

Longitudinal Analysis of a Very Short Measure of Positive Youth Development

G. John Geldhof · Edmond P. Bowers ·
Megan Kiely Mueller · Christopher M. Napolitano ·
Kristina Schmid Callina · Richard M. Lerner

Received: 12 September 2013 / Accepted: 15 January 2014 / Published online: 21 February 2014
© Springer Science+Business Media New York 2014

Abstract As developmental scientists cease to perceive adolescence as a period of inevitable turmoil and adopt the Positive Youth Development (PYD) perspective, psychometrically sound measurement tools will be needed to assess adolescents' positive attributes. In this article we examine the longitudinal stability of the very short version of the PYD scale developed as part of the 4-H Study of PYD. Using a sample of 7,071 adolescents (60 % female) followed between Grades 5 and 12, our results suggest general stability of PYD across adolescence, both in terms of mean levels and rank-order stability. We also show that both a global measure of PYD and the individual Five Cs of PYD consistently correlate with important criterion measures (i.e., contribution, depressive symptoms, and problem behaviors) in expected ways. Although our results suggest weak relationships among our three criteria, we especially note that across adolescence PYD becomes more strongly correlated with contribution but less strongly

correlated with depressive symptoms, and that confidence becomes more strongly related to depressive symptoms. We discuss implications for use of the present PYD measure in youth development programs.

Keywords Positive youth development · Short form · Confirmatory factor analysis (CFA) · Longitudinal measurement invariance · Adolescence

Introduction

Historically, research on adolescent development was framed by a “deficit perspective” that described the second decade of life as a period of “storm and stress” (Hall 1904), developmental disturbance (Freud 1969), or crisis (Erikson 1968). This perspective treats adolescents as problems to be managed (Roth and Brooks-Gunn 2003b), with positive development during this period of life indexed as the absence of, or decreases in, problem behaviors. The continued influence of the deficit perspective on research aims, policy, and practice is reflected in the contemporary prevalence of measures of risk and problem behaviors used by program and service organizations to assess youth functioning (Benson 2006). Looking at the field of youth development in the early years of this century, it appeared it was easier to determine what youth should avoid (e.g., violence, drugs mental health problems) than to identify youth characteristics and experiences that were indicators of thriving, positive development, or well-being (Moore et al. 2004).

Partly in response to this focus on the problems and deficits among young people, a new approach to adolescent development emerged over the past two decades—the positive youth development (PYD) perspective (Lerner

Electronic supplementary material The online version of this article (doi:10.1007/s10964-014-0093-z) contains supplementary material, which is available to authorized users.

G. J. Geldhof (✉)
College of Public Health and Human Sciences, 410 Waldo Hall,
Oregon State University, Corvallis, OR 97331, USA
e-mail: John.Geldhof@oregonstate.edu

E. P. Bowers · K. S. Callina · R. M. Lerner
Tufts University, Medford, MA, USA
e-mail: iaryd.pubs@gmail.com

M. K. Mueller
Cummings School of Veterinary Medicine, Tufts University,
Grafton, MA, USA

C. M. Napolitano
University of Zurich, Zurich, Switzerland

et al. 2009, 2010, 2011, 2013). The PYD perspective moves beyond a deficit view of youth and, instead, emphasizes that all youth have personal strengths that can be developed. Thus, the PYD perspective emphasizes that researchers should find ways to optimize positive functioning rather than simply focusing on the prevention of negative developmental outcomes.

Several theoretical frameworks now approach development from the PYD perspective (for a review, see Lerner et al. 2013). As these models become more popular it is important that they are empirically valid, can be widely applied, and include constructs that are specific and measurable. However, these models are just beginning to be tested. Recent work has attempted to evaluate youth development frameworks (Heck and Subramaniam 2009) and indicators of PYD (Dukakis et al. 2009), but further investigation of suitable models is needed. In this article we discuss one popular model of PYD, the Lerner and Lerner 5 Cs Model, Lerner et al. 2013) and extend findings that support its empirical validity across adolescence.

The Five Cs Model of PYD

The purpose of this article 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; Jellicic et al. 2007; Lerner et al. 2005; Phelps et al. 2009). This model operationalizes positive development through the assessment of five “Cs”—Competence, Confidence, Character, Connection, and Caring—found to be important for youth development according to the experiences of practitioners and based on extensive reviews of the adolescent development literature (Eccles and Gootman 2002; Lerner 2004; Roth and Brooks-Gunn 2003a, b). We present definitions of these Cs in Table 1, but note that all Five Cs are fully integrated such that PYD requires healthy development in all five domains (see also Dukakis et al. 2009).

The Five Cs were linked to the positive outcomes of youth development programs reported by Roth and 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). In turn, when young people manifest these Five Cs over the course of adolescence, they are more likely to be on a life trajectory marked by mutually-beneficial person ↔ context relationships 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 depressive symptoms. That is, as evidence for positive behavior increases, the PYD perspective hypothesizes that

Table 1 Definitions of the Five Cs of positive youth development

C	Definition
Competence	Positive view of one’s actions in domain specific areas including social, academic, cognitive, and vocational. Social competence pertains to interpersonal skills (e.g., conflict resolution). Cognitive competence pertains to cognitive abilities (e.g., decision making). School grades, attendance, and test scores are part of academic competence. Vocational competence involves work habits and career choice explorations, including entrepreneurship
Confidence	An internal sense of overall positive self-worth and self-efficacy; one’s global self-regard, as opposed to domain specific beliefs
Connection	Positive bonds with people and institutions that are reflected in bidirectional exchanges between the individual and peers, family, school, and community in which both parties contribute to the relationship
Character	Respect for societal and cultural rules, possession of standards for correct behaviors, a sense of right and wrong (morality), and integrity
Caring	A sense of sympathy and empathy for others

Derived from Lerner et al. (2005) and Roth and Brooks-Gunn (2003a)

there will be fewer indications of problematic behaviors (e.g., Benson et al. 2004; Pittman et al. 2001). Although recent research supports a general inverse relationship between PYD and risk and problem behaviors, these findings also indicate a more complex pattern of positive and negative developmental trajectories; these pathways are not simply inversely related (Lewin-Bizan et al. 2010; Phelps et al. 2007). Nevertheless, PYD is associated across development with positive indicators such as contribution, school engagement, successful intentional self-regulation, and hope (Lerner et al. 2013).

Structure and Validity of PYD as an Aggregate Score

The 4-H Study of PYD (e.g., Lerner et al. 2005, 2009, 2010, 2011) is a longitudinal study spanning from Grade 5 to, at this writing, Grade 12, and has provided the primary empirical support for the Five Cs Model. Using data from the first wave (Grade 5) of the 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 Cs. In a subsequent study, structural equation models were constructed to test the validity of the Five Cs model (Jellicic 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 latent construct. This structure suggests that the Five Cs can be summarized by a single indicator of positive development. This PYD construct significantly predicted criterion indices of positive functioning measured 1 year later, including community contributions, depressive

symptoms, and an aggregate measure of delinquency and substance use (Jeličić et al. 2007).

