Creation of Short and Very Short Measures of the Five Cs of Positive Youth Development

Geldhof, G J ; Bowers, Edmond P ; Boyd, Michelle J ; Mueller, Megan K ; Napolitano, Christopher M ; Schmid, Kristina L ; Lerner, Jacqueline V ; Lerner, Richard M

DOI: https://doi.org/10.1111/jora.12039

Creation of Short and Very Short Measures of the Five Cs of Positive Youth Development

G. John Geldhof, Edmond P. Bowers, Michelle J. Boyd,

Megan Kiely Mueller, Christopher M. Napolitano, Kristina L. Schmid

Jacqueline V. Lerner, Richard M. Lerner

Abstract

As developmental scientists seek to index the strengths of adolescents and adopt the Positive Youth Development (PYD) perspective, psychometrically sound measurement tools will be needed to assess adolescents' positive attributes. Using a series of EFA and CFA models, this research creates short and very short versions of the scale used to measure the Five Cs of PYD in the 4-H Study of Positive Youth Development. We created separate forms for earlier versus later adolescence and ensured that items displayed sufficient conceptual overlap across forms to support tests of factorial invariance. We discuss implications for further scale development and advocate for the use of these convenient tools, especially in research and applications pertinent to the Five Cs model of PYD.

Keywords: positive youth development, scale validation, short form, five Cs

The Creation and Validation of Short and Very Short Measures of PYD

Over the past two decades, the development and use of positive indicators of child well-being have increased substantially (Lippman et al., 2011; O'Hare, 2012). The primary movement framing this work has been the positive youth development (PYD; J. Lerner, Phelps, Forman, & Bowers, 2009; J. Lerner et al., 2012) perspective. PYD has been conceptualized in several ways and several theoretical frameworks have been posited (for a review, see J. Lerner et al., 2012). As these models become more popular with individuals working to enhance the positive growth of young people (Beets et al., 2009; Duerden, Witt, Fernandez, Bryant, & Theriault, 2012;; Kurtines et al., 2008), it is important that they are empirically useful, can be widely applied, and include constructs that are specific and measurable. However, these models are just beginning to be tested and there is still a great need for indicators of many instantiations of PYD (Lippman et al., 2011; O'Hare, 2012). Recent work has attempted to evaluate youth development frameworks (Heck & Subramaniam, 2009) and indicators of PYD (Dukakis, London, McLaughlin, & Williamson, 2009; Lippman, Moore, & McIntosh, 2011), but further investigation of suitable models and measures is needed, especially for the Five Cs model of PYD, which Heck and Subramanaim (2009) note is the most widely used approach in both research and youth programs (see too, Eccles & Gootman, 2002).

The purpose of this report is to provide information about a questionnaire developed to assess PYD based on the Lerner and Lerner Five Cs Model of PYD (Bowers et al., 2010; Lerner et al., 2005). This approach employs several measures to index PYD, which have been operationalized through the assessment of Five Cs—competence, confidence, character, connection, and caring. The Five Cs were hypothesized as a way of conceptualizing PYD (and of integrating all the separate indicators of it, such as academic achievement or self-esteem), based

on both the experiences of practitioners and on reviews of the adolescent development literature (Eccles & Gootman, 2002; Roth & Brooks-Gunn, 2003a, b). In addition, these "Cs" are prominent terms used by practitioners, adolescents involved in youth development programs, and the parents of these adolescents in describing the characteristics of a "thriving youth" (King et al., 2005; Roth & Brooks-Gunn, 2003a, b). The Five Cs Model attempts to provide a common terminology for indices of youth well-being that are often absent in the literature (Bradshaw & Guerra, 2008).

The Five Cs have been linked to the positive outcomes of youth development programs assessed by Roth and Brooks-Gunn (2003a, b). In turn, when a youth manifests these Five Cs over the course of adolescence he or she is more likely to be on a life trajectory marked by mutually-influential person ↔ context relations that contribute to self, family, community, and civil society (i.e., contribution – the "sixth C" – emerges; Lerner, 2004). The young person is also less likely to be on a trajectory of risk and problem behaviors, such as substance abuse, delinquency, and depression. That is, as evidence for positive behavior increases, the PYD perspective hypothesizes that there will be fewer indications of problematic behaviors (e.g., Benson et al., 2006; Pittman, Irby, & Ferber, 2001).

The strengths of the Five Cs Model of PYD as a philosophy for youth programming are evident in its adoption both domestically (e.g., Robinson, Esters, Dotterer, McKee, & Tucker, 2012) and internationally (e.g., Haskins, 2010) and in its expanding empirical base. For example, in their review of youth development frameworks, Heck and Subramaniam (2009) compared the strengths and limitations of the five PYD models in terms of their effectiveness, which they evaluated by the criteria of validity (scientific evidence), utility (extent of use and availability of instruments), and universality (applicability to various populations). While Heck and

Subramaniam noted that each of the approaches had benefits and drawbacks, their review indicated that the Five Cs model of PYD is the most empirically supported framework to date. Therefore, a valid, useful, and widely applicable measure of PYD derived from the Five Cs model may greatly benefit both researchers and practitioners who have adopted this approach.

Prior Measurement of the Five Cs Model of PYD

Using data from the first wave (Grade 5) of the 4-H study, Lerner et al. (2005) proposed and tested a higher-order measure of PYD that consisted of five first-order latent constructs, each representing one of the Five Cs of PYD. In a subsequent study, confirmatory factor analyses tested the validity of the Five Cs model (Jeličić et al., 2007). Results suggested that the Five Cs can be cast in terms of latent constructs, which in turn load on a higher-order PYD construct. More recently, Phelps et al. (2009) extended Lerner et al.'s (2005) Grade 5 findings by assessing the structure and development of PYD from Grade 5 to Grade 7 of the 4-H Study. The authors wanted to determine if there was evidence of a latent construct of PYD that generalized across the early years of adolescent development and whether it could be operationalized by lower-order latent constructs representing the Five Cs. Results indicated that the Five Cs Model of PYD continued to be a robust construct that can be defined comparably in Grades 6 and 7 as it was in Grade 5. Finally, Bowers et al. (2010) examined whether the structure of PYD in middle adolescence (Grades 8 through 10) was comparable to the structure of PYD identified in early adolescence. Using a hierarchy of second-order confirmatory factor analysis models to address this issue, Bowers and colleagues found that, while the overall structure of PYD was maintained across Grades 8 to 10, the scales relevant to measuring the Five Cs were slightly different for two of the Cs during middle adolescence as compared to early adolescence. Athletic competence was no longer a relevant indicator of competence during middle adolescence; however, physical

appearance significantly loaded on the latent construct of confidence. Thus, the structural definition of PYD has been confirmed within the 4-H Study data set from the beginning of the adolescent period through the middle portion of this time of life (i.e., from approximately 10-16 years).

Limitations of Prior Measurement Models of the Five Cs of PYD

Prior research using data from the 4-H Study of PYD (Bowers et al., 2010; Phelps et al., 2009) established the existence of a valid measure of PYD across early to middle adolescence. The methodology of these studies, however, had limitations that require further investigation. First, although the results of Bowers et al. (2010) suggested that the initial Five Cs model verified for Grades 5 to 7 should be modified for middle adolescents, the conclusion was drawn based on results obtained with a sample that is different from the one used in Phelps et al. (2009). Although there is overlap in these two study samples, it is possible that the original measurement structure does not fit middle adolescents because the model was tested on data from a different sample.

In establishing measurement invariance to address earlier limitations, Bowers and colleagues accomplished a critical step for further research and practice using the Five Cs model of PYD. However, the findings of Phelps et al. (2009) and Bowers et al. (2010) are hindered by their reliance on only three waves of what is now a longitudinal study of eight waves of data. In order to best test the structure of the Five Cs Model of PYD and whether the model is invariant over time, one needs to derive a model based on the data from the same participants across all eight waves of the data. As the PYD perspective is adopted in more youth-serving programs, the need for a measure that can be utilized across different ages, as well as in different contexts, becomes paramount.

