developmental relations associated with sparks to the continued development of supportive relationships and subsequent markers of thriving to our findings. Furthermore, our modest extension of the measure of sparks developed by Arnold and Gagnon (2019) shows strong psychometric properties. Reliability was strong ($\omega = 0.96$), and we found significant relations with theoretically related variables. The changing of response formats

from the five-point scale to our LMS and the addition of an item to build on content-related evidence of validity yielded a sound measurement tool for future research.

It is also important to discuss the relationship between dosage and sparks. As previously noted, dosage was not a significant predictor of sparks. Of course, we exercise caution in interpreting null effects. Null effects do not prove the relation does not exist. It is possible there is an effect and that our methods and statistical analysis failed to detect the effect. Nevertheless, we could speculate about the finding by comparing previous research on dosage. In the literature, dosage is operationalized in various ways, including intensity (e.g., hours/session), frequency (e.g., sessions/week), and/or duration (e.g., years) of participation, among others. Each of these indicators of dosage captures some aspect of the nature of behavioral participation (i.e., how much of a program did a youth attend?) Indicators of dosage have consistently explained more variation in developmental outcomes between program participants and non-participants than within program participants (Roth et al., 2010). The empirical literature on dosage within program participants is mixed, such that some studies find significant relations, others do not, and some report mixed findings depending on measurement and focal outcomes (e.g., Denault et al., 2009; Gardner et al., 2008; Mahoney & Vest, 2012). This is unsurprising as the literature strongly supports that indicators that capture participation beyond behavioral presence, such as psychological engagement, better explain what youth get from programs (Ramey et al., 2019). Moreover, previous research also supports that the quality of the program matters substantially (e.g., a small dosage of a high-quality program will yield better outcomes than a large dosage of a low-quality program; Yohalem & Wilson-Ahlstrom, 2010) and, thus, behavioral dosage in and of itself simply does not capture a full portrait of youth involvement in a program.

Limitations

Limitations exist. Perhaps most notable is recollection bias (Snelgrove & Havitz, 2010). As Snelgrove and Havitz (2010) point out, “One of the most common critiques of retrospective methods is that individuals cannot accurately recall past events, states of mind, or mental processes due to cognitive limitations” (p. 339). Greenhoot et al. (2011) discuss issues of retrospective methods in detail. Our use of existing data was thus inherently limited, and future research might be conducted with a sample of youth currently involved in 4-H. Another limitation is self-selection. Our questionnaire was distributed electronically to Texas 4-H alumni, who then chose whether they completed the questionnaire. This self-selection is a limitation, as is the related problem of the low response rate (Culp et al., 2022). Studies such as ours rely on 4HOnline to secure contact information for 4-H alumni, yet contact information is rarely updated after end of membership in 4-H (grade 12). Only alumni with internet access could receive a link to the questionnaire.

Generalization of outcomes across OST programs should be done carefully. To assist with such generalizations, Larson and colleagues (2006) have sought to inventory the various types of

"developmental experiences youth encounter when involved in different extracurricular and community-based organized activities" (Larson et al., 2006, p. 849). In their review, 4-H falls under the community-oriented category. Activities within this category are aimed at "connecting youth to community adults, institutions, and careers" (Larson et al., 2006, p. 850). In the case of our study, on average, our participants reported having been a member of Texas 4-H for eight years. This level of involvement in an OST program as specialized as Texas 4-H is far more likely to yield different PYD outcomes than involvement in other community-oriented activities and, most certainly, across other OST programs.

Future Research

While results from our study are directly related to Texas 4-H, the following recommendations for future research are applicable to all OST programs. Generalizing results from any correlational study must be done with appropriate caution. Formal theory development (Jaccard & Jacoby, 2020; Zetterberg, 1965), replicated results, longitudinal studies, and experimental designs in future studies are needed to confirm the significant relations we found and to ensure that those relations are causes and effects.

Our study should be replicated with the inclusion of measures pertaining to experience quality during programs, thriving, and developmental outcomes. Our findings confirm there are significant relationships between sparks and program quality and the measures of program involvement, yet the strength of relations between tested measures and thriving is still somewhat unknown. Questions about the strength of relations give rise to managers' concerns about efficiency and investments of limited resources. At what cost of investment in improving program quality (and other predictors), for example, can managers expect a unit increase in sparks? Similarly, collecting data from youth who are active members of Texas 4-H would provide valuable context regarding the discovery and development of sparks and program quality. Seeking to answer how immediate experiences should be structured to facilitate the discovery of sparks would also be an important contribution.