More recently, Phelps et al. (2009) and Bowers et al. (2010) extended the above findings by assessing the structure and development of PYD from Grade 5 to 7, and from Grade 8 to 10 of the 4-H Study, respectively. While the fitted models required estimating a large number of residual covariances among indicators, the results of these studies supported previous findings and indicated that the higher-order structure of PYD continued to be robust across adolescence. The global measure of PYD also displayed high stability across adolescence, especially between Grades 5 and 7 (i.e., stability correlations of approximately .88), although the scales relevant to measuring the Five Cs were slightly different for two of the Cs during middle adolescence than in early adolescence. That is, reflective of developmental change, athletic competence was no longer a relevant indicator of competence during middle adolescence; however, physical appearance significantly loaded on the latent construct of confidence.

A series of dual-trajectory analyses presented by Lewin-Bizan et al. (2010) similarly extended Jelicic et al.’s (2007) findings that suggested relationships between a global measure of PYD (taken as the average of all Five Cs of PYD) and important criterion variables. Using latent class growth analysis, Lewin-Bizan et al. (2010) estimated growth trajectories for PYD, problem behaviors, and depressive symptoms using data from Grades 5 to 10 of the

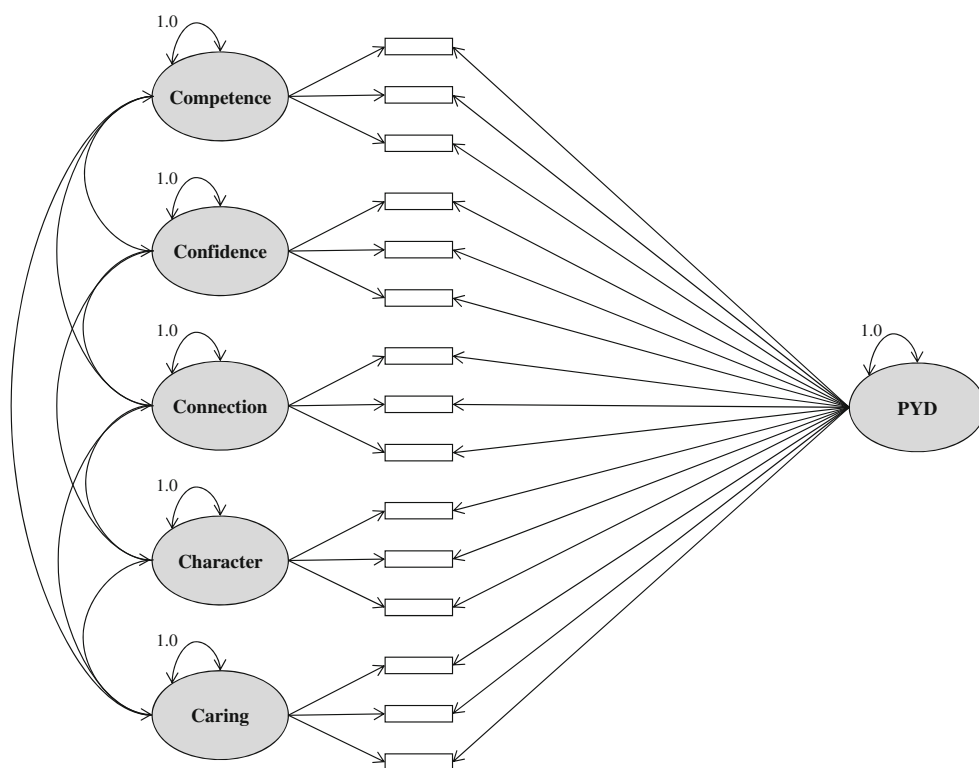
4-H Study of PYD. By cross-tabulating trajectory membership, their findings suggested that participants on more optimal trajectories of PYD tended to be on more optimal trajectories of both criteria. That is, these results suggest longitudinally stable negative relationships between PYD as indexed by the Five Cs model and indicators of problem behavior development.

A Bifactor Model of the Five Cs of PYD

Despite robust evidence supporting the validity of an aggregate/higher-order measure of PYD, recent evidence suggests that the individual Cs may provide a more nuanced picture of PYD than that provided by a global measure of PYD. For instance, Geldhof and colleagues (in press; see also von Eye et al. 2011) describe a bifactor model of PYD in which all indicators of PYD load onto two constructs: a global measure that aggregates across all Cs and one of five specific constructs that represent the variance in each C after controlling for global PYD (see Fig. 1). In such a bifactor model the global PYD construct provides a rough, heuristic estimate of each adolescent’s positive functioning that aggregates across all PYD indicators. The residual C constructs then represent the covariance among the items within each C that is not related to global PYD.

The bifactor model of PYD has both empirical and theoretical advantages over the previously implemented higher-order model. Empirically, the bifactor model

Fig. 1 A hypothetical bifactor model of PYD



provides statistically better fit to the 4-H data than does a higher-order model (see Geldhof et al., in press). Theoretically, a bifactor model also maps more directly onto the 5 Cs model of PYD discussed by Lerner, Lerner, and their colleagues (e.g., Lerner et al. 2005). A higher-order model of PYD necessarily assumes that each participant has a set (albeit latent) level of PYD that causes their level of each C. These latent C scores then cause individuals to score in certain ways on each indicator in the model. A bifactor model instead treats global PYD as one of multiple sources of item true score variance. Because PYD cuts across items, this global measure is akin to the variance of group means in multilevel modeling. In addition, the bifactor model allows indicators to covary systematically with each other in ways related to each unique C, thus suggesting that each item has two sources of true-score variance. Accordingly, a bifactor model also allows both the global PYD scale *and* the residual C constructs to covary independently with important criterion measures (e.g., contribution, depressive symptoms).

The Current Study

In the present article we extend previous research by examining the bifactor structure of PYD as indexed by the Five Cs of PYD across all eight waves of the 4-H study. We use the very short (17-item) form of the PYD measure developed by Geldhof and colleagues (in press). This use is predicated on two considerations. First, the very short version of this measure allows us to test the structure of the bifactor model across all waves of the 4-H Study; if we can find evidence of the validity of the measure with so few items in the scale, then we would have confidence that measures with larger item sets could also profit from use of bifactor modeling approaches across the span of adolescence. This potential use of the very short measure relates to our second reason for using this tool. Community-based youth development programs seek tools that provide valid indicators of youth development for use in program evaluations (Dukakis et al. 2009). The present research can ascertain if the very short form constitutes such a tool for programs seeking to promote PYD through their curricula. Accordingly, we test the measure's structure and factorial invariance across all eight waves, while also presenting findings that speak directly to the stability and criterion validity of the overarching measure of PYD as well as for the residual C constructs that represent systematic item variance not directly related to the overarching PYD measure. Specifically, we examine the mean-level and correlational stability for all PYD constructs as well as examining the correlations between these constructs and the criterion variables of contribution, depressive

symptoms, and problem behaviors across 8 years of adolescence.