Previous research has also made potentially untenable assumptions that have not yet been empirically tested. From a modeling standpoint, previous research has assumed that acceptable model fit for a higher-order CFA implies that a high-order model is appropriate. While the good fit presented in previous publications suggests that a higher-order model is reasonable, no previous research has directly tested whether a higher-order CFA model fits the 4-H data as well as a model without the higher-order PYD construct. Previous findings have also exclusively relied on item parcels when examining the factor structure of PYD. Parceling is generally appropriate when testing substantive hypotheses via CFA and SEM (e.g., Little, Cunningham, Shahar, & Widaman, 2002), especially when the number of indicators to be modeled is large. Item parcels do not provide information about the appropriateness of individual items, however, and the quality of the individual items used to measure PYD remains unexamined in the literature.

In addition to the above limitations, measures used by researchers and practitioners must be practical and have utility for users. Phelps et al. (2009) and Bowers et al. (2010) utilized a measure of the Five Cs that included over 80 items. Therefore, the time and energy commitment by researchers, practitioners, and youth may be exhausted in trying to index PYD for empirical and applied purposes. Often researchers want to examine the relations of several contexts to a wide breadth of both positive and negative youth outcomes and only have access to their sample for a relatively small amount of time due to the constraints of context (e.g., class time) or the individual (e.g., age of participants). Youth-serving professionals are also often constrained by time commitments to obtain data about their impact on youth. Often, these professionals are volunteers with additional home and family commitments who also want to provide an enjoyable and rewarding experience for the youth in their care. Even for the comprehensive model indexed

by the Five Cs PYD measure, 80 items may be too many to include in a survey to measure one construct, given these limiting factors. Therefore, a shortened scale to measure the Five Cs of PYD would be of great practical benefit to practitioners and researchers.

In addition, a shorter measure of the Five Cs of PYD can be included on waves of (longitudinal) surveys in other studies of adolescent youth (ages 10-18). If a shorter PYD was included as a complement to other measures of youth development, then the short form of Five Cs measure would added to the literature and could be used by researchers in many disciplines. The availability of this measure from use across a wide range of disciplines could greatly advance scholarship on the positive development of youth.For instance, the American Academy of Pediatrics supports work framed by the "Cs" of positive development (Ginsburg & Jablow, 2011).

In the present research we revise and reanalyze the measure of the Five Cs of PYD used in the 4-H study with an eye toward the above limitations. We analyze data from multiple waves of the 4–H study, empirically test whether a higher-order structure fits the data better than a model without a higher-order construct, and shorten the measure significantly such that we can analyze the quality of individual items instead of relying on the aggregate properties of item parcels. Our goals in the present research are therefore to revise and evaluate the validity of a shortened measure that captures the Five Cs of PYD. By achieving these goals, we will provide a reliable, valid, and useful tool for usable by researchers and practitioners alike.

Method

Participants

We analyzed data from 7,071 adolescents who participated in the 4-H Study and completed a measure of PYD in at least one wave of data collection. We recruited participants from 42 states. The mean age of participants was 10.94 (*SD* = .42) in the Grade 5 assessment and

17.71 (*SD* = .76) in Grade 12. Participants were slightly more likely to be female than male (60% female). With respect to race/ethnicity, the sample was 65.8% White; 7.4% Black; 9.4% Latino; and 17.5% other (including Asian, Native American, Multiethnic or multiracial, listed as "other," or not provided). Participants resided in diverse communities, with 35.7% living in rural areas; 16.3% in urban areas; and 25.7% in suburban areas (22.2% had missing data for locale).

Participants' parents provided data regarding the socioeconomic status of their families. In Grade 5, 20% of mothers attended or completed high school; 24.8% completed some college; and 18.6% had a bachelor's degree or higher (35.8% did not respond); average per capita income at Grade 5 was about $13,657 (*SD* = $8,348), and increased to $23,401 (*SD* = $13,798) in Grade 12.

Measures

Positive Youth Development. We operationalized PYD by the Five Cs of PYD discussed above. The Five Cs model identifies PYD as composed of the Five Cs noted above. We present a brief description of our measures below, although our measure drew items from several primary sources and are described in more detail elsewhere (e.g., Bowers et al., 2010; Lerner et al., 2005; Phelps et al., 2007).

Competence. In Grades 5 through 7 we measured competence using 19 items representing academic, social, and physical competence (six items per subscale) as well as academic grades (one item). Based on the findings from Bowers et al. (2010), beginning in Grade 8 and continuing through Grade 12, competence was comprised of four items measuring scholastic competence, five items measuring social acceptance, five items measuring physical competence, and again academic grades.

The scholastic, social, and physical competence items asked participants to select the type of person they were more like between two choices (e.g., "Some teenagers feel that they are just as smart as others their age, BUT Other teenagers aren't so sure and wonder if they are as smart"; or "Some teenagers are popular with others their age, BUT Other teenagers are not very popular") and then to decide if it was "really true" or "sort of true" for him or her (see Harter, 1983; 1988). Cronbach's alphas for the full competence subscale ranged from 0.80 to 0.86 across Grades 5 through 12.

Confidence. For Grades 5 through 7 we measured confidence using items that represent self-worth, physical appearance, and positive identity (six items each). Based on findings from Bowers et al. (2010), beginning in Grade 8 and continuing through Grade 12, the confidence subscale included items that measured self-worth (five items), physical appearance (five items), and positive identity (six items) that had a similar structure and response format.

The self-worth and physical appearance items asked respondents to select the type of person they were more like between two choices (e.g., "Some kids *like* the kind of *person* they are, BUT Other kids often wish they were someone else") and then to decide if it was "really true" or "sort of true" for him or her. Positive identity items were scored on a five-point Likert scale with response options ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Cronbach's alphas for the full confidence subscale ranged from 0.80 to 0.92 across Grades 5 through 12.

Character. In Grades 5 through 7 we measured character using 21 items representing social conscience (six items), values diversity (four items), conduct behavior (six items), and personal values (five items). Based on findings from Bowers et al. (2010), beginning in Grade 8

and continuing through Grade 12, the character subscale included 20 items with conduct behavior measured by only five items.

A sample social conscience item stated, "How important is each of the following to you in your life? Helping to make the world a better place to live in," with response options ranging from 1 = *not important* to 5 = *extremely important*. A sample values diversity item instructed respondents to think about people who know them well and indicate how they would rate the young person on characteristics including, "Respecting the values and beliefs of people who are of a different race or culture than I am" with a response format ranging from 1 = *not at all like me* to 5 = *very much like me*. The conduct behavior items asked respondents to select the type of person they were more like between two choices (e.g., "Some kids usually do the *right* thing BUT Other kids often *don't* do the right thing") and then to decide if it was "really true" or "sort of true" for him or her. The personal values items assessed the importance of certain values in the young person's life, including "standing up for what I believe, even when it's unpopular to do" with response options ranging from 1 = *not important* to 5 = *extremely important*. Cronbach's alphas for the full character subscale ranged from 0.89 to 0.93 across Grades 5 through 12.

