

Construct-Validity of the Engagement with Challenge Measure for Adolescents: Structural- and Criterion-Validity Evidence

E. Whitney G. Moore¹, David Hansen²

¹Health, Sport & Exercise Sciences Department, University of Kansas, Lawrence, USA

²Psychology & Research in Education Department, University of Kansas, Lawrence, USA

Email: ewgmoore@ku.edu

Received July 18th, 2012; revised August 19th, 2012; accepted September 12th, 2012

For adolescents, *engaging with challenge* is a key developmental task, hypothesized to support development of adult-like competencies (e.g., agency and self-direction; Larson, 2000). This study aimed to assess the construct-validity (structural- and concurrent-validity) of a new self-report measure assessing adolescents' engagement with challenge to help researchers understand how different settings and the conditions in these settings support adolescents' development. The sample consisted of 337 adolescents in 10 FFA programs along with the adult advisors in each program. Adolescents completed a questionnaire, which included the Engagement with Challenge measure and the following criterion variables: number of contests completed, participation frequency, and leadership roles. In addition to the self-reported criterion variables, the adult advisor evaluated Engagement with Challenge for each FFA student member in that program using a single item. The findings of this study provided strong evidence for the structural-validity of the engagement with challenge construct measured by the new scale, including having passed confirmatory factor analysis configural, weak, and strong invariance tests across four grade groupings. The findings also provided further evidence of criterion-validity, as Engagement with Challenge correlated in the *a priori* hypothesized direction and magnitude. Suggestions for analysis with the new measure and for future research are presented.

Keywords: Adolescent; Engagement; Challenge; Cognitive Development; Out-of-School Time

Introduction

There is a dearth of instruments assessing key adolescent learned competencies related to self-initiated or self-governed thought and behavior. As a form of intrinsic motivation (Csikszentmihalyi, 1990; Larson & Rusk, in press; Ryan & Deci, 2000), learning to become engaged with challenges is hypothesized to be a key indicator of progress toward becoming an agentic adult (Larson, 2000)—one who has developed a propensity to enjoy and become motivated by and absorbed in challenges that emerge while working on goals or projects over time. Learning to engage with challenge has also been hypothesized by some scholars to promote enjoyment and thriving in adulthood (Csikszentmihalyi & Rathunde, 1993). Relatively recent neuroscience findings on brain develop suggests adolescents may be primed to learn certain competencies, such as engagement with challenge, although its development is not a spontaneous or guaranteed outcome of adolescence. Engagement with challenge, we suggest, is an increasingly valuable commodity in contemporary economies, as suggested, for example, by the Secretary's Commission on Achieving Necessary Skills (SCANS; 1991) report. Thus, having a valid measure of engagement with challenge is an important step that can support research into how different experiences and settings of adolescents' lives support the development of this competency, as well as others. The aim of this research, then, is to evaluate the construct-validity of a new self-report measure of adoles-

cents' engagement with challenge, focusing on evidence for the structural-validity and criterion-validity of the measure.

Theoretical Foundation for Engagement with Challenge

The theoretical roots for the engagement with challenge construct stem from the literature on intrinsic motivation (Csikszentmihalyi, 1990; Larson, 2000; Ryan & Deci, 2000). At its most basal level, humans possess an incentivizing system (e.g., Steinberg et al., 2006) for self-initiated behavior and thought to attend to novel stimuli and explore their environment (Richards, Reynolds, & Courage, 2010). When individuals experience enjoyment or pleasure (reward) as a result of novelty seeking or exploration they are more likely to be motivated (incentivized) to repeat similar behaviors or thoughts. Theoretically, novelty seeking and exploration has considerable adaptive utility when it is linked to enjoyment, as it can propel learning and contribute to a sense of personal "control" of one's environment (e.g., ability to predict and cause behaviors and events that lead to goal attainment; Csikszentmihalyi, 1990; Hidi, 2006; Larson & Rusk, in press; Panksepp, 1998; Ryan & Deci, 2009). Ryan and Deci (2000) define intrinsic motivation as an "inherent tendency to seek out novelty and challenges" (pg. 70). Although intrinsic motivation is couched in terminology suggesting it is an innate or "hardwired" predisposition, we suggest it is more consistent with recent neuroscience findings to view intrinsic

motivation as developing from an interaction of the brain's incentivizing (e.g., positive emotion, enjoyment) and cognitive regulating (e.g., future planning, goals, challenges) systems that simultaneously encode experiences, which results over time in enduring patterns of thought, emotion, and behavior consistent with someone we describe as intrinsically motivated. Conceptualized this way, intrinsic motivation serves the evolutionary function of promoting the widest range of learning for adapting, surviving, and thriving within an environment that is in flux.

The engagement with challenge construct represents a particular form of intrinsic motivation, which Csikszentmihalyi (1990) and Larson (2000) posit is a commonly shared characteristic among those who develop skill and expertise in a particular domain, such as music or sport. Research suggests that individuals who become self-motivated by challenges, compared to those who are motivated by more extrinsic reasons, tend to sustain and expend effort on tasks longer when faced with difficulties, select more challenging tasks, learn from experience more proficiently, and find more creative solutions to problems (Deci & Ryan, 2000; Larson & Rusk, in press). Additionally, scholars have suggested that learning to become self-motivated by challenges can serve as a source of enduring enjoyment, satisfaction, and thriving throughout adulthood (Csikszentmihalyi & Rathunde, 1993; Ryan, 1995). Beyond the personal benefits, contemporary work and society highly value and reward individuals who demonstrate competency for engaging with challenge, as it can lead to innovation, novel solutions to problems, etc. (e.g., Secretary's Commission on Achieving Necessary Skills, 1991; SCANS). We speculate that this learned competency was less integral to successful functioning in past industrialized economies that relied on a workforce with discrete skill sets that applied to a fixed set of tasks and problems, but that today's "high-performance" economies has created a greater demand for individuals with competencies for engaging with challenge. As espoused here then, engagement with challenge is a learned competency consisting of persistent linkages between enjoyment and challenges that arise in pursuit of a goal or outcome; it is evidenced as an inclination or propensity to enjoy and become motivated by and absorbed in tasks' challenges rather than motivated primarily by the extrinsic rewards of performing a task (e.g., praise, public recognition).