References

Arnold, M. E. (2018). From context to outcomes: A thriving model for 4-H youth development programs. *Journal of Human Sciences and Extension*, *6*(1), 141–160. https://doi.org/10.54718/NBNL5438

Arnold, M., & Gagnon, R. (2018). *The 4-H thriving model program evaluation instruments.* https://helping-youth-thrive.extension.org/wp-content/uploads/2021/04/Program-Evaluation-Instrument-Scoring-Guide.pdf

Arnold, M., & Gagnon, R. (2019). Illuminating the process of youth development: The mediating effect of thriving on youth development program outcomes. *Journal of Human Sciences and Extension*, *7*(3), 24–51. https://doi.org/10.54718/GHUP2927

Ben-Eliyahu, A., Rhodes, J., & Scales, P. (2014). The interest-driven pursuits of 15 year old's: "Sparks" and their association with caring relationships and developmental outcomes. *Applied Developmental Science*, *18*(2), 76–89. https://doi.org/10.1080/10888691.2014.894414

Benson, P. L., & Scales, P. C. (2009). The definition and preliminary measurement of thriving in adolescence. *The Journal of Positive Psychology*, *4*(1), 85–104. https://doi.org/10.1080/17439760802399240

Benson, P. L., & Scales, P. C. (2011). Thriving and sparks. In Leveque, R. J. R (Eds.), *Encyclopedia of adolescence* (pp. 2963–2976). https://doi.org/10.1007/978-1-4419-1695-2_199

Bowers, E. P., Geldhof, G. J., Johnson, S. K., Hilliard, L. J., Hershberg, R. M., Lerner, J. V., & Lerner, R. M. (Eds.). (2015). *Promoting positive youth development: Lessons from the 4-H study*. Springer.

Bryan, H. (2000). Recreation specialization revisited. *Journal of Leisure Research*, *32*(1), 18–21. https://doi.org/10.1080/00222216.2000.11949879

Collier, J. (2020). *Applied structural equation modeling using AMOS*. Routledge. https://doi.org/10.4324/9781003018414

Csikszentmihalyi, M. (1993). *The evolving self*. HarperCollins Publisher.

Culp, K., III, Averbeck, M., Ragland, K. (2022, April 25-28). *Kentucky 4-H impact study: Alumni and volunteers.* National Extension Conference on Volunteerism, Tulsa, OK, United States.

Denault, A. S., Poulin, F., & Pedersen, S. (2009). Intensity of participation in organized youth activities during the high school years: Longitudinal associations with adjustment. *Applied Development Science*, *13*(2), 74–87. https://doi.org/10.1080/10888690902801459

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.

Durlak, J. A., & Weissberg, R. P. (2012). *Afterschool programs that follow evidence-based practices to promote social and emotional development are effective*. Expanded Learning and Afterschool Project.

Eccles, J., & Gootman, J. A. (Eds.). (2002). *Community programs to promote youth development*. The National Academies Press. https://doi.org/10.17226/10022

Farb, A. F., & Matjasko, J. L. (2012). Recent advances in research on school-based extracurricular activities and adolescent development. *Developmental Review*, *32*(1), 1–48. https://doi.org/10.1016/j.dr.2011.10.001

Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*, *75*(2), 159–210. https://doi.org/10.3102/00346543075002159

Gardner, M., Roth, J., & Brooks-Gunn, J. (2008). Adolescents' participation in organized activities and developmental success 2 and 8 years after high school: Do sponsorship,

duration, and intensity matter?. *Developmental Psychology*, *44*(3), 814–830. https://psycnet.apa.org/doi/10.1037/0012-1649.44.3.814

Geldhof, G. J., Bowers, E. P., & Lerner, R. M. (2013). Special section introduction: Thriving in context: Findings from the 4-H study of positive youth development. *Journal of Youth and Adolescence*, *42*(1), 1–5. https://doi.org/10.1007/s10964-012-9855-7

Greenhoot, A. F., Laursen, B., Little, T., & Card, N. (2011). *Retrospective methods in developmental science*. In B. Laursen, T. Little, & N. Card (Eds.), *Handbook of developmental research methods* (pp. 196–210). The Guilford Press.

Harter, S. (2012). *The construction of the self* (2nd ed.). The Guilford Press.

Jaccard, J. & Jacoby, J. (2020). *Theory construction and model building skills* (2nd ed.). The Guilford Press.