Caring. In the 4-H Study of Positive Youth Development, Caring was originally measured using the Eisenberg Sympathy Scale (ESS; Eisenberg et al., 1996) in the first wave of data collection at Grade 5. For each item, participants indicated the degree to which a statement described him or her (e.g., When I see someone being picked on, I feel kind of sorry for them) with response options ranging from 1 = *not like you* to 3 = *really like you*. Beginning in Grade 6, we introduced a nine-item Caring measure that was developed by modifying the five ESS items and adding four items adapted from the Empathic Concern (EC) subscale of the Interpersonal

Reactivity Index (IRI; Davis, 1980, 1983). An example of these new items was, "How well does each of these statements describe you? When I see someone being taken advantage of, I want to help them," with a five-point Likert scale ranging from 1 = *not well* to 5 = *very well*. The Caring items were revised based on the evolving definitions of the Five Cs; according to Phelps et al. (2009), the original five ESS items were not thought to accurately capture the Caring construct. Beginning in Grade 7, we dropped the original five-item ESS in favor of the new 9-item assessment of Caring. Cronbach's alphas for the full caring subscale ranged from 0.80 to 0.88 across Grades 5 through 12.

Connection. We measured connection using items that represent connection to family (six items in Grades 5 to 7; 5 items in Grades 8 to 12, based off of findings from Bowers et al., 2010), neighborhood (five items), school (seven items), and peers (four items). A sample connection to family item stated, "How much do you agree or disagree with the following? In my family, I feel useful and important." A sample connection to neighborhood item assessed respondents level of agreement with the following statement, "In my neighborhood, there are lots of people who care about me." A sample connection to school item stated, "How much do you agree or disagree with the following? I care about the school I go to." All items were scored on a five-point Likert scale with response options ranging from 1 = *strongly disagree* to 5 = *strongly agree*, except the connection to peers scale. Items measuring connection to peers asked respondents to indicate the accuracy of statements, including "I trust my friends" with response options ranging from 1 = *never true* to 5 = *always true*. Cronbach's alphas for the full connection subscale ranged from 0.89 to 0.92 across Grades 5 through 12.

Outcomes. To ensure that the factor structure of PYD is robust (i.e., does not change) in the presence of important outcomes, we included unit-weighted composites (i.e., scale scores) for

the following scales in all confirmatory factor analyses. Outcome measures were included for each wave of data and, while not a primary focus of this research, the substantive correlations between PYD and these outcomes are presented in the results below.

Contribution. We measured contribution using two equally weighted subscales: ideology and actions. Each subscale included 6 items, all administered using a five-point Likert scale. The ideology subscale measured the extent to which contribution was an important facet of youth's identity and future self. An example ideology item stated, "It is important to me to contribute to my community and society." The action subscale was comprised of three components: helping, leadership, and service. Items from the helping, leadership, and service components measured the frequency of time youth spent helping others (i.e., friends and neighbors), acting in leadership roles (i.e., being a leader in a group or organization within the last 12 months), and providing service to their communities (i.e., volunteering, mentoring or peer advising, and participating in school government), respectively. The composite contribution scores ranged from 0 to 100, with higher scores indicating higher levels of contribution. For the current study sample, the Cronbach's alphas for the contribution scale were .40 at Grade 5 and .68 at Grade 6; however, the alphas ranged from .75 to .81 across Grades 7 through 12.

Depression. We measured depressive symptomatology using the 20-item self-report Center for Epidemiological Studies Depression scale (CES-D; Radloff, 1977). Respondents indicated how often they experienced particular symptoms during the past week. Example items included: "I was bothered by things that usually don't bother me" and "I felt sad." Four items were positively worded and included: "I felt hopeful about the future" and "I enjoyed life." The response options ranged from 0 = *rarely or none of the time (less than 1 day)* to 3 = *most or all of the time (5-7 days)*. Items were summed for a total score, with a maximum value of 60.

Cronbach's alphas for the CES-D scale ranged from 0.81 to 0.89 across Grades 5 through 12 of the present study.

Risk behaviors. We assessed indicators of substance use and delinquency derived from items included in the Search Institute's Profiles of Student Life-Attitudes and Behaviors (PSL-AB) scale (Leffert et al., 1998) and the Monitoring the Future (2000) questionnaire to indicate adolescent risk behaviors.

Substance use. At Grade 5, five items assessed the frequency of substance use during the past 12 months. Specifically, we asked students whether or not they had ever smoked cigarettes; used chewing tobacco or snuff; had any beer, wine, wine coolers, or liquor to drink – more than just a few sips; used marijuana (grass, pot) or hashish (hash, hash oil); and used any other drug, such as ecstasy, speed, LSD, heroin, crack or cocaine. In addition to the previously mentioned items, students in Grades 6 through 12 indicated whether they had ever sniffed glues, sprays or gases. We then added a final item asking whether respondents had ever taken steroid pills or shots without a doctor's prescription in Grades 7 through 12. The response options for all substance use items ranged from 0 = *never* to 3 = *regularly*.

Delinquency. We assessed Grade 5 and 6 delinquency using four items that indicated the frequency of delinquent behavior during the past 12 months. Specifically, we asked students how many times they had stolen something from a store; gotten into trouble with the police; hit or beat up someone; and damaged property just for fun (such as breaking windows, scratching a car, putting graffiti on walls, etc.). At Grade 7 and continuing through Grade 12, an additional item assessed how many times the student carried a weapon (such as a gun, knife, club, etc.). The response format for the delinquency items ranged from 0 = *never* to 4 = *five or more times*. In

order to create more meaningful categories, responses were recoded as 0 = *never,* 1 = *once or twice*, 2 = *3-4 times*, and 3 = *five or more times.*

We calculated averages for the substance use and delinquency items, respectively, and transformed (multiplied by five) to have a range from 0 to 15. A composite measure was then calculated by summing the averages of both subscales for a maximum score of 30. For the current study sample, the Cronbach's alphas for the risk behaviors scale were .65 at Grade 5 and .71 at Grade 6; the alphas ranged from 0.76 to 0.86 across Grades 7 through 12.

Analyses

Our goal was to analyze the factor structure of PYD and to reduce the 80+ item PYD measure used in the 4-H Study into shorter formats that can be more easily implemented in large research studies. We accomplished this objective through a series of exploratory and confirmatory factor analyses and invariance tests that resulted in four scales: separate short (34 items; PYD-SF) and very short (17 item; PYD-VSF) PYD scales for both early and middle to late adolescents, respectively. These analyses emphasized parsimonious representation of the Five Cs, utilized a consistent format across the early versus middle to late adolescent forms, and underscored the importance of item heterogeneity.

We began by fitting exploratory factor analyses (EFAs) using two waves of the 4-H data set, which contained data for the early adolescent and the middle to late adolescent versions of our PYD measure, respectively. We extracted factors using maximum likelihood with geomin (oblique) rotation in M*plus*. We then used the results from these analyses to select two items from each subscale of each C to be used in our PYD-Short Form scale (PYD-SF).

We selected items for the PYD-SF according to the following criteria, listed in order of importance: any item retained in the PYD-SF had to display a strong factor loading onto at least

one construct in the EFAs; items in the early adolescent questionnaire that displayed strong conceptual overlap with items in the middle to late adolescent questionnaire (and vice versa) were preferred over items that did not display such overlap; we retained two items from each subscale of each C to ensure construct heterogeneity; and items with strong target factor loadings were preferred to items with weaker loadings.

We next performed a series of confirmatory factor analyses (CFAs) of the PYD-SF using data from separate waves of the 4-H study. These analyses (a) empirically tested whether a higher-order PYD construct fit the data, (b) ensured that the factor structure of the PYD-SF was longitudinally stable, and (c) determined whether responses to the early-adolescent version of the PYD-SF could be compared to the middle to late-adolescent scale. A subsequent set of models then examined whether the 34-item PYD short measures could be reduced to 17-item very short measures.

All CFA analyses implemented full-information maximum likelihood (FIML) which produces unbiased parameter estimates under the assumption that data are missing at random. While in any data set some proportion of the data will not be missing at random, our use of FIML allows us to capture that percentage of the variance that is recoverable given the analyzed data. FIML thus makes the same assumptions as other modern approaches to missing data (e.g., multiple imputation). Table 1 presents the number of cases and average percentage of missing data for each wave.