Adolescence and Engagement with Challenge

While the functional importance of becoming skilled at engaging with challenge in society may vary over time and culture, relatively recent findings from the cognitive and neuroscience literature suggest adolescents may be "neuro-developmentally primed" to learn such a competency. With the onset of puberty, adolescents' brains begin to undergo a major reorganization or restructuring, which ultimately leads to the full integration of emotion-motivation (e.g., incentivizing system) and cognitive regulation (e.g., executive control) systems (Keating, 2004; Luna et al., 2001; Spear, 2000; Steinberg et al., 2006). The emotion-motivation system experiences a period of heightened arousal, especially during early adolescence, resulting in greater emotional and motivational awareness, arousal, and "excitatory tone" (Spear, 2000; Steinberg et al., 2006). The cognitive regulation system itself (e.g., the PFC), which is linked to higher-order thinking skills—such as developing strategies for accomplishing complex work or managing one's

time and effort—experiences a pruning of extraneous neurons and increased myelination of axons (Griedd, 2004; Keating, 2004; Spear, 2000). (Myelination supports greater efficiency and speed of synaptic firing and is an indication of learning that has occurred over time). Perhaps most poignant for becoming accomplished at engaging with challenge, the adolescent brain develops new, more comprehensive axonal pathways between the emotion-motivation and higher-order cognitive regulation systems (Goldman-Rakic, 1988; Luna et al., 2001). These myelinated pathways represent an emerging functional integration of two systems—each system able to contribute its own specialized learned functions and skills—that results in greater self-governance over thought, motivation, behavior, and learning. We suggest, then, that adolescents' brains are primed to learn complex competencies, including engagement with challenge, which reflect the integration among these two systems.

Becoming skilled at and garnering enjoyment from engaging with challenges, however, should not be viewed as a spontaneous outcome of contemporary adolescence. Such a process is not an automatic developmental outcome of "normal" adolescence. Experience is an essential, yet surprisingly overlooked, influencer of brain development and functioning (Greenough, Black, & Wallace, 1987; Markham & Greenough, 2004). In laboratory settings, for example, stimulating or enriching environments induce morphological changes in the brains of animals (e.g., greater density of neural connections), which are associated with greater learning or brain plasticity (Markham & Greenough, 2004). While a direct association between adolescent brain development and experiences that support learning to engage with challenge—or other positive competencies (e.g., strategic planning/thinking) for that matter—have not yet been established, research on adolescents' experiences in structured, out-of-school youth programs suggest this is a setting that supports development.

A Setting That Supports Learning to Engage with Challenges

One early study identified organized, out-of-school youth programs as a setting in which adolescents reported experiencing both high levels of challenge and high levels of enjoyment (Csikszentmihalyi & Larson, 1984). Subsequent research has consistently pointed to these youth programs as a prominent setting in adolescents' lives—compared to other settings such as class or hanging out with friends—in which they report higher rates of a wide range of developmental experiences, including challenge and enjoyment (Hansen, Larson, Dworkin, 2003; Hansen & Larson, 2007; Larson, Hansen, & Moneta, 2006). Qualitative research has also found that, over a course of time between three to eight months, depending on the particular youth program, adolescents describe becoming motivated to take on greater challenges that emerged while working on projects, particularly projects that produced "real-world," tangible products, such as creating murals for public display (Larson & Angus, 2011; Larson, Hansen, & Walker, 2005; Larson & Walker, 2006).

There is no single activity or program that would support every adolescent's learning to engage with challenge, however. Adolescents' interests in activities, for example, change and diversify over time (Hofer, 2011; Krapp, 1999). Thus, activities or programs that may support learning to engage with challenge for one adolescent may not be same for another. Similarly,

engaging with challenge implies the tendency to enjoy and to be motivated by increasingly more difficult challenges. Given Western adolescents have considerable discretion over participation in activities and programs—some for example simultaneously participate in multiple programs (Larson, Hansen, & Moneta, 2006)—they can select activities or programs that fit with their current desire to experience a particular level of challenge. Some programs offer progressively more difficult or skilled activities, which could support an individual adolescent's increasing desire for greater challenge. Research, then, suggests that organized youth programs tend to offer opportunities for adolescents to experience engagement with challenge, and thus provide a reasonable setting in which to evaluate a new quantitative measure of engagement with challenge.

Methods

Data come from a larger study of rural adolescents' development in FFA.¹ The FFA is an agriculturally-based, national youth program that offers high school students opportunities to participate in different competitions (e.g., electrical wiring), hold official elected offices (e.g., Chapter president), work with peers and adults on projects, socially interact with adults and other youth, as well as design and implement long-term agriculture-related projects (e.g., Supervised Agricultural Experience). Throughout the year, students have opportunities to plan and implement events/projects (e.g., community service projects), set personal and group goals, and receive feedback on their work. Participation in FFA activities is voluntary, and adolescents are free to participate as much or as little as they choose.

The FFA was selected for this study because it provides an extensive array of activities that can appeal to a broad range of adolescent interests and because activities vary in the degree of training and challenge needed, which we reasoned could provide progressive opportunities to engage with challenge for a wide adolescent population. FFA Chapters are located in high schools and have at least one adult "advisor", typically the agricultural teacher. FFA Chapter activities occur after school hours, on weekends, and often during the summer months. Because the FFA curriculum and activity options are consistent across Chapters, the focus only on the FFA program helped reduce one potential source of systematic variance in the measured variables that could result from differences in the types of youth programs (e.g., sports and arts).

Procedures

Convenience sampling, with purposive selection of FFA Chapters, was used to select 10 out of a potential of 58 FFA Chapters within a three hour driving distance from the university in which the co-author was employed at the time.² The selection procedures were based on criteria on Chapters' performance on the "program standards" and "quality indicators" collected by the overarching regulatory FFA body, Facilitating Coordination in Agricultural Education (FCAE). All FFA Chapters report on these indicators to the FCAE. We chose 28 indicators that reflected the degree a particular Chapter fol-

lowed the aims and curriculum of the FFA, and reflected the degree to which a particular Chapter was engaged and active in FFA: eight indicators of classroom instruction, four of the supervised agricultural experience (SAE), and 16 of the FFA. Collectively we refer to these indicators as Chapter "fidelity".

A score for each chapter was computed (range 0 - 28) based upon the number of standards and indicators a Chapter met, one point for each indicator (1 = "met the standard"). Based on the computed scores, the 58 chapters were divided into three groups: Chapters with 20 or more points, 15 - 19 points, and less than 15 points. Chapters scoring less than 15 points were excluded from selection because we reasoned that they would not provide sufficient enough opportunities for adolescents to experience engagement with challenge, and because the purpose of the larger study was to examine the impact of students' active participation (e.g., greater dosage) in the program. Finally, only those Chapters with (a) at least one agriculture instructor who had been teaching for three years or more in the current school and (b) had a Chapter membership of least 40 members were included. These two final criteria ensured advisors had sufficient experience leading an FFA Chapter and that there were sufficient student participants for statistical analysis. After applying the criteria, 16 chapters were eligible for selection: eight had 20 or more points and eight had 15 - 19 points. We then randomly selected five chapters from each of the two groups of eight for this study.