Jacobs, J. E., Bleeker, M. M., & Constantino, M. J. (2003). The self-system during childhood and adolescence: Development, influences, and implications. *Journal of Psychotherapy Integration*, *13*(1), 33–65. https://doi.org/10.1037/1053-0479.13.1.33

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (5th ed.). The Guilford Press.

Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist*, *55*(1), 170–183. https://doi.org/10.1037/0003-066X.55.1.170

Larson, R. W., Hansen, D. M., & Moneta, G. (2006). Differing profiles of developmental experiences across types of organized youth activities. *Developmental Psychology*, *42*(5), 849–863. https://doi.org/10.1037/0012-1649.42.5.849

Lerner, R. M. (2012). Developmental science: Past, present, and future. *International Journal of Developmental Science*, *6*(1–2), 29–36. https://doi.org/10.3233/DEV-2012-12102

Lerner, R. M., Lerner, J. V., Almerigi, J., Theokas, C., Phelps, E., Gestsd´ottir, S., Naudeau, S., Jelicic, H., Alberts, A., Ma, L. Smith, L. M., Bobek, D. L., Richman-Raphael, D., Simpson, I., Christiansen, E. D., & von Eye, A. (2005). Positive youth development, participation in community youth development programs, and community contributions of fifth grade adolescents: Findings from the first wave of the 4-H study of positive youth development. *Journal of Early Adolescence*, *25*(1), 17–71. https://doi.org/10.1177/0272431604272461

Lerner, R. M., Lerner, J. V., Bowers, E., & Geldhof, G. J. (2015) *Positive youth development and relational developmental system*s. In W. F. Overton, P. C. Molenaar, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Theory and method* (7th ed.; pp. 607–651). Wiley.

Lerner, R. M., Lerner, J. V., Murry, V. M., Smith, E. P., Bowers, E. P., Geldhof, G. J., & Buckingham, M. H. (2021). Positive youth development in 2020: Theory, research, programs, and the promotion of social justice. *Journal of Research on Adolescence*, *31*(4), 1114–1134. https://doi.org/10.1111/jora.12609

Lerner, R. M., Tirrell, J. M., Dowling, E. M., Geldhof, G. J., Gestsdóttir, S., Lerner, J. V., King, P.E., Williams, K., Iraheta, G., & Sim, A. T. (2019). The end of the beginning: Evidence

and absences studying positive youth development in a global context. *Adolescent Research Review*, *4*(1), 1–14. https://doi.org/10.1007/s40894-018-0093-4

Lewis, K., Mun Kok, C., Worker, S., & Miner, G. (2021). Exploring the relationship between program experience and youth developmental outcomes. *Journal of Human Sciences and Extension*, *9*(3), 68–106. https://doi.org/10.54718/YKAF2288

Linver, M. R., Roth, J. L., & Brooks-Gunn, J. (2009). Patterns of adolescents' participation in organized activities: Are sports best when combined with other activities? *Developmental Psychology*, *45*(2), 354–367. https://doi.org/10.1037/a0014133

Lodge, M. (1981). *Magnitude scaling*. Sage. https://www.doi.org/10.4135/9781412984874

Mahoney, J. L., & Vest, A. E. (2012). The over-scheduling hypothesis revisited: Intensity of organized activity participation during adolescence and young adult outcomes. *Journal of Research on Adolescence*, *22*(3), 409–418. https://doi.org/10.1111/j.1532-7795.2012.00808.x

Meilgaard, M. C., Civile, G. V., & Carr, B. T. (2016). *Sensory evaluation techniques* (4th ed.). CRC Press. https://doi.org/10.1201/b16452

Overton, W. F., & Lerner, R. M. (2014). Fundamental concepts and methods in developmental science: A relational perspective. *Research in Human Development*, *11*(1), 63–73. https://doi.org/10.1080/15427609.2014.881086

Pierce, S., Gould, D., & Camire, M. (2017). Definition and model of life skills transfer. *International Review of Sport and Exercise Psychology*, *10*(1), 186–211. https://doi.org/10.1080/1750984X.2016.1199727

Privitera, G. (2017). *Research methods for the behavioral sciences* (2nd ed.). Sage.

Ramey, H. L., Lawford, H. L., & Rose-Krasnor, L. (2019). Psychological engagement and behavioral activity participation as predictors of positive youth development. *Journal of Youth Development*, *14*(3), 88–109. https://doi.org/10.5195/jyd.2019.769

Reeve, J., & Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, *36*(4), 257–267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Ryan, R., & Deci, E. (2002). 1: Overview of self-determination theory: An organismic dialectical perspective. In E. Deci & R. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). University of Rochester Press.