Insert Table 1 about here

Results

Exploratory Factor Analyses

We performed separate EFAs for each of the Five Cs using the Grade 6 and Grade 11 data, producing a total of 10 EFA models. All final EFA models displayed acceptable fit (i.e., RMSEA < .08, CFI and TLI > .90), and generally suggested one factor for each modeled subscale (e.g., separate factors for academic, social, and physical competence). Reverse-coded items also tended to form method-effect factors with the reverse-coded method effect being more pronounced in the Grade 6 analyses than in the Grade 11 analyses.

Creation of the PYD-SF

Using results from the EFAs described above, we next selected two items per subscale to represent each of the Five Cs in the two PYD-SF surveys. Caring did not contain subscales, however, and six caring items were retained. Caring items were selected such that three items in the early-adolescent PYD-SF closely matched items in the sympathy scale used to represent caring in Grade 5 of the 4-H data set (and were also included in the Grade 6 data).

Due to the unanticipated method effect for reverse-coded items, we omitted all items that loaded onto a reverse-coded method factor. In addition, when subscales clearly differentiated into two separate factors in our EFA models we purposefully retained one item from each factor to ensure fully heterogeneous construct measurement in the PYD-SF. All items included in the final PYD-SF scales are freely available online from http://ase.tufts.edu/iaryd/researchPositive4HpydResources.htm

Validation of the PYD-SF

We used three sets of CFA models to examine the validity of the PYD-SF and test whether the early-adolescent PYD-SF could be directly compared to the middle/late adolescent PYD-SF. To test the structure of the PYD-SF across a broad range of ages, and to ensure

factorial invariance across the early and middle/late adolescent forms, we examined the factor structure of the PYD-SF in Grades 6 and 7, Grades 7 and 8, and in Grades 9 and 12, respectively[4].

We first estimated a CFA that only examined the caring and sympathy items in Grade 6 (the only wave of data in which both the sympathy and caring scales were administered), specifying each as a latent construct and estimating residual covariances between caring and sympathy items that contained strong conceptual overlap. The CFA displayed acceptable model fit ($\chi^2(23) = 87.29$, $p < .001$, RMSEA = .05, 90% C.I. [.04, .06]; CFI = .99, TLI = .98), although a subsequent model indicated that the latent correlation between the constructs (estimated as .48) was significantly different from 1.00 ($\Delta \chi^2(1) = 211.49$, $p < .001$). Sympathy and caring were therefore modeled separately in the Grade 6 data, and sympathy items were omitted from the PYD-SF to promote scale parsimony.

A subsequent CFA therefore specified six latent constructs for the Grade 6 data (the Five Cs plus sympathy) and five latent constructs for the Grade 7 data (the Five Cs only). Because two indicators per subscale were included in the PYD-SF, we estimated residual covariances among same-subscale indicators within and across time.

Based on model modification indices we allowed a dual loading between one character item and competence ("Some kids usually act the way they know they are supposed to") and a residual covariance between one competence item ("Some kids are popular with others their age") and one confidence item ("Some kids think that they are attractive or good looking"). The resulting CFA displayed acceptable model fit ($\chi^2(2647) = 5594.99$, $p < .001$, RMSEA = .02, 90% C.I. [.02, .02]; CFI = .94, TLI = .93). We next determined whether a higher-order PYD construct could parsimoniously represent the 5 Cs. Our results suggested that imposing a higher-order

PYD factor fit the data reasonably well ($\chi^2(2700) = 5801.68$, $p < .001$, RMSEA = .02, 90% C.I. [.02, .02]; CFI = .92, TLI = .92), but fit significantly worse than a model without the higher-order PYD construct even after relaxing several model constraints suggested by high modification indices (e.g., allowing residual covariances among specific Cs; $\Delta \chi^2(53) = 206.69$, $p < .001$).

While much of the existing literature treats PYD as a higher-order construct, other research has suggested that a bifactor model might be more appropriate than a higher-order model when modeling multi-dimensional scales such as PYD (e.g., von Eye, Martel, Lerner, Lerner, & Bowers, 2011). In a bifactor model, a more global construct of interest (e.g., PYD) is modeled as a direct function of items rather than only being modeled as a function of lower-order latent constructs. Thus, each item indicates a lower-order construct *and* a more general construct by loading onto each simultaneously. Accordingly, the term 'bifactor' represents reflects the dual nature of each item.

The bifactor model relaxes the assumption that relations among lower-order factors, and the relations between these factors and important criterion measures, can be fully explained by a single higher-order construct. The bifactor model also alleviates the assumption that indicators are only related to PYD because they indicate the Five Cs. The bifactor model instead allows indicators to separately load onto their respective lower-order constructs and onto the more general higher-order construct. The higher-order construct is modeled to be orthogonal to the lower-order constructs, such that the lower-order constructs represent residual constructs after controlling for the higher-order construct (e.g., competence that is not directly related to PYD). These lower-order constructs can then be allowed to correlate with each other and with important criterion measures.

We compared the above PYD-SF five-factor model to a bifactor model, where the dual loading and residual covariance specified in the five-factor CFA were included in the bifactor model. Results indicated that the bifactor model fit the data very well ($\chi^2(2563) = 4500.22$, $p <$ .001, RMSEA = .02, 90% C.I. [.02, .02]; CFI = .96, TLI = .95) and suggested that the five-factor CFA fit the data significantly worse than the bifactor model ($\Delta \chi^2(84) = 1094.78$, $p < .001$), even after adjusting for model parsimony (Δ BIC = -440.12; Δ aBIC = -707.00). We then established partial longitudinal weak and strong factorial invariance for this model using the criterion suggested by Cheung & Rensvold (2002; i.e., Δ CFI < .01 for each level of invariance). Results from the strong invariance CFA are presented in the following tables: Table 2 presents standardized factor loadings, while Tables 3 and 4 present latent correlations among the PYD constructs and between the PYD constructs and the outcomes, respectively. In our models the interpretation of the reliability of each scale is not entirely straightforward because each item has multiple sources of true score variance. Despite this limitation, we provide composite reliability estimates (ω) for each C of PYD in Table 2. We provide ω estimates in lieu of the more common coefficient α because ω is computed as a function of factor analysis estimates and provides a more accurate estimate of reliability. Both ω and α estimate the same parameter, however, and readers can interpret ω as they would α.

Insert Tables –2 - 4 about here

A second set of models established the factor structure of the PYD-SF in Grades 7 and 8 of the 4-H study. As above, a bifactor model displayed good model fit ($\chi^2(2366) = 4622.25$, $p <$

.001, RMSEA = .02, 90% C.I. [.02, .02]; CFI = .95, TLI = .94) and fit the data significantly better than a CFA without the PYD factor ($\Delta \chi^2(84) = 1210.09$, $p < .001$) even after adjusting for model parsimony (Δ BIC = -576.83; Δ aBIC = -834.19). Weak and strong factorial invariance were established across time, and because participants completed the early-adolescent PYD-SF in Grade 7 but the middle/late PYD-SF in Grade 8, these results indicate that the latent PYD constructs are directly comparable across our early and middle/late adolescent forms. Results from the strong-invariance CFA are presented in the same tables as the Grade 6 and 7 models above (i.e., Tables 2 through 4).

A final set of PYD-SF CFA models established the factor structure of the PYD-SF in Grades 9 and 12 of the 4-H study and tested the structure's factorial invariance across these waves. The bifactor model again displayed good model fit ($\chi^2(2366) = 4178.43$, $p < .001$, RMSEA = .02, 90% C.I. [.02, .02]; CFI = .93, TLI = .92) and fit the data significantly better than a CFA without the PYD factor ($\Delta \chi^2(81) = 732.33$, $p < .001$) even after adjusting for model parsimony (Δ BIC = -126.34; Δ aBIC = -383.66). Weak and strong factorial invariance were established across time and details of the invariance tests and results from the strong-invariance CFA are presented in the same tables as the Grade 6 and 7 models above (i.e., Tables 2 through 4).