Data were collected from both the student FFA members and the adult advisors of each Chapter; questionnaires were given to students and advisors during the spring of 2006 (Time 1) and 2007 (Time 2). Two weeks prior to the data collection, an informational flyer was sent to parents of FFA members of a given Chapter along with instructions for declining a child's consent if a parent so desired. Adolescent assent was obtained on the day of the questionnaire administration and FFA advisors provided their consent. While this study focuses solely on evaluating the construct-validity of the new Engagement with Challenge measure at one point in time, data from both collection times were used to best inform the multiple imputation process (Little, in press; Wu, Lang, & Little, 2009). Evaluating the measure longitudinally in this study, however, was deemed premature and was reserved for a future study. In addition, simultaneously including longitudinal analysis in this study would expand its scope and length and address a fundamentally different research question, the longitudinal invariance of the Engagement with Challenge construct, which requires cross-sectional structural-validity. Thus, this analysis of the construct-validity of the measure, focusing on structural- and criterion-validity evidence, analyzed only data from Time 1.

Participants

Students

The sample consisted of 337 high school students (67% male) from 10 FFA Chapters. The mean age of participants was 16.21 (range = 14 to 19). There were 121 freshman, 60 sophomores, 54 juniors, and 102 seniors. The majority of students, 60.7%, lived in a rural area ("on a farm in the country" or "not on a farm but in the country"), 37.7% reported living "in a small town or city (less than 10,000 people)," and 1.6% of students reported living in a "medium size city (between 10,000 and

¹FFA stands for Future Farmers of America but the national FFA organization no longer include the full title, preferring simply the FFA.

²Lead author was employed at the University of Illinois at the time the larger study was conducted.

200,000 people).” No students reported living in a “large city (more than 200,000 people).” The sample was 85.8% White, 2.7% Native American, 0.3% Hispanic, 6.0% reported multiple ethnicities or “other”, and 5.2% chose not to respond. Students self-reported this demographic information.

Adult Advisors

Eleven adult advisors participated in this study; one school had two advisors. Within each FFA Chapter, the adult advisor rated each youth’s engagement with challenge. In the one Chapter with two advisors, each advisor rated half of the youth. Ten of the 11 advisors were male and all were White. The average number of years these advisors had been teaching was 28.82, with a range from 5 to 37.

Measures

Engagement with Challenge

The engagement with challenge scale was developed from the qualitative research of Larson and colleagues (Larson & Hansen, 2005; Larson, Hansen, & Walker, 2005; Pearce & Larson, 2006). In that research, organized youth programs and their adolescent participants and adult leaders were studied over a naturally occurring cycle of the program (e.g., academic year). Observational, interview, and survey data were collected. From the emerging grounded-theory of that qualitative work, items for the engagement with challenge scale were generated. The aim was to create a concise measure that not only captured the theoretical, higher-order linkages between intrinsic motivation and challenge but also to use adolescents’ language and expressions that reflected the construct; for this latter aim, items drew directly on youths’ words from the interview data when possible. Nine items for the scale were originally generated from the qualitative work. In a different study with a group of 20 adolescents attending a conference at a university, we asked participants to complete the scale and comment on any items that seemed awkward or unclear. Once completed, a researcher met with participants in groups of 4 - 5 students to elicit their feedback about the items. Based on the students’ feedback, three of the nine items were deemed to be confusing or vague and thus were dropped. As a result of this vetting process, the final scale comprised six items, three that were positively worded and three that were negatively worded.

The engagement with challenge scale is a self-report measure of a form of intrinsic motivation, assessing the adolescents’ enjoyment of the challenges that occur while working on a goal or project. Students were given the following instructions, “Please read each statement and circle the number that is most correct about your participation in FFA in the last 12 months”. The six items in the scale are: 1) There are always things I’m trying to work on and achieve in this program; 2) I feel challenged in a good way in this program; 3) The activities in this program are boring; 4) I’m not working toward anything in this program; 5) What we do in this program is both difficult and enjoyable; 6) The goals people are working on in this program are not important to me. Items 3, 4, and 6 are reverse coded. Students’ indicated agreement with each statement on a 6-point scale: 1 = “Strongly Disagree”, 2 = “Disagree”, 3 = “Slightly Disagree”, 4 = “Slightly Agree”, 5 = “Agree”, 6 = “Strongly Agree”.

Criterion Variables

Student-Reported Lead Role

The FFA offers multiple opportunities for students to serve in a leading capacity at the Chapter, state, and national level (e.g., Chapter president, district representative). Holding a lead role represents a substantial investment of time and effort, as well as a commitment to the FFA culture. As such, holding a lead role should also correlate with student’s level of engagement with challenge. Students were asked to indicate, “yes” (1) or “no” (0), if they had a lead role in the FFA Chapter. Holding a lead role was hypothesized to be moderately and positively correlated with scores on engagement with challenge.

Student-Reported Frequency of Participation

Conceptually, a students’ level of engagement with challenge in the FFA Chapter should be reflected in the amount of time they spend in the activities. One indicator of time is the frequency of students’ participation in the FFA Chapter. Frequency of participation was assessed with one item: “How often did you participate in FFA in the last 12 months?” Response categories were, 1 = “Less than once a month”, 2 = “Once a month”, 3 = “A few times a month”, 4 = “A couple of days a week”, and 5 = “Almost every day”. More frequent participation in the activities of the FFA Chapter was expected to moderately correlate with greater engagement with challenge.

Student-Reported Participation in Competitions

The FFA offers competitions in which students demonstrate skills and knowledge, such as electrical wiring, at the local, regional, state, and national levels. Because of the additional time commitment required to participate in FFA contests, it was expected to positively and moderately correlate with engagement with challenge. Students reported on each FFA contest in which they participated (1 = participated, 0 = did not participate). The total number of contests was computed by summing across each student’s reported contest participation to make a single sum variable.