Method

Full details of the 4-H Study of PYD have been presented elsewhere (Lerner et al. 2005, 2009, 2010, 2011). Therefore, we present here only the features of the methods relevant to the present research, which includes data from all waves of the 4-H Study. A discussion of the overall method of the 4-H Study is provided in the introductory article of this special issue (Bowers et al., in press).

Participants

We analyzed data from 7,071 adolescents (60 % female) who participated in the 4-H Study. As Table 2 shows, the mean age of participants was 10.94 ($SD = .42$) in the Grade 5 assessment and 17.71 ($SD = .76$) in Grade 12. With respect to race/ethnicity, the sample was 65.8 % White; 7.6 % Black; 9.4 % Latino; and 14.4 % other (including Asian, Native American, Multiethnic/multiracial, or "other"). 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). Our sample's demographic characteristics were not completely stable across all waves of the study, however, with female and White participants over-represented in later waves of the data.

In addition, participants' parents provided data regarding the socioeconomic status of their families. In Grade 5, 20 % of mothers had attended or completed high school; 24.8 % had 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 ranged to \$23,401 ($SD = \$13,798$) in Grade 12 (also see Table 2).

Procedure

In Waves 1–3 of the 4-H Study, data collection from youth was conducted by trained study staff or, at more distant locations, hired assistants. A detailed protocol was used to ensure that data collection was administered uniformly and to ensure the return of all study materials. After Wave 1, youth who were absent on the day of the survey or were from schools or programs that did not allow on-site testing were contacted by e-mail, mail, or phone, and were asked to complete and return the survey to us. Beginning in Wave 5, youth completed the survey online unless they requested a paper survey. Parents completed online or paper surveys. Paper surveys were delivered to their homes by their

Table 2 Participant demographics in the 4-H Study of PYD, by Grade

	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Age <i>M(SD)</i>	10.94 (0.42)	12.01 (0.43)	13.00 (0.47)	14.02 (0.53)	14.98 (0.57)	15.82 (0.70)	16.83 (0.76)	17.71 (0.76)
Mother's education (%)								
High school or less	20.8	18.2	13.2	6.4	2.8	1.2	1.0	1.6
Some college	24.8	22.1	19.4	9.9	9.1	4.2	4.3	1.4
BA or higher	18.6	17.8	18.2	10.3	8.5	6.8	4.4	3.2
% Missing	35.8	42.0	49.3	73.4	79.6	87.8	90.3	93.8
Mean per capita income (<i>SD</i>)	13,656.86 (8,348.46)	13,635.81 (8,621.05)	16,553.42 (10,631.93)	19,137.40 (13,216.27)	19,981.29 (12,938.41)	24,331.31 (18,664.92)	24,981.29 (17,316.31)	23,401.44 (13,798.49)

children or through the mail (in the latter case, return postage was provided).

Measures

Positive Youth Development

We operationalized PYD using the very short measure of the Five Cs of PYD (PYD-VSF) discussed by Geldhof and colleagues (in press). The Five Cs model identifies PYD as consisting of the Five Cs discussed in Table 1. Our measure of PYD drew items from several primary sources, including: the Search Institute's Profiles of Student Life-Attitudes and Behaviors (PSL-AB) scale (Benson et al. 1998; Leffert et al. 1998); the Self-Perception Profile for Children (Harter 1983); the Self-Perception Profile for Adolescents (Harter 1986, 1988); Teen Assessment Project (TAP) Survey Question Bank (Small and Rodgers 1995); Eisenberg Sympathy Scale (Eisenberg et al. 1996); and the Empathic Concern Subscale of the Interpersonal Reactivity Index (IRI; Davis 1980). At Grades 5–7, we used the Self-Perception Profile for Children (Harter 1983). Beginning at Grade 8 and continuing through Grade 12, we used the Self-Perception Profile for Adolescents (Harter 1986, 1988); we used the 1986 version for Grade 8 and the 1988 version for Grades 9–12). While previous research has suggested that the structure of PYD changes across adolescence (e.g., Bowers et al. 2010), we included all scales used to measure PYD in any wave of the 4-H study in our analyses to maintain longitudinal consistency.

As discussed by Geldhof and colleagues (in press), the PYD-VSF measures *Competence* using three items, one item representing Academic Competence, Social Competence, and Physical Competence, respectively. All Competence items asked participants to select the type of person they were more like between two choices (e.g., "Some teenagers do very well at their class work, BUT Other teenagers don't do very well at their class work.") and then to decide if it was "really true" or "sort of true" for him/her. The PYD-VSF similarly measures *Confidence* using three items, one representing Self Worth, Positive Identity, and Physical Appearance, respectively. The Self-Worth and Physical Appearance items followed the same response format (e.g., "Some teenagers are happy with themselves most of the time, BUT Other teenagers are often not happy with themselves"), while the Positive Identity item was scored using a five-point Likert-type scale (e.g., "All in all, I am glad to be me").

The measure of *Character* in the PYD-VSF includes four items, with one representing Social Conscience, Values Diversity, Conduct Behavior, and Personal Values, with all items except the one indicating Conduct Behavior scored on a five-point Likert-type scale (e.g., "Helping to

make the world a better place to live in”). The Conduct Behavior item follows a similar response format to the Competence items above (“Some teenagers do things they know they shouldn’t do, BUT Other teenagers hardly ever do things they know they shouldn’t do”). Items representing *Caring* and *Connection* were all scaled on a five-point Likert-type scale, with *Caring* represented by three items representing empathic responding (e.g., “When I see someone being picked on, I feel sorry for them”), and the *Connection* scale containing four items that represent connection to participants’ families, neighborhoods, schools, and peers, respectively (e.g., “I have lots of good conversations with my parents,” and “Adults in my town or city make me feel important”).

Criteria

To ensure that the factor structure of PYD remained stable (i.e., did not change) in the presence of important outcomes, single-item composites for the following criteria scales (i.e., contribution, depressive symptomatology, and problem behaviors) were included in all confirmatory factor analyses. Single-item composites were chosen to ensure that model fit was not negatively impacted by minor misspecification of the factor structure of the outcome scales. Outcome measures were included for each wave of data and, while not a primary focus of this research, the substantive relationships between PYD and these outcomes are considered in the results below. Although these results may be somewhat attenuated by the fact that the outcome composites contain a degree of measurement error, we note that internal consistency reliability (e.g., Cronbach’s coefficient α ; Cronbach 1951; Guttman 1945; Kuder and Richardson 1937) estimates the proportion of true score versus error variance in a unit-weighted composite (de Gruijter and van der Kamp 2008). Thus, the fact that our criterion measures displayed acceptable reliability suggests that any bias due to measurement error is minimal.