Roth, J. L., & Brooks-Gunn, J. (2016). Evaluating youth development programs: Progress and promise. *Applied Developmental Science*, *20*(3), 188–202. https://doi.org/10.1080%2F10888691.2015.1113879

Roth, J. L., Malone, L. M., & Brooks-Gunn, J. (2010). Does the amount of participation in afterschool programs relate to developmental outcomes? A review of the literature. *American Journal of Community Psychology*, *45*(3-4), 310–324. https://doi.org/10.1007/s10464-010-9303-3

Scales, P., Benson, P., & Roehlkepartain, E. (2011). Adolescent thriving: The role of sparks, relationships, and empowerment. *Journal of Youth and Adolescence*, *40*, 263–277. https://doi.org/10.1007/s10964-010-9578-6

Schutz, H., & Cardello, A. (2001). A labeled affective magnitude (LAM) scale for assessing food liking/disliking. *Journal of Sensory Studies*, *16*(2), 117–159. http://doi.org/10.1111/j.1745-459X.2001.tb00293.x

Search Institute. (2020). *Developmental relationships*. https://www.search-institute.org/developmental-relationships/developmental-relationships-framework/

Simpkins, S. D. (2015). When and how does participating in an organized after-school activity matter? [Editorial]. *Applied Developmental Science*, *19*(3), 121–126. https://doi.org/10.1080/10888691.2015.1056344

Skrocki, A. C. (2021). *Academic motivation and success of texas high school graduates who were and were not involved in Texas 4-H* [Master's thesis, Texas A&M University]. https://hdl.handle.net/1969.1/196077

Skrocki, A., Locke, D., Ellis, G., & Williams, M. (2023). Academic performance of Texas 4-H alumni. *Journal of Human Sciences and Extension*, *11*(2), Article 14. https://doi.org/10.55533/2325-5226.1383

Smith, C., & Hohmann, C. (2005). *Full findings from the youth PQA validation study*. High/Scope Educational Research Foundation. https://forumfyi.org/wp-content/uploads/2019/05/YPQA-Full-Findings-Val-Study-05.pdf

Snelgrove, R., & Havitz, M.E. (2010). Looking back in time: The pitfalls and potential of retrospective methods in leisure studies, *Leisure Sciences*, *32*(4), 337–351. https://doi.org/10.1080/01490400.2010.488199

Stevens, S. S. (1959). Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *Journal of Experimental Psychology*, *57*(4), 201–209. https://doi.org/10.1037/h0048957

Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, *4*(1), 25–29. https://doi.org/10.1037/h0071663

Vallerand, R. J., Blanchard, C., Mageau, G. A., Koestner, R., Ratelle, C., Leonard, M., Gagne, M., & Marsolais, J. (2003). Les passions de l'a ̂me: On obsessive and harmonious passion. *Journal of Personality and Social Psychology*, *85*(4), 756–767. https://doi.org/10.1037/0022-3514.85.4.756

Vandell, D. L., Larson, R. W., Mahoney, J. L., & Watts, T. W. (2015). Children's organized activities. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (7th ed., pp. 1–40). Wiley.

Wright, S. (1934). The method of path coefficients. *Annuals of Mathematical Statistics*, *5*(3), 161–215. https://doi.org/10.1214/aoms/1177732676

Yohalem, N., & Wilson-Ahlstrom, A. (2010). Inside the black box: Assessing and improving quality in youth programs. *American Journal of Community Psychology*, *45*(3-4), 350–357. https://doi.org/10.1007/s10464-010-9311-3

Zetterberg, H. (1965). *On theory and verification in sociology* (3rd ed.). The Bedminster Press.

Alexandra Skrocki is a Ph.D. student in the Parks, Recreation and Tourism Management Department at Clemson University. Her applied research interests include positive youth development and evaluation of out-of-school time programs. Please direct correspondence about this article to Alexandra Skrocki at askrock@g.clemson.edu.

Gary Ellis is a professor and Bradberry chair for the Agricultural Leadership, Education, and Communications Department at Texas A&M. He has an extensive track record of publications, presentations, and keynotes addressing immediate experiences for youth, tourists, and other recreationists.

Andrea Ettekal is an associate professor for Applied Youth Development within the Agricultural Leadership, Education, and Communications Department at Texas A&M. Her research investigates positive youth development through out-of-school time programs.

Darlene Locke is an associate professor and extension specialist (youth development) with Texas A&M AgriLife Extension. She has served with the Texas 4-H Youth Development program for more than 30 years, with 11 of those as director of the Texas 4-H Conference Center.