Analyzing the standardized factor loadings for all the above CFA models (Table 3) clearly shows that some items most strongly represent domain-general PYD, others most strongly represent the residual C factors, and still others load onto both constructs. PYD is indicated by items from all Five Cs although the social competence, physical competence, and physical appearance do not contribute a meaningful amount of variance to the PYD construct, for instance. Similarly, the residual Five C constructs are indicated by nearly all of their respective

indicators, but items from the conduct behavior subscale do not meaningfully load onto the residual character construct. These differences are discussed as they pertain to scale implementation in the discussion section below.

Creation and Validation of the PYD-VSF

For our final set of analyses we selected one item per subscale to be retained in the very short form of the 5 Cs measure of PYD (PYD-VSF), using standardized factor loadings from the PYD-SF CFA models (Table 3) as a guide. We omitted all items that displayed a dual loading or unexpected residual covariance with another item *a priori* and, because most items loaded more strongly onto the individual C constructs than on PYD, items with stronger loadings onto PYD were preferred to items with weaker PYD loadings. As before, the caring items were an exception in that caring was not comprised of any subscales. We instead selected three caring items: one that displayed a stronger loading onto the general PYD construct, one that displayed a stronger loading onto the residual caring construct, and one that loaded strongly onto both constructs. All items retained in the PYD-VSF are available from the first author upon request.

We examined the PYD-VSF with a series of bifactor CFA models that matched those used for the PYD-SF. We then compared these CFA models to nested models in which the factor loadings, intercepts, latent means, and latent covariance matrices were equated to those in the parallel PYD-SF models. These tests ensured that the PYD-VSF adequately captured the same constructs as the larger PYD-SF. Likelihood ratio tests for these comparisons indicated invariance of all parameter estimates across the short and very short forms. (Grades 6 and 7: $\Delta \chi^2(235) = 282.525$, $p > .001$; Grades 7 and 8: $\Delta \chi^2(235) = 166.264$, $p > .001$; Grades 9 and 12: $\Delta \chi^2(235) = 150.165$, $p > .001$).

Discussion

As the PYD perspective is adopted by more researchers of adolescent development, youth policy makers, and youth-serving professionals, the need for an easily-used measure of youth thriving becomes paramount. Individuals from each of these fields are often limited by issues of practicality and utility that require concise, yet comprehensive, valid and reliable indices of youth development. For example, researchers are often constrained by ethical and scientific concerns that limit the length of time they can survey youth as part of a study, and practitioners are limited by competing desires of funders, supervisors, parents, and youth when attempting to evaluate the impact of their programs.

In this study we used a series of EFA and bifactor CFA models to create short (PYD-SF) and very short (PYD-VSF) versions of the scale used to measure the 5 Cs of PYD in the 4-H Study of Positive Youth Development (e.g., Bowers et al., 2010; Lerner et al., 2005; Phelps et al., 2009). Our analyses addressed major limitations facing research that has previously examined the Five Cs of PYD; we examined the measure across more waves of measurement; reduced the length of the scale; examined items rather than parcels; and empirically tested the tenability of a higher-order factor structure. Our analyses emphasized parsimonious representation of the Five Cs, utilized a consistent format across the early versus middle/late adolescent forms, and underscored the importance of item heterogeneity (e.g., Cattell, 1961). We created separate forms for early versus middle/late adolescents and ensured that items displayed sufficient conceptual overlap across forms to support tests of factorial invariance. Despite the parsimony of our shortened scales and the psychometric benefits of having strong conceptual overlap across forms, our scale's bifactor structure is not as straightforward as the structure of many research instruments. The following sections briefly discuss how the short versions can be most optimally analyzed.

Using the PYD-SF and PYD-VSF

Bifactor CFA. Our scale is best analyzed using the bifactor technique described above. When possible, we therefore recommend bifactor CFA and SEM analyses for researchers who wish to analyze relatively large datasets. As with the results presented in this article, researchers analyzing PYD with a bifactor model should expect lower factor loadings than those generally found when utilizing CFA and SEM because each loading only represents part of an item's true score variance. An item's C-specific and PYD factor loadings should be considered together when interpreting the quality of individual items.

While CFA or SEM is generally preferred, CFA is not the optimal statistical method to answer every research question. In the following sections we therefore consider how scale composites might be used to test substantive hypotheses using the PYD-SF. Scale scores can also be created using the PYD-VSF, although the low number of items per factor could substantially reduce estimates of scale reliability.

PYD scale score. Scale scores are easily-computable and provide a single number which researchers and practitioners can use to gauge an individual's global level of positive development. Despite the bifactor structure used in the present analyses, we encourage the use of an overall PYD scale score when the sophistication of a bifactor CFA model is unreasonable. Not all items included in the PYD short versions adequately represent PYD as a construct, however, and caution is warranted when computing an individual's PYD composite score. In particular, items from the physical competence, social competence, and physical appearance subscales did not strongly reflect our general PYD construct, and these items should be omitted when computing PYD composite scores.

Scale scores for individual Cs. Researchers and practitioners may alternatively wish to examine an individual's (or sample's) scores for each C separately. For instance, a practitioner working to increase a child's self-worth may only need to measure that child's level of confidence. In such cases, C-specific composites can be created by summing across all items that represent a given C. Creating such composites implicitly assumes that the overlap between individual Cs and PYD is not important. Including items that do not strongly represent the residual C constructs in our models is justified here, because those items simply represent the component of each C that is also related to PYD.

Comparison to Previous Findings

Our findings generally support the Five Cs model of PYD and the factor structure's longitudinal invariance as previously reported by Bowers et al. (2010). The bifactor structure of our models contradicts the higher-order structure of PYD found in previous research, however. As with previous research, we found that higher-order CFA models displayed acceptable fit, but empirical tests showed that the higher-order models fit significantly worse than models without a higher-order PYD factor. These findings therefore build on previous research, indicating that a higher-order CFA model can provide a rough estimate of PYD's factor structure but that a bifactor model is more appropriate.

Utilizing a bifactor model structure provides several benefits to researchers that are not possible with higher-order CFA models. The bifactor model relaxes the assumptions that relations among lower-order factors and other measures are fully explained by a single construct and that indicators are only related to PYD because they indicate one of the Five Cs. Therefore researchers are able to examine relations among lower-order constructs with each other and other youth outcomes that may be important. For example, the results presented above show that many

Cs consistently correlate with contribution, depression, and risk, even after "controlling" for PYD by specifying a bifactor model.

Previous findings have also exclusively relied on item parcels when examining the factor structure of PYD, clouding the quality of individual items. The present study, however, examined the quality and appropriateness of the individual items that have been used to measure PYD in the 4-H Study. Therefore, analyses utilizing the bifactor model with individual items have provided both novel and more nuanced results than prior work (Bowers et al., 2010; Phelps et al., 2009). For example, PYD is indicated by items from all Five Cs, although the social competence, physical competence, and physical appearance items do not contribute a meaningful amount of variance to the PYD construct. Thus, consistent with prior work, the Five Cs model of PYD is retained, but only particular indices are useful when differentiating adolescents on overall PYD.

Limitations and Future Directions

While we made every effort to ensure the psychometric quality of the PYD-SF and PYD-VSF, its administration is not without drawbacks. The scales' primary limitation stems from the fact that items were drawn from multiple sources and accordingly are scored using different metrics. Future research is needed to determine whether it is appropriate to administer all items using a similar format (e.g., all items using a 5-point Likert scale) or if the differential scoring formats are integral to the structure of our scales. This limitation may make our proposed scales especially difficult to administer, and our lab is actively exploring more facile ways to administer these items. We have already begun collecting data that that involves the administration of all items in the PYD-VSF using a 5-point Likert-type response scale. We also have plans for

translating these items to follow a single response format (e.g., scores for all items ranging agree to disagree).