Contribution At each grade of the 4-H Study, participants responded to twelve items which were weighted and summed to create a composite measure of contribution. These items were derived from existing instruments with known psychometric properties and used in large-scale studies of adolescents, including the Search Institute’s Profiles of Student Life-Attitudes and Behaviors (PSL-AB) scale (Benson et al. 1998; Leffert et al. 1998) and the TAP Survey Question Bank (Small and Rodgers 1995).

Contribution was comprised of two equally weighted subscales—ideology and actions—and each subscale included 6 items. The ideology subscale measured the extent to which contribution was an important facet of youth’s identity and future self. An example ideology subscale item

stated, “It is important to me to contribute to my community and society” with response options ranging from 1 = *strongly disagree* to 5 = *strongly agree*. An example item that assessed one’s future ideological orientation gauged the perceived chances that the young person would be involved in community service in the future, with a response format that ranged from 1 = *very low* to 5 = *very high*. The action subscale of contribution 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/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 7–12.

Depressive Symptoms We measured depressive symptomatology using the 20-item, self-report Center for Epidemiological Studies Depression scale (CES-D; Radloff 1977). This scale conceptualizes depressive symptomatology by several components: “depressed mood, feelings of guilt and worthlessness, feelings of helplessness and hopelessness, psychomotor retardation, loss of appetite, and sleep disturbance” (Radloff 1977, p. 386). 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, and higher scores were indicative of higher depressive symptomatology (i.e., greater frequency and number of symptoms of depression). Cronbach’s alphas for the CES-D scale ranged from 0.81 to 0.89 across Grades 5–12 in the present study.

Problem 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 survey (2000). Although somewhat distinct concepts, we aggregated our measures of substance use and delinquency into a composite measure of adolescent problem behaviors. Our use of a single aggregate of problem behaviors is consistent with prior research that has

examined data from the 4-H Study (e.g., Phelps et al. 2007).

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 6–12 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 7–12. The response options for all substance use items ranged from 0 = *never* to 3 = *regularly*.

Delinquency. We assessed Grade 5 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*.

For consistency, the delinquency items were rescaled so that their values ranged from 0 to 3. The averages for the substance use and delinquency items, respectively, were calculated and transformed to range from 0 to 15. A composite measure was then calculated by summing the averages of both subscales for a maximum score of 30, with higher scores indicating higher levels of problem behaviors. For the current study sample, the Cronbach's alphas for the problem behaviors scale were .65 at Grade 5 and .71 at Grade 6; the alphas ranged from 0.76 to 0.86 across Grades 7–12.

Analyses

We implemented a series of overlapping longitudinal CFA models that established the factorial invariance of the PYD-VSF across all waves of the 4-H study. As suggested by Geldhof and colleagues (in press; see also von Eye et al. 2011), we modeled the PYD-VSF using a bifactor model that specified each item as loading onto both an overarching PYD factor, as well as loading onto a factor representing that item's respective C independent of PYD (see Fig. 1). As noted previously, this bifactor model has several practical and theoretical advantages over the higher-

order model implemented in previous research. We tested weak (i.e., loading) and strong (i.e., intercept) invariance across time, meaning our analyses allowed us to examine longitudinal changes in the means, variances, and correlations among both a general measure of PYD, as well as each individual C. Thus, all latent means and variances presented in our final models can be interpreted as latent means and variances, relative to those in Grade 5. Due to the very large model size, we ran four sets of CFA models that considered Grades 5–7, Grades 7–9, Grades 9 and 10, and Grades 10–12, respectively. All analyses implemented FIML estimation which is robust in the presence of missing data when the MAR assumption is met. Missing data percentages are summarized in Table 3. Finally, we examined the latent correlations among PYD and each criterion variable.

Results

Validation of the PYD-VSF

We established longitudinal invariance for the first model (Grades 5–7) using standard procedures (see Little 2013), then established invariance for the three later models by fixing parameters in these models to equal the corresponding estimates in the strong invariance model that examined Grades 5–7. Establishing invariance across all waves scaled all latent means and variances in a comparable metric and allowed us to then explore developmental trends in the latent parameters across adolescence, despite the fact that different waves were examined in different models.

All initial CFAs for the PYD-VSF displayed acceptable model fit, and we established partial factorial invariance across all waves of data using the change in CFI criterion suggested by Cheung and Rensvold (2002; i.e., $\Delta\text{CFI} < .01$). We present fit for these models in Table 4. The finding of partial invariance indicates that some factor loadings and intercepts changed across grades and this finding can be interpreted to mean that the item-construct relationships, and the expected score of some items when their respective latent construct was zero, changed across waves of testing. Table 5 presents completely standardized factor loadings (i.e., from a model where both indicators and constructs were standardized) from the strong invariance models, and Table 6 presents raw-metric item intercepts from the same models. As these tables show, a majority of the changes occur in Grade 9, suggesting that the qualitative meaning of the latent constructs changes slightly as adolescents enter high school. Fewer than half of the factor loadings or intercepts changed within any given wave, however, meaning that it is still reasonable to

Table 3 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
Depressive symptoms	14.99	8.38	6.65	4.50	8.63	5.16	9.66	6.83
Problem behaviors	7.40	3.49	14.53	4.50	18.23	9.58	9.21	6.35
PYD items								
Minimum	9.03	9.43	1.10	0.95	6.69	1.22	6.40	1.43
Maximum	39.08	23.88	18.59	10.58	40.56	9.44	9.78	9.21
Ave. % missing	21.30	17.89	9.24	5.69	20.36	3.74	7.65	3.33

PYD positive youth development

Table 4 Model fit for PYD-VSF Models

	χ^2	df	RMSEA	CFI	TLI	$\Delta \chi^2$	Δ df	<i>p</i>	Δ CFI
<i>Grades 5, 6, and 7</i>									
CFA	1,985.02	1,194	.02 (.01, .02)	0.96	0.94				
Weak invariance	2,132.95	1,245	.02 (.01, .02)	0.95	0.94	147.93	51	<0.001	-0.005
Strong invariance	2,187.09	1,264	.02 (.01, .02)	0.95	0.94	54.15	19	<0.001	-0.002
Final	2,215.24	1,276	.02 (.01, .02)	0.95	0.94	28.15	12	0.005	-0.001
<i>Grades 7, 8, and 9</i>									
CFA	2,424.79	1,311	.02 (.02, .02)	0.95	0.93				
Weak invariance	2,682.44	1,401	.02 (.02, .02)	0.94	0.92	257.66	90	<0.001	-0.008
Strong invariance	2,919.64	1,437	.02 (.02, .02)	0.93	0.91	237.20	36	<0.001	-0.009
Final	2,942.74	1,451	.02 (.02, .02)	0.93	0.91	23.09	14	0.059	0.000
<i>Grades 9 and 10</i>									
CFA	1,115.71	562	.02 (.02, .02)	0.96	0.94				
Weak invariance	1,230.86	607	.02 (.02, .02)	0.95	0.93	115.15	45	<0.001	-0.006
Strong invariance	1,327.47	626	.02 (.02, .02)	0.94	0.93	96.61	19	<0.001	-0.006
Final	1,340.32	634	.02 (.02, .02)	0.94	0.93	12.84	8	0.117	0.000
<i>Grades 10, 11, and 12</i>									
CFA	2,303.42	1,311	.02 (.02, .02)	0.95	0.92				
Weak invariance	2,555.19	1,400	.02 (.02, .02)	0.94	0.92	251.77	89	<0.001	-0.009
Strong invariance	2,750.86	1,439	.02 (.02, .02)	0.93	0.91	195.66	39	<0.001	-0.009
Final	2,778.83	1,458	.02 (.02, .02)	0.93	0.91	27.97	19	0.084	0.000

compare latent parameters between any two concurrent waves.