Advisor Ratings of Each Student’s Engagement with Challenge

Advisors reported their perception of each student’s engagement with challenge compared to other same-age peers. Each advisor rated every student in the chapter using a single item. Given the number of FFA members in most Chapters, it was not reasonable to ask the advisor to rate every student member in a manner equivalent to the student-report measure with six items. The advisor was given the following definition: “Engagement with challenge refers to how motivated and engaged a student has been in challenging activities of FFA. For engagement with challenge, rate how each youth compares to other youth their age using the scale below.” Each student’s name was listed by the scale and the advisor ranked each student by deciles from 1 = “0% - 10%”, the lowest decile, to 10 = “91% - 100%”, the highest decile. A student ranked in the 50th percentile, for example, would indicate that the advisor considered the student to demonstrate an average level of engagement with challenge compared to other students his or her age. This response scale was used to provide the advisor with as objective reference points as possible and to increase the likelihood of approximating a normal distribution with a single item. While not an equivalent measure to students’ reports on the engage-

ment with challenge measure, advisors' ratings were expected to have a moderately positive correlation with students' reports.

Results

All analyses were completed within a structural equation modeling (SEM) framework. Preliminary analyses were first conducted in order to analyze the data's tenability for meeting parametric analysis assumptions and to compute descriptive statistics. Confirmatory factor analysis (CFA) for engagement with challenge followed the preliminary analysis to evaluate the structural-validity (Clark & Watson, 1995; Loevinger, 1957) of the scale—testing the general fit of the hypothesized construct's measurement model and if the instrument measured the same construct across grade groups. SEM analyses were then conducted to examine the concurrent-validity of the engagement with challenge scale by assessing the construct's correlations with the criterion variables.

Preliminary Analysis

The preliminary data evaluation indicated that the data had 1% missingness and that the highest fraction missing for any individual item was .053. Given this high recoverability of the missing data, modern missing data techniques were employed (Little, 2010). Specifically, multiple imputation was conducted using Amelia (R Development Core Team, 2010). Multiple imputation was selected in order to retain the maximum sample N (e.g., no listwise deletion of cases), while producing more accurate and stable estimates of standard errors—that are neither inflated nor left out of the estimation process (Little, 2010)—even compared to the full information maximum likelihood method (Olinsky, Chen, & Harlow, 2003). One set of sufficient statistics were then calculated across the 100 imputed datasets to use in all subsequent SEM analyses (Little, 2010), which were completed using Mplus, Version 5.21 (Muthen & Muthen, 2009). **Table 1** displays the descriptive statistics used in the SEM analyses. Stronger positive correlations existed between the three positively worded engagement with chal-

lenge items than between these three items and the negatively worded engagement with challenge items, and vice versa (See **Table 2**).

Measurement Invariance (CFA)

The purpose of this important step—establishing measurement invariance—is to confirm that the latent construct's measurable manifestations (e.g., scores on the items) were effectively measuring the same construct across the four grade groups (i.e., 9th, 10th, 11th, and 12th grade). Thus, before SEM analyses can be conducted to examine differences in variances, covariances or means, the measurement model must be established to ensure subsequent comparisons across groups are accurately based on comparing the same construct of interest. As the final model attained from the CFA is then used for the SEM analysis, the criterion variables were included throughout the CFA process. However, as the criterion variables were all single-indicators, their measurement invariance was not tested; their inclusion was solely in preparation for evaluating the criterion validity of the engagement with challenge scale after the CFA was complete. Therefore, the emphasis and results of the invariance testing is on the findings for the engagement with challenge construct across the four grade groupings.

Configural Model

The initial CFA step was to establish the appropriate measurement or configural model. The configural model is the hypothesized model for the measure of interest (i.e., engagement with challenge). If an initially hypothesized configural model does not have adequate model fit, then either alternatively hypothesized models or post-hoc adjustments that align with theory should be made. The configural model included the three positively worded and the three negatively worded engagement with challenge items (e.g., items 3, 4, and 6) after being reverse coded to positively load on the engagement with challenge construct. The measurement model for a single latent construct demonstrated (with all items coded in the same direction) inadequate fit (RMSEA = 0.128, 90%CI: .111 - .146; TLI = .859;

Table 1.
Means and standard deviations by grade.

¹ Engagement with Challenge (EWC) Items	9th Grade		10th Grade		11th Grade		12th Grade	
	M	SD	M	SD	M	SD	M	SD
Item 1	4.17	1.24	3.90	1.34	3.94	1.33	3.86	1.30
Item 2	4.32	1.34	4.06	1.36	4.08	1.24	4.30	1.17
Item 3	4.51	1.54	4.54	1.41	4.43	1.42	4.75	1.31
Item 4	4.82	1.37	4.66	1.46	4.55	1.40	4.50	1.36
Item 5	4.34	1.31	4.22	1.31	4.39	1.10	4.41	1.13
Item 6	4.79	1.35	4.66	1.35	4.61	1.34	4.71	1.18
Frequency of Participation	2.65	1.30	2.82	1.15	2.82	1.23	3.04	1.32
² Advisor Rating of Student's EWC	5.44	2.48	5.66	2.49	6.20	2.43	7.13*	2.30
Lead Role	5.05	2.37	5.36	2.42	5.88	2.43	6.98*	2.33
Contest Number	1.55	2.29	1.97	2.67	1.93	2.21	2.75*	2.78

Note: ¹Response on a 1 - 6 scale; ²Response on a 1 - 10 scale; *The 12th grade latent means were significantly different than the 9th, 10th, and 11th grade latent means.

Table 2.
Bi-variate correlations among study variables across all grades (N = 337).

	1	2	3	4	5	6	7	8	9	10	11
1) Grade	--										
2) Eng1 (EwC)	-.16**	--									
3) Eng2 (EwC)	-.05**	-.62**	--								
4) Eng3+	-.06**	-.29**	-.39**	--							
5) Eng4+	-.10**	-.46**	-.42**	.55**	--						
6) Eng5 (EwC)	-.01**	-.47**	-.61**	-.46**	-.41**	--					
7) Eng6+	-.02**	-.34**	-.31**	.51**	-.53**	-.41**	--				
8) Frequency	-.09**	-.42**	-.27**	-.14**	-.30**	-.20**	-.17**	--			
9) Advisor	-.19**	-.24**	-.22**	-.21**	-.23**	-.29**	-.18**	.26**	--		
10) Contests	-.07**	-.22**	-.12**	-.08**	-.19**	-.18**	-.03**	.26**	.31**	--	
11) Leadership	-.38**	-.26**	-.17**	-.15**	-.20**	-.19**	-.16**	.39**	.42**	.31**	--

Note: +Identifies the eventual non-engaged with challenge construct's items. * $p < .05$, ** $p < .001$.