Another limitation stems from the fact that we omitted all items that loaded onto an unexpected method factor for reverse coded items in the EFA models. The presence of such a factor may be reflective of social desirability or poor emotional regulation. Indeed, youth seemed to respond in an adverse manner to any items with negative valences (e.g., "I don't feel sorry for other people when they are having problems," as a measure of caring), and this factor was more pronounced in younger grades, when individuals are less adept at emotional regulation (Rossman, 1992; Ryan, 1989; Zimmer-Gembeck & Skinner, 2011). Future studies should attempt to model responses to these reverse-coded items as a function of emotional regulation and social desirability.

As with many large-scale longitudinal studies, our results are also based on a highly selective sample. Our sample largely consisted of white middle-class participants, bringing the generalizability of our results into question. The PYD model is derived from a relational developmental systems perspective and it is reasonable to question whether the relations among the Five Cs or even the factor structure of PYD could be moderated by contextual factors.

Finally, some of our variables displayed a high amount of missingness. While our use of FIML allowed us to re-capture some of this missing information, testing the assumption that data are missing at random is logically not possible. Data that are missing not at random are missing because of their own value (e.g., participants with low character selectively not responding to the character items) or are missing because of an unmeasured factor (e.g., participants' attitudes toward psychological research). Thus, any amount of missingness will introduce uncertainty into

analyses, suggesting the need to replicate the validity of our measures in future independent samples.

In summary, as developmental scientists seek to study the strengths of adolescents and become interested in adopting the PYD perspective, psychometrically sound measurement tools will be needed to assess adolescents' positive attributes. The PYD-SF and PYD-VSF provide such measurement instruments for the Five Cs model of PYD. Use of these measures can be easily implemented by researchers and practitioners alike. Accordingly, these scales may be able to make important contributions to science and practice in the field of youth development.

Acknowledgements

This research was supported by grants from the Thrive Foundation for Youth, the John Templeton Foundation, and the National 4-H Council.

Author Affiliations

G. John Geldhof, Eliot-Pearson Department of Child Development, Tufts University; Edmond P. Bowers, Eliot-Pearson Department of Child Development, Tufts University; Michelle J. Boyd, Eliot-Pearson Department of Child Development, Tufts University; Megan Kiely Mueller, Eliot-Pearson Department of Child Development, Tufts University; Christopher M. Napolitano, Eliot-Pearson Department of Child Development, Tufts University; Kristina L. Schmid, Eliot-Pearson Department of Child Development, Tufts University; Jacqueline V. Lerner, Department of Counseling, Developmental, and Educational Psychology, Boston College; Richard M. Lerner, Eliot-Pearson Department of Child Development, Tufts University.

G. John Geldhof, Institute for Applied Research in Youth Development, Lincoln Filene Center, Medford, MA 02155. E-mail: John.Geldhof@Tufts.edu.

References

Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K. K., Burns, K., Washburn, I. J., & Durlak, J. (2009). Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawaii. *American Journal of Public Health, 99*(8), 1438-1445.

Benson, P. L., Scales, P. C., Hamilton, S.F., & Sesma, A. (2006). Positive youth development: Theory, research and application. In R. M. Lerner (Vol. Ed.), *Theoretical models of human development*. Volume 1 of *Handbook of Child Psychology* (6th ed. Pp. 894-941), Editors-in-chief: W. Damon & R. M. Lerner. New York: Wiley.

Bowers, E. P., Li, Y., Kiely, M. K., Brittian, A., Lerner, J. V., & Lerner, R. M. (2010). The Five Cs model of positive youth development: A longitudinal analysis of confirmatory factor structure and measurement invariance. *Journal of Youth and Adolescence, 39*, 720-735.

Bradshaw, C.P. & Guerra, N.G. (2008). Future directions for research on core competencies. In N. G. Guerra, & C. P. Bradshaw (Guest Editors), Core competencies to prevent problem behaviors and promote positive youth development. *New Directions for Child and Adolescent Development, 122*, 89-92.Cattell, R. B. (1961). Theory of situational, instrumental, second order, and refraction factors in personality structure research, *Psychological Bulletin, 58,* 160-174.

Cheung, G. W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.

Duerden, M.D., Witt, P.A., Fernandez, M., Bryant, M.J., & Theriault, D. (2012). Measuring life skills: Standardizing the assessment of youth develop[ment indicators. *Journal of Youth Development, 7*(1), 99-117.

Dukakis, K., London, R. A., McLaughlin, M., & Williamson, D. (2009). *Positive youth development: Individual, setting and system level indicators.* (Issue brief: Positive youth development indicators). Stanford, CA: John W. Gardner Center for Youth and Their Communities.

Eccles, J., & Gootman, J. (Eds). (2002). *Community programs to promote youth development*. Washington, D.C.: National Academy Press.

Erikson, E.H. (1968). *Identity: youth and crisis.* Oxford, England: Norton & Co.

Freud, A. (1969). Adolescence as a developmental disturbance. In G. Caplan & S. Lebovici (Eds.), *Adolescence* (pp. 5-10). New York: Basic Books.

Gestsdóttir, S., & Lerner, R. M. (2007). Intentional self regulation and positive youth development in early adolescence: Findings from the 4-H Study of Positive Youth Development. *Developmental Psychology, 43,* 508-521.

Ginsburg, K. R., & Jablow, M. M. (2011). *Building resilience* (2nd. Ed.). Elk Grove Village, IL: American Academy of Pediatrics.

Hall, G. S. (1904). *Adolescence: Its psychology and its relations to physiology, anthropology, sociology, sex, crime, religion, and education* (Vols. 1 & 2). New York: Appleton.

Harter, S. (1983). *Supplementary description of the Self-Perception Profile for Children: Revision of the Perceived Competence Scale for Children*. Unpublished manuscript, University of Denver.

Harter, S. (1988). *Manual for the Self-Perception Profile for Adolescents.* Denver, CO: University of Denver.

Haskins, D. (2010). Coaching the whole child: Positive development through sports. United Kingdom: The National Coaching Foundation.

Heck, K. E., & Subramaniam, A., (2009). *Youth development Frameworks*. [Monograph]. Davis, CA: 4-H Center for Youth Development, University of California.

Jeličić, H., Bobek, D., Phelps, E., D., Lerner, J. V., Lerner, R. M. (2007). Using positive youth development to predict contribution and risk behaviors in early adolescence: Findings from the first two waves of the 4-H Study of Positive Youth Development. *International Journal of Behavioral Development*, *31*(3), 263-273.

King, P. E., Dowling, E. M., Mueller, R. A., White, K., Schultz, W., Osborn, P., Dickerson, E., Bobek, D. L., Lerner, R. M., Benson, P. L., & Scales, P. C. (2005). Thriving in Adolescence: The voices of youth-serving practitioners, parents, and early and late adolescents. *Journal of Early Adolescence, 25(1),* 94-112.

Kurtines, W. M., Ferrer-Wreder, L., Berman, S. L., Cass Lorente, C., Briones, E., Montgomery, M. J., Albrecht, R., Garcia A., & Arrufat, O. (2008b). Promoting positive youth development: The Miami Youth Development Project (YDP). *Journal of Adolescent Research, 23*, 256-267.

Leffert, N., Benson, P. L., Scales, P. C., Sharma, A. R., Drake, D. R., & Blyth, D. A. (1998). Developmental assets: Measurement and prediction of risk behaviors among adolescents. *Applied Developmental Science*, *2*(4), 209-230.