Standardized factor loadings suggested that the general PYD factor was indicated by items from all Five Cs, but was weakly indicated by Social Competence, Physical Competence, and Physical Appearance. The residual Five C constructs were also indicated by nearly all of their respective indicators, although the conduct behavior item did not meaningfully load onto the residual Character construct. In addition, the Connection to Peers item loaded weakly onto the residual Connection construct, suggesting that the residual Connection construct emphasized Connection to Ecological Resources (family, neighborhood, and school).

Stability and Relationships Among the PYD Constructs

Latent means and variances for all constructs are presented in Table 7, while latent correlations from the PYD-VSF strong invariance model are presented in Table 8. Because we established partial invariance across all waves, latent means and variances can be directly compared across waves and these parameter estimates can be used to approximate developmental trajectories for each construct across adolescence. The only exception is Character, which seemed to change qualitatively during high school (i.e., three out of four factor loadings and three out of four intercepts were not invariant across the middle waves of data collection). The development of the Character

Table 5 Standardized factor loadings from the PYD-VSF strong invariance models

Grade	5	6	7	8	9	10	11	12
<i>Competence</i>								
Academic	0.32	0.36	0.36	0.34	0.33	0.32	0.33	0.35
Social	0.45	0.48	0.48	0.48	0.45	0.44	0.44	0.46
Physical	0.47	0.55	0.55	0.48	0.41	0.46	0.44	0.46
<i>Confidence</i>								
Self-worth	0.50	0.55	0.52	0.59	0.58	0.55	0.58	0.60
Positive identity	0.28	0.31	0.29	0.38	0.52	0.37	0.38	0.39
Physical appearance	0.42	0.45	0.47	0.61	0.64	0.64	0.63	0.65
<i>Character</i>								
Social conscience	0.68	0.70	0.55	0.54	0.50	0.51	0.54	0.56
Values diversity	0.34	0.41	0.31	0.31	0.42	0.41	0.42	0.43
Conduct behavior	0.00	0.00	-0.03	0.00	-0.10	-0.12	-0.12	-0.12
Personal values	0.61	0.61	0.48	0.46	0.41	0.31	0.32	0.36
<i>Caring</i>								
CARE2	Na	0.64	0.59	0.60	0.50	0.52	0.50	0.51
CARE7	Na	0.79	0.71	0.72	0.77	0.76	0.74	0.77
CARE9	Na	0.79	0.72	0.72	0.76	0.76	0.74	0.74
<i>Connection</i>								
Family	0.28	0.32	0.25	0.27	0.27	0.31	0.31	0.32
Neighborhood	0.57	0.57	0.52	0.58	0.56	0.50	0.50	0.50
School	0.44	0.45	0.39	0.44	0.44	0.41	0.40	0.41
Peers	0.12	0.12	0.11	0.12	0.23	0.22	0.22	0.22
<i>PYD</i>								
Academic	0.33	0.37	0.42	0.41	0.29	0.43	0.39	0.42
Social	0.17	0.19	0.21	0.22	0.13	0.22	0.19	0.20
Physical	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Self-worth	0.34	0.37	0.40	0.37	0.34	0.38	0.51	0.36
Positive identity	0.43	0.48	0.52	0.55	0.37	0.44	0.41	0.42
Physical appearance	0.16	0.17	0.20	0.21	0.13	0.23	0.20	0.21
Social conscience	0.36	0.37	0.43	0.45	0.46	0.47	0.43	0.45
Values diversity	0.22	0.27	0.30	0.32	0.30	0.29	0.27	0.27
Conduct behavior	0.38	0.44	0.49	0.46	0.61	0.49	0.43	0.45
Personal values	0.40	0.39	0.46	0.47	0.46	0.49	0.46	0.50
CARE2	Na	0.30	0.36	0.38	0.39	0.40	0.40	0.41
CARE7	Na	0.35	0.41	0.44	0.42	0.42	0.42	0.44
CARE9	Na	0.36	0.43	0.46	0.42	0.42	0.43	0.43
Family	0.44	0.50	0.50	0.51	0.49	0.45	0.41	0.42
Neighborhood	0.33	0.33	0.40	0.40	0.37	0.38	0.34	0.34
School	0.38	0.38	0.44	0.45	0.43	0.45	0.40	0.41
Peers	0.32	0.33	0.40	0.40	0.40	0.28	0.25	0.25

Estimates from analyses where target grade is earliest (e.g., Grade 7 comes from the model of Grades 7, 8, and 9)

Bold indicates a temporary change in the raw-metric factor loading

Bold and italic represents relatively stable changes in the raw-metric factor loading

construct must therefore be qualified by the fact that the Values Diversity subscale became more important for Character, while the Conduct Behavior (which only weakly indicated character) and Personal Values subscales

became less important for Character across waves of testing.

Results suggested rank-order and mean-level stability for PYD across waves while the residual C constructs

Table 6 Raw-metric intercepts from the PYD-VSF strong invariance models

Grade	5	6	7	8	9	10	11	12
<i>Competence</i>								
Academic	3.20	3.20	3.20	3.20	3.20	3.38	3.38	3.38
Social	3.32	3.32	3.32	3.32	3.32	3.32	3.32	3.32
Physical	2.73	2.73	2.73	2.73	2.55	2.73	2.73	2.73
<i>Confidence</i>								
Self-worth	3.39	3.39	3.39	3.39	3.39	3.39	3.39	3.39
Positive identity	4.30	4.30	4.30	4.30	4.24	4.38	4.38	4.38
Physical appearance	2.96	2.96	2.96	2.96	2.96	2.96	2.96	2.96
<i>Character</i>								
Social conscience	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03
Values diversity	3.47	3.93	3.89	3.85	3.85	3.85	3.85	3.85
Conduct behavior	2.89	2.89	2.89	2.89	2.73	2.73	2.73	2.73
Personal values	4.04	4.04	4.04	4.04	4.04	4.17	4.17	4.17
<i>Caring</i>								
CARE2	Na	3.78	3.78	3.78	3.95	3.95	3.95	3.95
CARE7	Na	3.80	3.80	3.80	3.80	3.80	3.80	3.80
CARE9	Na	3.82	3.82	3.82	3.82	3.82	3.82	3.82
<i>Connection</i>								
Family	4.31	4.31	4.31	4.31	4.31	4.31	4.31	4.31
Neighborhood	3.74	3.74	3.74	3.74	3.74	3.74	3.74	3.74
School	3.99	3.99	3.99	3.99	3.99	4.09	4.09	4.09
Peers	4.36	4.36	4.36	4.36	4.36	4.47	4.47	4.47