CFI = .909). A second model was then fit with the negatively worded items' residual variances correlated to account for the methodological similarity of these three items. This decision is a supported approach for handling "method variance" shared among specific indicators. Additionally, the negatively worded items greater correlation with each other than with the positively worded items provided evidence that this methodological specific variance existed. The updated measurement model resulted in improved model fit (RMSEA = .054, 90% CI: .021 - .079; TLI = .975; CFI = .986). Despite this improved model fit and the non-significant p -value of the Chi-square model value, examination of the model parameters revealed that item 6 consistently had a factor loading below .5 across the four grade groupings and a large residual variance. By comparison, the rest of the factor loadings were .62 and above, which suggested that item 6 did not add much to the measurement of the latent construct (i.e., was not an effective manifestation of the construct). Therefore, a configural measurement model was fit with only items 1 - 5 loading onto the latent construct. Results with this revised model improved, albeit slightly, model fit (RMSEA = .054, 90%CI: .013 - .082; TLI = .978; CFI = .988) and, more importantly, it resulted in consistent strong loadings for all five items and across all four groupings (Table 3, Panel A provides the fit indices for each of these configural models).

This final five-indicator configural model, then, was used for all subsequent invariance testing.

Weak and Strong Invariance

Invariance testing was conducted to assess the consistency of the engagement with challenge's measurement across the four grade groupings. Weak invariance tests the tenability of equating (i.e., constraining) the construct loadings across multiple groups, and strong invariance tests the tenability of equating the construct intercepts (e.g., item means) across multiple groups. The quality of these constraints is best assessed by the change in CFI being less than .01 compared to the prior, less constrained model's CFI (Δ CFI < .01; Cheung & Rensvold, 2002). The invariance tests were completed across the four grade-levels (9th, 10th, 11th, and 12th grade), with grade 9 as the

reference group (see Table 3, Panel B for model fit indices).

The weak invariance model indicated that enforcing the equality constraint of the item loadings did not significantly diminish the model's fit (Δ CFI = .004), that is, the engagement with challenge's factor loadings were equivalent across grade groups. There was sufficient evidence, then, supporting the tenability of weak invariance—constraining each of the engagement with challenge's five items' loading across the four grade groupings. Test of the scale's strong measurement invariance—equating the item intercepts (i.e., individual item's means) across the four grade-groupings—was conducted next (Table 3, Panel B and Figure 1). Results indicated the strong invariance model's constraints were also tenable based upon the change in CFI (Δ CFI = .006). Therefore, no systematic difference on item means between grade groups was evidenced. Overall, engagement with challenge passed these CFA invariance tests, providing support that the latent variable was measured consistently across grade groupings.

At this point in the analysis, the reliability of the five-item engagement with challenge scale was calculated using equations that take the CFA factor loadings and residual variances into account. Two assessments used for examining a latent construct's internal reliability are the average variance extracted (minimum criterion value is .5; Zinbarg, Revelle, Yovel, & Li, 2005) and the composite reliability (minimum criterion value is .6; Zinbarg, Revelle, Yovel, & Li, 2005). The strong invariance model's standardized loading and residual variance values for the engagement with challenge's five items were used in calculating both of these reliability test values. The average variance extracted passed with a .55 value, and the composite reliability passed with a .85 value. Thus, both tests' values provided additional support for the good reliability of the engagement with challenge measure.

Testing for Homogeneity of Variance/Covariance

Prior to evaluating the criterion-validity of the engagement with challenge measure—i.e., correlating the measure with the criterion variables—the variances and covariances of the engagement with challenge construct across the four grade groupings

Table 3.
Measurement model invariance and homogeneity testing.

Model	$\chi^2 (df)$	p	$\Delta\chi^2 (\Delta df)$	p	RMSEA	90%CI	TLI	CFI	ΔCFI	Maintain Fit, Tenable Constraint?
Panel A—Configural Models										
Null Model, 6 items	2284.02 (180)	0	--	--	.342	.330 - .355	0	0	--	
Configural Invariance 6 items	306.71 (116)	0	--	--	.128	.111 - .146	.859	.909	--	Not adequate
Configural Invariance 6 items, negcorr	134.25 (104)	.025	--	--	.054	.021 - .079	.975	.986	+.077	Yes
Configural Invariance, 5 items, negcorr	103.58 (80)	.039	--	--	.054	.013 - .082	.978	.988	+.002	Yes
Configural Invariance, 5 items, no corr	216.25 (84)	0	--	--	.126	.105 - .146	.882	.931	-.057	No
Panel B—Measurement Invariance & Homogeneity Summary										
Weak Invariance, 5 items with 1 negcorr	134.02 (104)	.025	30.44 (24)	.17	.054	.020 - .079	.978	.984	--	Yes
Strong Invariance, 5 items with 1 negcorr; eng items constrained only	157.29 (116)	.007	23.27 (12)	.03	.060	.033 - .092	.973	.978	--	Yes
Homogeneity Var&Cov Omnibus Test	189.90 (149)	.013	32.62 (33)	.49	.052	.025 - .074	.979	.979	--	Yes
Homogeneity of Latent Means: Engagement (9 - 12), Frequency (9 - 12), Leadership (9 - 11), Advisor Ratings (9 - 11) & Contests (9 - 12)	216.36 (162)	.003	8.90 (3)	.03	.058	.035 - .077	.975	.972	--	Yes

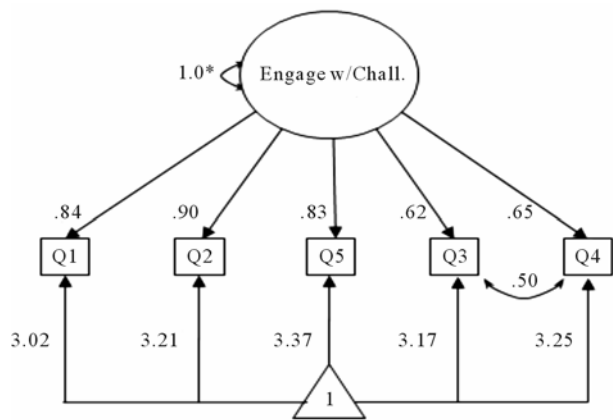


Figure 1.
Strong invariance model for engagement with challenge. Note: This model shows the standardized values from the strong invariance model for the engagement with challenge construct. Constraining the loadings and the intercepts across the grade groups was tenable; therefore only one value for each of these parameters is shown.

needed to be tested. If the variances were found to not be homogeneous, then the variability of the participants' report of engagement with challenge was being moderated by their grade. That is, there would be systematic variance in the engagement with challenge construct associated with grade.