Lerner, J. V., Bowers, E. P., Minor, K., Boyd, M. J., Mueller, M. K., Schmid, K. L., Napolitano, C. M., Lewin-Bizan, S., & Lerner, R. M. (In press). Positive youth development: Processes, philosophies, and programs. In R. M. Lerner, M. A., Easterbrooks, & J. Mistry (Eds.), *Handbook of Psychology, Volume 6: Developmental Psychology* (2nd edition). Editor-in-chief: I. B. Weiner. Hoboken, NJ: Wiley.

Lerner, J. V., Phelps, E., Forman, Y., & Bowers, E. P. (2009). Positive youth development. In R. M. Lerner, L. Steinberg, R. M. Lerner, L. Steinberg (Eds.), *Handbook of adolescent psychology, Vol 1: Individual bases of adolescent development (3rd ed.)* (pp. 524-558). Hoboken, NJ US: John Wiley & Sons Inc.

Lerner, R. M. (2004). *Liberty: Thriving and civic engagement among American youth*. Thousand Oaks, CA: Sage.

Lerner, R. M., Lerner, J. V., Almerigi, J., Theokas, C., Phelps, E., Gestsdóttir, S. Naudeau, S., Jeličić, H., Alberts, A. E., 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.

Lerner, R. M., Lerner, J. V., Lewin-Bizan, S., Bowers, E. P., Boyd, M., Mueller, M., Schmid, K., Napolitano, C. (2011). Positive youth development: Processes, programs, and problematics. *Journal of Youth Development, 6(3),* 40-64.

Lewin-Bizan, S., Lynch, A. D., Fay, K., Schmid, K., McPherran, C., Lerner, J. V., & Lerner, R. M. (2010). Trajectories of positive and negative behaviors from early- to middle-adolescence. *Journal of Youth and Adolescence, 39*(7), 751-763.

Li, Y., Lerner, J. V., & Lerner, R. M. (2010). Personal and ecological assets and academic competence in early adolescence: The mediating role of school engagement. *Journal of Youth and Adolescence,* doi:10.1007/s10964-010-9535-4.

Lippman, L. H, Moore, K. A., & McIntosh, H. (2011). Positive indicators of child well-being: A conceptual framework, measures and methodological issues. *Applied Research Quarterly, 6*,

425-449. Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9,* 151-173.

Monitoring the Future (2000). National survey on drug use, 1975-2000. Bethesda, MD: National Institute on Drug Abuse.

Moore, K. A., Lippman, L., & Brown, B. (2004). Indicators of child well-being: The promise for Positive Youth Development. *Annals of the American Academy of Political and Social Science. Special Issue: Positive Development: Realizing the Potential of Youth, 591*, 125-145.

O'Hare W.P. (2012). Development of the child indicator movement in the United States. *Child Development Perspectives, 6(*1), 79-84.

Phelps, E., Balsano, A., Fay, K., Peltz, J., Zimmerman, S., Lerner, R., M., & Lerner, J. V. (2007). Nuances in early adolescent development trajectories of positive and of problematic/risk behaviors: Findings from the 4-H Study of Positive Youth Development. *Child and Adolescent Clinics of North America*, *16*(2), 473–496.

Phelps, E., Zimmerman, S., Warren, A. E. A., Jeličič, H., von Eye, A., & Lerner, R. M. (2009). The structure and developmental course of Positive Youth Development (PYD) in early adolescence: Implications for theory and practice. *Journal of Applied Developmental Psychology, 30*(5), 571-584.

Pittman, K., Irby, M., & Ferber, T. (2001). Unfinished business: Further reflections on a decade of promoting youth development. In P. L. Benson & K. J. Pittman (Eds.), *Trends in youth development: Visions, realities and challenges* (pp. 4-50). Norwell, MA: Kluwer.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385-401.

Robinson, A., Esters, L.T., McKee, R., Tucker, M.A., & Dotterer, A. (2012). An exploratory study of the Five Cs model of positive youth development among Indiana 4-H youth. *Journal of Youth Development, 7*(1), 82-98.

Rossman, B. R. (1992). School-age children's perceptions of coping with distress: Strategies for emotion regulation and the moderation of adjustment. *Journal of Child Psychology and Psychiatry, 33*(8), 1373-1397.

Roth, J. L., & Brooks-Gunn, J. (2003a). What is a youth development program? Identification and defining principles. In. F. Jacobs, D. Wertlieb, & R. M. Lerner (Eds.), *Enhancing the life chances of youth and families: Public service systems and public policy perspectives: Vol. 2 Handbook of applied developmental science: Promoting positive child, adolescent, and family development through research, policies, and programs* (pp. 197-223). Thousand Oaks, CA: Sage.

Roth, J. L., & Brooks-Gunn, J. (2003b). What exactly is a youth development program? Answers from research and practice. *Applied Developmental Science*, *7*, 94-111.

Ryan, N. M. (1989). Identification of children's coping strategies from the school-agers' perspective. *Research in Nursing Health, 12*, 111–122.

Schmid, K. L., Phelps, E., & Lerner, R. M. (2011). Constructing positive futures: Modeling the relationship between adolescents' hopeful future expectations and intentional self-regulation in predicting positive youth development. *The Journal of Adolescence, 34*, 1127-1135.

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18,* 210-222.

von Eye, A., Martel, M. M., Lerner, R. M., Lerner, J. V., Bowers, E. P. (2011). Integrating theory and method in the study of Positive Youth Development: The sample case of gender-specificity and longitudinal stability of the dimensions of intention self-regulation (Selection, Optimization, and Compensation). In R. M. Lerner, J. V. Lerner, & J. B. Benson (Eds.). *Advances in Child Behavior and Development: Positive Youth Development: Research and Applications for Promoting Thriving in Adolescence* (pp. 352-403). Amsterdam: Elsevier.

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods, 8,* 16-37.

Zimmer-Gembeck, M. J., & Skinner, E. A. (2011). Review: The development of coping across childhood and adolescence: An integrative review and critique of research. *International Journal of Behavioral Development, 35*(1), 1-17.

Table 1
Percent missing data for each construct, by Grade

	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Contribution	39.52	21.65	6.91	5.69	10.25	5.71	10.65	5.24
CES-D	14.99	8.38	6.65	4.50	8.63	5.16	9.66	6.83
Risk Behavior	7.40	3.49	14.53	4.50	18.23	9.58	9.21	6.35
PYD Items								
Minimum	8.85	3.84	0.52	0.87	6.04	0.95	6.29	0.95
Maximum	55.27	24.65	18.72	11.45	42.29	11.24	10.45	9.21
Avg. % Missing	22.90	9.77	10.36	5.72	20.51	4.13	7.75	3.55