Estimates from analyses where target grade is earliest (e.g., Grade 7 comes from the model of Grades 7, 8, and 9)

Bold indicates a temporary change in the intercept

Bold and italic represents relatively stable changes in the intercept

Table 7 Latent means and variances for strong invariance PYD-VSF models

Grade	5	6	7	8	9	10	11	12
<i>Latent means</i>								
Competence	0.000	0.000	0.000	0.181	0.293	−0.280	−0.417	−0.417
Confidence	0.000	−0.288	−0.288	−0.150	0.000	−0.448	−0.448	−0.448
Character	0.000	−0.257	0.000	0.000	0.000	0.000	0.000	0.000
Caring	0.000	0.000	0.174	0.174	0.174	0.174	0.174	0.174
Connection	0.000	−0.392	−0.737	−0.737	−0.737	−0.737	−0.862	−0.862
PYD	0.000	0.214	0.000	−0.168	0.000	0.000	0.246	0.246
<i>Latent variances</i>								
Competence	1.000	1.000	1.000	1.000	0.763	0.763	0.763	0.763
Confidence	1.000	1.000	1.000	1.701	1.679	1.375	1.375	1.375
Character	1.000	1.000	0.574	0.574	0.430	0.430	0.430	0.430
Caring	1.000	1.000	0.751	0.751	0.846	0.846	0.608	0.608
Connection	1.000	1.000	0.743	1.000	1.000	0.748	0.748	0.748
PYD	1.000	1.000	1.277	1.444	1.277	1.277	0.998	0.998

Estimates from analyses where target grade is earliest (e.g., Grade 7 comes from the model of Grades 7, 8, and 9)

displayed lower levels of stability. The unique C constructs generally showed moderate rank-order stability during early adolescence that became stronger in later waves.

Character and Caring also displayed general mean-level stability, although the remaining C constructs tended to show mean-level declines across waves, especially during

Table 8 Stability and latent correlations among the PYD constructs

Grade	5	6	7	8	9	10	11	12
<i>Competence with</i>								
Competence (at T – 1)	Na	0.44***	0.47***	0.53***	0.76***	0.83***	0.66***	0.83***
Confidence	1.00 ^a	0.80***	0.76***	0.78***	1.00***	0.83***	0.92***	0.75***
Character	0.25***	0.17**	0.06	0.06	0.24***	0.02	0.24**	0.25**
Caring	Na	-0.01	-0.06	0.01	-0.04	-0.13*	-0.12	0.10
Connection	0.42***	0.46***	0.40***	0.33***	0.56***	0.66***	0.63***	0.55***
<i>Confidence with</i>								
Confidence (at T – 1)	Na	0.39***	0.66***	0.43***	0.83***	0.87***	0.70***	0.70***
Character	0.25***	0.17*	-0.15*	-0.06	0.12***	-0.16**	-0.01	-0.04
Caring	Na	-0.06	-0.24***	-0.19***	-0.12*	-0.20***	-0.19**	-0.14*
Connection	0.42***	0.39***	0.17*	0.32***	0.34***	0.43***	0.42***	0.52***
<i>Character with</i>								
Character (at T – 1)	Na	0.25**	0.43***	0.62***	0.64***	0.69***	0.79***	0.62***
Caring	Na	0.56***	0.44***	0.48***	0.50***	0.51***	0.55***	0.52***
Connection	0.30***	0.54***	0.35***	0.41***	0.42***	0.10	0.18*	0.16
<i>Caring with</i>								
Caring (at T – 1)	Na	Na	0.31***	0.50***	0.53***	0.38***	0.51***	0.49***
Connection	Na	0.18**	0.24***	0.20***	0.14*	0.10	-0.06	-0.06
<i>Connection with</i>								
Connection (at T – 1)	Na	0.59***	0.78***	0.64***	0.78***	0.62***	0.72***	0.93***
<i>PYD</i>								
PYD (at T – 1)	Na	0.81***	0.70***	0.78***	0.81***	0.90***	0.82***	0.81***

Estimates from analyses where target grade is earliest (e.g., Grade 7 comes from the model of Grades 7, 8, and 9)

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

^a Fixed parameters

the high school years. These results suggest that general adaptive functioning is relatively stable across adolescence while the specific aspects of PYD (i.e., levels of each C, controlling for PYD) display intra-individual variability.

Replicating previous findings from the 4-H Study that treated PYD as a second-order latent construct, our bifactor models suggested strong correlations among several Cs. Competence and Confidence were highly correlated within time of testing, suggesting that participants who rate themselves as being competent also rate themselves as having high self-confidence, even after controlling for interindividual differences in PYD. Character and Caring similarly correlated with each other while displaying consistently low correlations with Competence and Confidence. Connection displayed consistently high correlations with Competence and Confidence and consistently weak correlations with Caring. Connection also displayed moderate to high correlations with Character early in adolescence that decreased in magnitude during later adolescence. Thus, the residual C constructs tended to aggregate in two groups, efficacious (Competence and Confidence) and socioemotional (Character and Caring), with Connection correlating positively with each group.

Correlations with Criterion Variables

Similar to previous research that examined PYD as a higher-order latent construct, our bifactor measure of general PYD displayed consistently significant correlations with all three criterion variables across all eight waves of the 4-H Study. Table 9 presents the latent correlations among PYD, depressive symptoms, problem behaviors, and contribution. Unsurprisingly, PYD correlated negatively with indices of negative development (i.e., depressive symptoms and problem behaviors) and correlated positively with our indicator of positive development (i.e., contribution). The magnitude of the correlations between PYD and both contribution and problem behaviors increased across waves, suggesting that across the 4-H Study, PYD became increasingly related to behavioral indices of positive development. However, the correlation between PYD and depressive symptoms attenuated slightly across waves, suggesting a weak bifurcation of the constructs across adolescence.

In contrast to the findings for general PYD, the residual C constructs representing efficacious development (Confidence and Competence) displayed weak to moderate correlations with behavioral measures of positive development (contribution and problem behaviors) but increasingly strong (negative) correlations with depressive symptoms. While unsurprising, given the tight coupling between depressive symptoms and having a negative self-image (e.g., Harter and Jackson 1993), we must additionally interpret these relationships in the context of each residual

C being completely orthogonal to general PYD. Thus, while our general index of PYD slowly decoupled with depressive symptoms across adolescence, depressive symptoms became increasingly correlated with participants' efficacious self-concepts.