The tenability of equatable variances and covariances of the engagement with challenge construct across all four grade groupings was assessed with an omnibus test of the chi-square difference for nested models. This omnibus test was used be-

cause it is the most conservative test of model fit difference available and is the most appropriate test for latent space significance testing since the latent parameter estimates are unbiased (Kline, 2012). The Chi-square difference test criterion was with an alpha value of .001. Results of this omnibus test indicated that constraining the variances and covariances did not lead to a significant change in model fit ($\Delta X^2_{(33)} = 32.62, p = .49$). Therefore, this more parsimonious model was supported, providing evidence for the homogeneity of the engagement with challenge construct's variances/covariances across grades, which allows for the criterion-validity tests.

Criterion-Validity

Criterion-validity was tested by evaluating the correlations between the engagement with challenge construct and the following criterion variables: 1) FFA student leadership role held; 2) student's frequency of participation; 3) number of contests the student entered that year; and 4) the FFA Advisor's perception of engagement with challenge score for each individual student. The direction and magnitude of these correlations was hypothesized *a priori*. As hypothesized, results indicated that the engagement with challenge construct was significantly correlated with each of the criterion variables and was in the expected direction and magnitude. For example, frequency of participation was moderately and positively correlated with the latent engagement with challenge construct ($r = .490; \Delta X^2_{(1)} = 89.0, p < .001$). The criterion variables were also significantly correlated with each other. **Figure 2** depicts the correlations among the latent engagement with challenge construct and the criterion-variables. These correlations then offer support for the

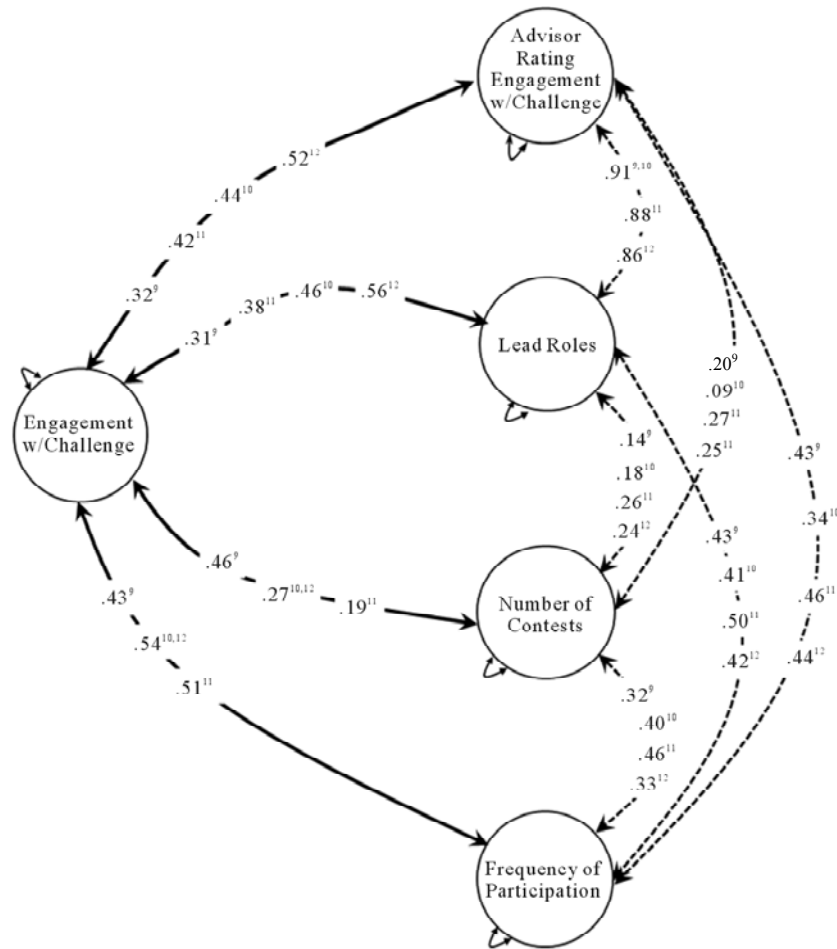


Figure 1. Engagement with challenge & criterion validity variables correlation model.

criterion-validity of the engagement with challenge measure.

The final set of analyses evaluated the homogeneity of the model's means across grades. The final model from the previous step was used to evaluate (still using a nested model approach) if the means across grade groupings for both the latent engagement with challenge mean and the means for the criterion variables were equatable. The homogeneity of the means model was found to fit the data significantly worse than the parent model (homogeneous variance-covariance model; $\Delta X^2_{(14)} = 42.61, p < .001$). Follow-up tests, however, indicated there were significant mean-level differences by grade only for two of the single-indicator criterion variables and not for differences between latent means by grade for the engagement with challenge construct.

Discussion

The purpose of this study was to evaluate the construct-validity of a new self-report instrument of Engagement with Challenge for adolescents. This study evaluated two types of construct validity: structural and criterion. The theoretical focus of the engagement with challenge construct is on an adolescent's intrapersonal synchronous pairing of motivation with challenges, within the context of working on tasks, projects, or goals. Overall, the findings provided good initial evidence for

the construct-validity of this five-item Engagement with Challenge measure.

Structural-Validity

Evidence for the structural-validity of the Engagement with Challenge measure was demonstrated through the fit of the measurement model and the achievement of measurement invariance across grade groupings. In trying to fit the hypothesized measurement model with the original six items to the data, the analyses indicated inadequate fit. There were two sources of this misfit.

Analysis revealed that the first source of misfit was attributable to the inclusion of negatively and positively worded items (i.e., "method effects"). There is a growing research literature suggesting that wording half of the items in a scale in the opposite way of the remaining items (i.e., negatively worded) may result in statistical findings inappropriately indicating two latent factors when they were *a priori* conceptualized as single, unidimensional factor (Chen, Rendina-Gobioff, & Dedrick, 2020; DiStefano & Motl, 2006; Marsh, 1996; Schriesheim, Eisenbach, & Hill, 1991). Analysis to specifically examine this possibility revealed the presence of method effects from inclusion of negatively worded items. Removing this source of variance by allowing the residuals for the negatively worded

items to correlate resulted in a good fit of the hypothesized model to the data.