Table 2

Standardized Factor Loadings from the Strong Invariance Model of the PYD-SF

	Target		PYD		Target		PYD		Target		PYD	
Grade	6	7	6	7	7	8	7	8	9	12	9	12
Competence												
Reliability (ω)	0.73	0.74			0.72	0.74			0.73	0..76		
HART07	0.41	0.41	0.33	0.36	0.35	0.36	0.35	0.36	0.36	0.37	0.23	0.21
HART25	0.42	0.43	0.40	0.44	0.32	0.32	0.45	0.46	0.27	0.30	0.42	0.40
HART08	0.54	0.54	0.04	0.05	0.51	0.55	0.12	0.13	0.60	0.59	0.09	0.08
HART32	0.59	0.59	-0.10	-0.11	0.58	0.62	0.01	0.01	0.67	0.70	-0.05	-0.04
HART15	0.49	0.48	0.03	0.04	0.54	0.52	0.05	0.05	0.55	0.57	-0.13	-0.11
HART21	0.54	0.54	-0.09	-0.09	0.53	0.52	0.00	0.00	0.50	0.53	-0.11	-0.10
HART17	0.33	0.32			0.23	0.24			0.22	0.28		
Confidence												
Reliability (ω)	0.75	0.77			0.77	0.83			0.81	0.82		
HART18	0.56	0.58	0.29	0.29	0.52	0.54	0.42	0.42	0.59	0.61	0.32	0.31
HART30	0.60	0.65	0.22	0.23	0.59	0.65	0.34	0.36	0.61	0.65	0.30	0.30
ABME10	0.47	0.49	0.32	0.33	0.37	0.44	0.44	0.49	0.51	0.50	0.38	0.34
ABME13	0.32	0.35	0.26	0.28	0.24	0.27	0.34	0.36	0.37	0.37	0.34	0.31
HART34	0.65	0.69	-0.14	-0.15	0.67	0.73	0.07	0.07	0.64	0.66	-0.01	-0.01
HART10	0.46	0.51	0.10	0.11	0.58	0.72	0.14	0.16	0.70	0.72	0.07	0.07
Character												
Reliability (ω)	0.82	0.78			0.79	0.80			0.81	0.80		
ABME21	0.76	0.67	0.20	0.22	0.62	0.61	0.33	0.36	0.55	0.61	0.46	0.42
ABME22	0.76	0.67	0.19	0.21	0.68	0.66	0.31	0.34	0.58	0.59	0.46	0.40
ABME40	0.46	0.39	0.08	0.09	0.44	0.42	0.14	0.15	0.39	0.36	0.18	0.14
ABME41	0.50	0.43	0.13	0.14	0.45	0.45	0.19	0.21	0.42	0.42	0.29	0.24
HART17	0.06	0.05	0.49	0.51	0.06	0.05	0.48	0.49	-0.11	-0.14	0.53	0.56
HART29	0.07	0.06	0.55	0.58	0.04	0.04	0.53	0.52	-0.15	-0.17	0.65	0.63
ABME26	0.62	0.54	0.28	0.30	0.44	0.43	0.37	0.41	0.40	0.46	0.46	0.45
ABME29	0.64	0.55	0.31	0.33	0.45	0.42	0.40	0.42	0.36	0.42	0.46	0.46
Caring												
Reliability (ω)	0.90	0.88			0.87	0.88			0.88	0.87		
CARE2	0.68	0.65	0.18	0.20	0.57	0.57	0.33	0.37	0.53	0.53	0.40	0.40
CARE4	0.68	0.62	0.19	0.21	0.55	0.54	0.31	0.34	0.46	0.43	0.38	0.35
CARE6	0.69	0.62	0.12	0.13	0.56	0.54	0.22	0.24	0.55	0.51	0.29	0.26
CARE7	0.83	0.79	0.19	0.21	0.75	0.76	0.33	0.37	0.77	0.77	0.37	0.37
CARE8	0.82	0.75	0.19	0.20	0.74	0.76	0.29	0.34	0.75	0.74	0.36	0.35
CARE9	0.85	0.81	0.18	0.20	0.79	0.79	0.32	0.36	0.79	0.79	0.37	0.36
Connection												
Reliability (ω)	0.79	0.80			0.82	0.83			0.82	0.80		
FAM4	0.51	0.49	0.27	0.27	0.35	0.34	0.45	0.49	0.42	0.39	0.34	0.28
FAM5	0.54	0.48	0.32	0.30	0.32	0.31	0.46	0.49	0.43	0.40	0.38	0.31
NEIGH3	0.59	0.62	0.16	0.18	0.59	0.58	0.32	0.36	0.54	0.52	0.31	0.26
NEIGH4	0.55	0.61	0.14	0.16	0.60	0.57	0.28	0.30	0.57	0.55	0.29	0.25
CLAS05	0.49	0.49	0.27	0.28	0.39	0.38	0.45	0.49	0.43	0.44	0.41	0.37
CLAS10	0.43	0.44	0.25	0.27	0.33	0.33	0.41	0.46	0.39	0.40	0.39	0.35
PEER6	0.26	0.29	0.24	0.28	0.24	0.22	0.40	0.41	0.29	0.31	0.37	0.34
PEER7	0.32	0.35	0.24	0.27	0.30	0.28	0.39	0.42	0.35	0.36	0.35	0.32

Item labels represent the Early Adolescent item labels; Results for Sympathy are not shown.

Table 3

Stability and Latent Correlations among the 5 Cs - Strong Invariance PYD-SF Models

	Grade 6	Grade 7	Grade 7	Grade 8	Grade 9	Grade 12
Competence with						
Stability		0.67***		0.70***		0.69***
Confidence	0.84***	0.83***	0.80***	0.78***	0.87***	0.80***
Character	0.22***	0.21***	0.13*	0.15*	0.30***	0.23***
Caring	0.12**	0.11**	0.02	0.08	0.11	0.18**
Connection	0.53***	0.51***	0.43***	0.43***	0.62***	0.60***
Confidence with						
Stability		0.63***		0.57***		0.54***
Character	0.29***	0.24***	0.08	0.09	0.18**	0.11
Caring	0.10*	0.08*	-0.05	-0.03	-0.05	-0.01
Connection	0.60***	0.57***	0.40***	0.38***	0.47***	0.60***
Character with						
Stability		0.51***		0.72***		0.37**
Caring	0.51***	0.63***	0.59***	0.58***	0.51***	0.52***
Connection	0.61***	0.62***	0.54***	0.52***	0.45***	0.26**
Caring with						
Stability		0.65***		0.51***		0.43***
Connection	0.31***	0.47***	0.39***	0.32***	0.19**	0.11
Connection with						
Stability		0.83***		0.86***		0.76***
PYD Stability		0.72***		0.78***		0.71***

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 4

Latent Correlations Between PYD-SF and Key Outcomes - Strong Invariance Models

	Grade 6	Grade 7	Grade 7	Grade 8	Grade 9	Grade 12
PYD with						
Contribution	0.25***	0.23***	0.35***	0.41***	0.50***	0.51***
Depression	-0.40***	-0.37***	-0.47***	-0.53***	-0.39***	-0.32***
Risk	-0.46***	-0.42***	-0.47***	-0.53***	-0.56***	-0.56***
Competence with						
Contribution	0.36***	0.33***	0.26***	0.27***	0.25***	0.30***
Depression	-0.34***	-0.33***	-0.27***	-0.30***	-0.37***	-0.50***
Risk	0.02	-0.03	0.04	0.11*	0.06	0.02
Confidence with						
Contribution	0.31***	0.21***	0.11*	0.09*	0.11*	0.14**
Depression	-0.43***	-0.46***	-0.37***	-0.35***	-0.45***	-0.57***
Risk	-0.02	-0.05	0.06	0.18***	0.20***	0.08
Character with						
Contribution	0.49***	0.60***	0.55***	0.56***	0.42***	0.37***
Depression	-0.06	-0.01	0.09*	0.14**	0.12*	0.07
Risk	-0.08*	-0.19***	-0.13***	-0.07	0.03	0.16**
Caring with						
Contribution	0.28***	0.43***	0.37***	0.35***	0.20***	0.21***
Depression	-0.04	0.06*	0.15***	0.20***	0.12**	0.02
Risk	-0.10**	-0.10***	-0.04	-0.02	0.05	0.07
Connection with						
Contribution	0.50***	0.58***	0.53***	0.48***	0.43***	0.39***
Depression	-0.33***	-0.32***	-0.18***	-0.08	-0.32***	-0.52***
Risk	-0.11**	-0.16***	-0.03	0.00	-0.09	0.05
Contribution with						
Depression	-0.19***	-0.12***	-0.12***	-0.15***	-0.16***	-0.21***
Risk	-0.16***	-0.18***	-0.19***	-0.20***	-0.25***	-0.21***
Depression with						
Risk	0.28***	0.27***	0.28***	0.21***	0.24***	0.20***

* $p < .05$ ** $p < .01$ *** $p < .001$