While the correlations between our criterion variables and both PYD and the two efficacious Cs displayed a consistent pattern of development, the correlations between our criteria and the residual Cs that represent socioemotional functioning (Character and Caring) displayed greater stability. The socioemotional Cs moderately correlated (positively) with both contribution and depressive symptoms while being generally unrelated with problem behaviors across adolescence. The positive correlations between the socioemotional Cs and depressive symptoms are especially surprising, but again must be interpreted in light of the fact that these two Cs both represent socioemotional functioning independent of general PYD. While this finding must be replicated in future work, we can abductively speculate that self-reported Caring and Character that exceeds the levels expected from participants' level of general PYD may indicate emotional hypersensitivity, or an anxiety-producing over concern for (or about) others' thoughts and feelings that correlates with increased depressive symptomology.

The final residual C, Connection, displayed criterion correlations that were somewhat similar to those displayed by the efficacious Cs, and yet also somewhat like the socioemotional Cs. That is, Connection displayed moderately positive correlations with contribution and essentially zero correlations with problem behaviors across adolescence, mirroring the parallel correlations between the socioemotional Cs and those same criterion variables. Similar to the efficacious Cs, Connection also displayed negative correlations with depressive symptoms that increased in magnitude over time. Connection was therefore consistently related to contribution behaviors across adolescence and may buffer against depressive symptoms in late adolescence.

Discussion

Assessing the empirical usefulness of theoretical models of PYD through scrutiny of model-specific indicators of PYD remains an important task for developmental researchers. While existing evidence supports the empirical validity of such models, further empirical investigation is needed. In this article we discuss one popular model of PYD, the Lerner and Lerner 5 Cs Model, and extend findings that support its empirical validity across adolescence. We extend previous research by examining the bifactor structure of PYD as indexed by the Five Cs of PYD across all

Table 9 PYD-VSF correlations with criterion variables

Grade	5	6	7	8	9	10	11	12
<i>PYD with</i>								
Contribution	0.21***	0.41***	0.41***	0.44***	0.53***	0.58***	0.56***	0.57***
Depressive symptoms	-0.68***	-0.55***	-0.52***	-0.52***	-0.49***	-0.40***	-0.47***	-0.47***
Problem behaviors	-0.46***	-0.52***	-0.50***	-0.54***	-0.60***	-0.61***	-0.57***	-0.66***
<i>Competence with</i>								
Contribution	0.19***	0.27***	0.20***	0.20***	0.25***	0.15***	0.21***	0.27***
Depressive symptoms	-0.18***	-0.23***	-0.21***	-0.26***	-0.42***	-0.38***	-0.40***	-0.41***
Problem behaviors	0.03	0.12**	0.09*	0.15**	0.10	0.02	0.08	0.09
<i>Confidence with</i>								
Contribution	0.19***	0.17**	-0.01	0.01	0.06	-0.05	-0.02	0.06
Depressive symptoms	-0.18***	-0.28***	-0.34***	-0.38***	-0.44***	-0.46***	-0.43***	-0.52***
Problem behaviors	0.03	0.15**	0.13**	0.25***	0.21***	0.13***	0.12**	0.17***
<i>Character with</i>								
Contribution	0.33***	0.39***	0.47***	0.49***	0.36***	0.38***	0.31***	0.34***
Depressive symptoms	0.11*	0.09*	0.23***	0.22***	0.22***	0.31***	0.21***	0.24***
Problem behaviors	-0.03	0.02	-0.05	-0.01	0.00	0.12**	0.07	0.26***
<i>Caring with</i>								
Contribution	Na	0.18***	0.32***	0.30***	0.15***	0.15***	0.17***	0.15**
Depressive symptoms	Na	0.09*	0.24***	0.26***	0.21***	0.20***	0.28***	0.13**
Problem behaviors	Na	-0.03	0.04	0.03	0.08	0.06*	0.05	0.14**
<i>Connection with</i>								
Contribution	0.35***	0.37***	0.50***	0.44***	0.38***	0.24***	0.29***	0.29***
Depressive symptoms	0.01	-0.11	-0.06	-0.05	-0.16**	-0.35***	-0.38***	-0.40***
Problem behaviors	0.03	0.06	0.07	0.01	-0.03	0.00	0.12*	0.16*
<i>Contribution with</i>								
Depressive symptoms	-0.09**	-0.19***	-0.11***	-0.14***	-0.16***	-0.14***	-0.16***	-0.22***
Problem behaviors	-0.08*	-0.17***	-0.19***	-0.21***	-0.27***	-0.27***	-0.25***	-0.22***
<i>Depressive symptoms with</i>								
Problem behaviors	0.25***	0.29***	0.27***	0.20***	0.23***	0.27***	0.12***	0.23***

Estimates from analyses where target grade is earliest (e.g., Grade 7 comes from the model of Grades 7, 8, and 9)

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

eight waves of the 4-H study, establishing factorial invariance across time while also presenting findings that speak directly to the stability and criterion validity of the overarching measure of PYD as well as for the residual C constructs that represent systematic item variance not directly related to the overarching PYD measure.

Our findings suggest moderate rank-order stability of PYD and the residual C constructs during early adolescence, which increases in magnitude to indicate strong rank-order stability by late adolescence. Most constructs displayed mean-level stability, although the Competence, Confidence, and Connection constructs displayed mean-level declines over time. The relative mean-level stability of the Character and Caring constructs mirrors their generally stable criterion correlations, while the mean-level changes in Competence, Confidence, and Connection also reflect the longitudinally increasing criterion correlation

between those constructs and depressive symptoms. Our general measure of PYD displayed stability both in terms of the magnitude of criterion correlations and mean levels across adolescence, however.

The differential development of the PYD and residual C constructs, and especially the differential development of the criterion correlations that each construct displayed, offers added support for continued examination of the Five Cs of PYD using a bifactor framework. While not addressed as a target hypothesis in this article, the factor loadings presented in Table 5 further highlight this point, reinforcing findings presented by Geldhof and colleagues (in press) that found differential relationships between specific indicators and the general PYD versus residual C constructs. For instance, PEER 6, the item measuring connection to peers indicated PYD much more strongly than it indicated the residual Connection construct, while

other Connection indicators represented the two constructs more equally.

Interpreting Cross-Wave Differences

The present study reports the first-ever examination of PYD and its relationships with important criterion variables across all eight waves of the 4-H Study of PYD, as well as the first attempt to interpret stability and criterion correlations for a bifactor model of PYD substantively. Our findings add potentially important information to the field's understanding of the Five Cs of PYD and their development across adolescence. Our results suggest several marked differences between early and late adolescence while highlighting the joint importance of both global PYD and the individual Cs.