However, examining the model parameters suggested an additional source of model misfit also be tested. Specifically, the model estimated parameter values indicated item six was potentially distinct from the other items (e.g., large residual variance). Inspection of the wording of this item provided insight as to why it was not a reasonable indicator of the construct. The item reads, "The goals people are working on in this program are not important to me." The focus of this item, then, is on importance of other people's goals in relation to self rather than on an individual's pairing of his or her own motivation to challenges that emerge in pursuit of a program's goal or project. The latter of which is the theoretical basis of the construct. Removal of item six had minimal impact on the model's fit statistics supporting that it was not informing the estimation of the engagement with challenge construct. Thus, the five-item engagement with challenge construct provided a strong and parsimonious explanation of the data, offering initial evidence for structural-validity of the measure.

An important further step in establishing the structural-validity was to determine if the Engagement with Challenge instrument measured the same latent construct for each of the four grade groupings (i.e., invariance). The CFA findings provided good evidence for the invariance of the construct across grade groupings—there were equivalent indicator loadings and intercepts. This provided evidence that the Engagement with Challenge instrument measured the same latent construct for each of the four grade groupings. Overall then, these initial findings provided strong evidence for the structural-validity of the five-item Engagement with Challenge measure. These CFA invariance results provide important support for the measurement model of a construct that should always be reported before the SEM analysis can be conducted and the results interpreted with confidence.

Criterion-Validity

The results of this study also provided evidence for the criterion-validity of the Engagement with Challenge measure. The latent engagement with challenge factor was significantly correlated in the hypothesized (*a priori*) direction and magnitude with the criterion-variables, $r = .31 - .56$; or sharing 10% - 32% of the variance. While the correlations between three of the four criterion variables were self-report, the correlation between the adult advisor's ratings of a student's engagement with challenge supports the interpretation that the criterion findings are not attributable to self-report biases. In addition, the three self-report criterion variables were based on tangible (as possible) behaviors, such as holding a lead role in the program, which were less likely to be influenced by attitudes or ability to accurately report. A particular challenge faced in this study, however, was that the criterion variables were all single indicators, which limited the ability to evaluate their construct validity.

It is important to note the strong correlations between the advisor's reports of each student's engagement with challenge and students' self-report of holding a lead role ($r = .86 - .91$). Further, the correlation between students report of engagement with challenge and lead roles was much smaller ($r = .38 - .40$). This pattern suggests that advisors may rely on a student holding a lead role as a "proxy" or indicator of his or her engagement with challenge (the correlation between advisor's ratings

and the other criterion variables is comparatively much smaller). Given the intrapersonal (psychological) nature of engagement with challenge, which is beyond the direct observation by others, this pattern suggests that the individual adolescent may provide a more direct assessment of this learned disposition.

Construct validity, evidenced in many forms (e.g., structural, criterion), is built over time, and no single study is sufficient to infer validity (Clark & Watson, 1995). Evidence in this study for the structural-validity of the new Engagement with Challenge measure is promising. There is, however, additional research is needed with the measure in order to expand the support for its structural- and criterion-validity. For reasons outlined in the methods section, this was a homogeneous sample of rural high school students in one type of youth program. Therefore, an obvious need is to replicate the structure of the measure and the patterns of associations with other additional variables using a more diverse sample of adolescents. In theory, the engagement with challenge construct should be salient for adolescents across socio-economic and geographic regions. It will also be important to evaluate the construct-validity with younger adolescents, for example in grades 6 - 8, to understand if the instrument measures the same latent construct or if there are developmental limits and differences.

Research with the Engagement with Challenge Measure

Early evidence from this study of construct-validity of the Engagement with Challenge measure suggests it has utility for future research. In theory, engagement with challenge is an important developmental task of adolescence (Larson, 2000). Quantitatively understanding its developmental course, then, is an important research question to address. Research on adolescent brain development suggests that early adolescents experience a heightened period of motivational-emotional arousal (Keating, 2004; Spear, 2000), which could provide an impetus for wanting experiences that satisfy this state. Do younger adolescents, then, experience greater engagement with challenge compared to older adolescents; and are they more likely to experience it across multiple settings or activities? We reason that this motivational-emotional biological "push" could serve to increase the chances that early adolescents will learn to engage with challenge across settings, which could allow them to explore a range of interests and activities through which they can develop, over-time, more specific pursuits. A related question is if engagement with challenge is a general competency that is transferable across settings of individuals' lives, or if it is a state that is setting dependent. Research could evaluate adolescents' engagement with challenge across multiple settings or activities, which could offer insight its stability.

Because, in theory, engagement with challenge results from experiences in different settings, future research could also examine how the particular environment or climate affects its development over time. Research literature on intrinsic motivation (Deci & Ryan, 1985; Larson & Rusk, in press; Ryan & Deci, 2000; Vallerand, 1997) suggests that autonomy supportive environments (e.g., where individuals have control and choice over activities) could promote the development of engagement with challenge. Similarly, a rich research line on the climate of sports teams and academic settings suggests that a mastery/task-involving climate can enhance intrinsic motivation and enjoyment (Cury, Elliot, Sarrazin, Da Fonseca, & Rufo,

2002; Elliot, Cury, Fryer, & Huguet, 2006; Morris & Kavusanu, 2009; Nicholls, 1989). We would expect then, that a highly mastery/task-involving climate would support the development of engagement with challenge.

This study's findings offer evidence for construct-validity and suggest that future research evaluate the antecedents, correlates, and consequences of engagement with challenge. Although important in its own right, engagement with challenge has conceptual ties to other skills or competencies. We suggest sustained engagement with challenge in a setting could grant adolescents access to other experiences for building skills. In a qualitative study of the development of adolescents strategic planning skills, Larson and Hansen (2005) reported that, as high school adolescents in a civic activism program became empowered and motivated to overcome the challenges of influencing the Chicago School board, they began to develop strategic planning skills (e.g., thinking that involves "system processes") for how to accomplish the goals of their project. Thus, future research could explore the longitudinal relationships between engagement with challenge and other skills, such as strategic planning.

Practical Implications and Limitations

Overall, the findings of this study provide promising early evidence for the construct-validity of the Engagement with Challenge measure. As with all cross-sectional data, however, the longitudinal invariance of the construct still needs to be assessed. In addition, this study offers little guidance on what should be the appropriate time interval between measurement points when using this instrument. The measurement time interval that is best for capturing the influence of engagement with challenge on adolescent's higher-order skills is an important research question to answer in order to be able to more accurately increase our understanding of how these skills are learned.

Finally, there are at least two practical implications of this study. The method effect attributed to the inclusion of negatively and positively worded items of the same construct implies that researchers using this measure need to statistically account for the correlated residuals among the two negatively worded items in order for the latent construct to retain its intended meaning. Fortunately, this can be more easily accomplished with modern statistical approaches, such as SEM used in this study, but it poses problems for statistical methods (e.g., standard linear regression) that are unable to account for this relation. A practical advantage of this new measure is that it is short, only five items, which increases its ease of use for researchers and practitioners with minimal risk of fatiguing adolescent participants. Overall then, the findings of this study suggest that the utility of this measure is promising for addressing salient research questions related to adolescents' normative development.

Conclusion

This study addressed a specific quantitative measurement need faced by researchers: assessing salient developmental tasks of adolescence. The findings provided support for the construct validity of Engagement with Challenge. The meaning and impact of engaging with challenge can now be further investigated to understand how characteristics of different set-

tings in adolescents' lives support or hamper its development.

Acknowledgements

We are grateful to colleagues and reviewers who provided comments and suggestions for improvement. Special acknowledgement is given to one reviewer's comments regarding testing for method effects.

REFERENCES

- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.
- Clark, L. & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309-319.
- Csikszentmihalyi, M., & Larson, R. W. (1984). *Being adolescent*. New York: Basic Books.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper Perennial.
- Cury, F., Elliot, A. J., Sarrazin, P., DaFonseca, D., & Rufo, M. (2002). The trichotomous achievement goal model and intrinsic motivation: A sequential mediational analysis. *Journal of Experimental Social Psychology, 38*, 473-381.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Eccles, J., Barber, B., Stone, M., & Hunt, J. (2003). Extracurricular activities in adolescent development. *Journal of Social Issues, 59*, 865-889.
- Eccles, J., & Gootman, J. A. (Eds.) (2002). *Community programs to promote youth development. Committee on community-level programs for youth*. Washington: National Academy Press.
- Elliot, A. J., Cury, F., Fryer, J. W., & Huguet, P. (2006). Achievement goals, self-handicapping, and performance attainment: A mediational analysis. *Journal of Sport and Exercise Psychology, 28*, 344-361.
- Jay, G. (2009) The teen brain: Primed to learn, primed to take risks. *Cerebrum*. URL (last checked 26 February 2009). <http://bit.ly/GieddPrimed>
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., & Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2*, 861-863.
- Goldberg, E. (2009). *The new executive brain: Frontal lobes in a complex world*. Oxford: Oxford University Press.
- Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (1998). Role of cognitively stimulating home environment in children's academic intrinsic motivation: A longitudinal study. *Child Development, 69*, 1148-1460.
- Guay, F., Boggiano, A. K., & Vallerand, R. J. (2001). Autonomy support, intrinsic motivation, and perceived competence: Conceptual and empirical linkages. *Personality and Social Psychology Bulletin, 6*, 643-650.
- Hansen, D. M., & Larson, R. (2007). Amplifiers of developmental and negative experiences in organized activities: Dosage, motivation, lead roles, and adult-youth ratios. *Journal of Applied Developmental Psychology, 28*, 360-374.
- Hansen, D. M., Larson, R., & Dworkin, J. (2003). What adolescents learn in organized youth activities: A survey of self-reported developmental experiences. *Journal of Research on Adolescence, 13*, 25-55.
- Keating, D. (2004). Cognitive and brain development. In R. Lerner, & L. Steinberg (Eds.), *Handbook of Adolescent Psychology* (2nd ed., pp. 45-84). Hoboken, NJ: Wiley & Sons.
- Kuhn, D. & Franklin, S. (2006). The second decade: What develops and how. *Handbook of Child Psychology* (6th ed., pp. 953-993), New York: Wiley.
- Larson, R. W., & Angus, R. M. (2011). Adolescents' development of skills for agency in youth programs: Learning to think strategically. *Child Development, 82*, 277-294.

- Larson, R., & Hansen, D. (2005). The development of strategic thinking: Learning to impact human systems in a Youth Activism program. *Human Development, 48*, 327-349.
- Larson, R. (2000). Towards a psychology of positive youth development. *American Psychologist, 55*, 170-183.
- Larson, R., Hansen, D., & Moneta, G. (2006). Differing profiles of developmental experiences in across types of organized youth activities. *Developmental Psychology, 42*, 849-863.
- Larson, R., Hansen, D., & Walker, K. (2005). Everybody's gotta give: Adolescents' development of initiative within a youth program. In J. Mahoney, R. Larson, & J. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after-school and community programs*. Hillsdale, NJ: Erlbaum.
- Larson, R. W., & Rusk, N. (in press) Intrinsic motivation and positive development. In R. M. Lerner, J. V. Lerner, & J. B. Benson (Eds.), *Advances in Child Development and Behavior: Positive Youth Development*. Oxford: Elsevier.
- Little, T. D. (2010). Structural equation modeling: Foundations and extended applications. Lawrence, KS: University of Kansas.
- Little, T. D. (in press). Longitudinal structural equation modeling: Individual difference panel models. New York: Guilford Press.
- Loevinger, J. (1954). The attenuation paradox in test theory. *Psychological Bulletin, 51*, 493-504.
- Markham, J. & Greenbough, W. (2004). Experience driven brain plasticity: Beyond the synapse. *Neuron Glia Biology, 1*, 351-363.
- Mascolo, M. F., Fisher, K. W., & Neimeyer, R. (1999). The dynamic co-development of intentionality, self, and social relations. In J. Brandstätter, & R. M. Lerner (Eds.), *Action and Self-Development: Theory and Research through the Life Span*. Thousand Oaks, CA: Sage.
- Morris, R. L., & Kayussanu, M. (2009). Approach-avoidance achievement goals in sport: Psychological correlates and a comparison with the dichotomous model. *International Journal of Sport and Exercise Psychology, 9*, 185-202.
- Muthen, L. K., & Muthen, B. O. (1998-2009). *Mplus (Version 5.21) [Base Program (32-bit)]*. Los Angeles, CA: Muthen & Muthen.
- National Research Council & Institute of Medicine (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.
- Olinsky, A., Chen, S., & Harlow, L. (2003). The comparative efficacy of imputation methods for missing data in structural equation modeling. *European Journal of Operation Research, 151*, 53-79.
- Pearce, N., & Larson, R. (2006). The process of motivational change in a civic activism organization. *Applied Developmental Science, 10*, 121-131.
- R Development Core Team (2010). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Richards, J. E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science, 19*, 41-46.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68-78.
- Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews, 24*, 417-463.
- Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 271-360). New York: Academic Press.
- Wu, W., Lang, K. M., & Little, T. D. (2009, October). A simple method of assessing fit and significance in multiply imputed data. *Annual Meeting of the Society of Multivariate Experimental Psychology*.
- Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's, Revelle's, and McDonald's: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*, 123-133.