Any developmental differences presented in this study must be interpreted with caution, however, as our longitudinal findings conflate maturational change with several non-developmental factors. For example, our sample moved from being relatively heterogeneous in earlier waves to having disproportionately large numbers of White, female and higher-SES participants. Future research should therefore replicate these longitudinal findings with a more diverse array of independent samples. Similarly, only a fraction of the 4-H Study's more than 7,000 participants provided data in any given wave, with the overlap between adjacent waves often being relatively small. Estimates of rank-order stability assume complete data between adjacent waves of data, and it is not entirely clear how the limited overlap between waves affected our stability estimates in this article. On one hand, we might argue that missingness occurred at random and should have little impact on our stability estimates. It is entirely possible that attrition was not completely random, however (e.g., Bowers et al. 2011), and if this situation was in fact the case, our stability estimates may be inflated as a result, for instance, of being estimated using only data from our most dedicated/consistent participants.

Nevertheless, both the present research and prior work pertinent to the psychometric characteristics of the very short form of the PYD measure (Geldhof et al., in press) indicate that the PYD-VSF may be used to provide researchers or practitioners with a useful index of global PYD, with an index for individual Cs, or an index of the role of a specific item in the development of youth. Of course, any of these uses must be rationalized by an appropriate theory-predicated question and must be employed in recognition of the limitations of any of these uses. For instance, Geldhof et al. (in press) explain that using the PYD-VSF to index global PYD is limited by the fact that some items (e.g., ones regarding physical or social competence or physical appearance) should not be used in

the computation of such scores. In turn, if individual items within the PYD-VSF are of interest, then the items' C-specific and PYD factor loadings should be considered together in any interpretation of findings about the significance for development of a particular item (Geldhof et al., in press). Finally, if C-specific composite scores are of interest, such scores can be computed under the assumption that overlap between individual Cs and global PYD is not relevant to the question being tested (Geldhof et al., in press).

In short, then, in the context of these constraints on its use, the PYD-VSF can be an efficient and useful index of features of youth thriving and of their links to positive and problematic features of development across the adolescent decade. The present findings about these links reflect two important themes. First, as illustrated in prior research derived from the 4-H Study (Lewin-Bizan et al. 2010; Phelps et al. 2007), the links between PYD and indices of either positive or problematic features of youth development are neither perfectly positively nor negatively correlated, respectively. The magnitude of these relations suggests that there are many different patterns of covariation among PYD and other positive and problematic features of youth development. Second, the nature of these links changes across the adolescent decade. This latter finding raises the still-to-be-addressed issue of the basis of developmental changes in these patterns of covariation. The relational developmental systems model framing the research conducted within the 4-H Study (Lerner et al. 2013, in press) suggests that variation in the relations between the actions of youth on their context and the actions of the context on youth accounts for this variation. However, this general idea must be linked to research that elucidates the specific features of youth and the specific features of contexts that, when they occur at specific points in adolescence, result in such variation. Although adequate explanations of the variation across development in the links between PYD and positive and problematic behaviors must await the results of such research, the present findings demonstrate, at a descriptive level, that the PYD-VSF is a useful measure of youth thriving per se and of the links between such thriving and other important features of behavior across the adolescent decade. This potential utility has, then, important implications for both research and applications to youth development programs.

Future Directions and Implications for Application

The present bifactor models add to our understanding of PYD and its relationships with important criteria, supplementing previous research that has largely implemented higher-order CFA models and mixture regression analyses. The PYD data from the 4-H Study are far from being fully

explored, however, and our findings only represent a single stepping stone in the larger path to understanding the development of positive functioning during adolescence and beyond. Future research must continue to consider not only which factors moderate relationships among the Five Cs of PYD and important criteria, but also continue exploring how important predictors of positive development (e.g., self-regulation, see Geldhof et al. 2010; McClelland et al. 2010) relate to the various facets of PYD when it is modeled with a bifactor structure. Fully understanding these relationships will require triangulation across multiple quantitative and qualitative methods, as well as a nuanced understanding of how context moderates the above processes. As such, even though the present results summarize 8 years of the 4-H Study of PYD, they mark only the beginning steps of an exciting and quickly growing body of literature that focuses on positive development both among youth and across the life span.

Nonetheless, the psychometric features of the PYD-VSF that have been demonstrated in this study indicate that this measure may provide an efficient means for practitioners to use in efforts to index the positive development of youth program participants across much of the second decade of life. Similarly, the PYD-VSF may be useful as a tool in evaluations of youth development programs having a theory of change related to the promotion of PYD. As such, the findings of the present research suggest that the PYD-VSF measure constitutes a “quick and clean” means to appraise the ways in which youth development programs may foster thriving among adolescent participants.

Conclusions

Across the last decade, researchers, youth development practitioners, and policy makers have all displayed burgeoning of interest in the bases and outcomes for individuals and society of the positive development of youth (e.g., see Lerner et al., in press). The research associated with this attention to thriving among young people has resulted in indexes of PYD becoming almost standard measures in the study of adolescent development and community-based youth development programs (Lerner et al. 2013). Derived from the decade-long 4-H Study of PYD, the present research provides evidence that it is possible to assess the development of thriving among adolescents with a theoretically-predicated measure that has good psychometric characteristics and that can be used efficiently in either research or applied settings.

Accordingly, given the interest among both researchers and practitioners to include PYD as one of a suite of measures appraising both the individual and ecological antecedents of thriving and, as well, the impact of PYD on

both positive and problematic behaviors of youth, a short but psychometrically sound index of PYD is an important, and indeed a highly valuable, addition to the measurement toolbox of developmental scientists and developmental practitioners. If scholars are to translate their research into applications of use in youth development programs and, in turn, if practitioners want to assess the success of their programs vis-à-vis the theories of change framing their work, tools that are valid, reliable, and feasible to use in the day-to-day work of actual youth development programs must be developed. The present research constitutes a major step forward in providing researchers and practitioners with just such a measure, and for a construct central in the contemporary study of youth development. As such, we expect that the present data set constitutes only a baseline for future research and applications about the relations between PYD and other indices of adolescent behavior and development.

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

Author contributions GJG helped conceive this article, draft the manuscript, and led all data analyses; EPB helped conceive this article and draft the manuscript; MKM assisted with data analysis and helped draft the manuscript; CMN assisted with data analysis and helped draft the manuscript; KSC assisted with data analysis and helped draft the manuscript; RML was PI of the 4-H Study of Positive Youth Development, helped conceive this article, and assisted in drafting the manuscript. All authors read and approved the final manuscript.

References

- Benson, P. L. (2006). *All kids are our kids* (2nd ed.). San Francisco: Jossey-Bass.
- Benson, P. L., Leffert, N., Scales, P. C., & Blyth, D. A. (1998). Beyond the “village” rhetoric: Creating healthy communities for children and adolescents. *Applied Developmental Science, 2*(3), 138–159.
- Benson, P. L., Mannes, M., Pittman, K., & Ferber, T. (2004). Youth development, developmental assets, and public policy. In R. M. Lerner, L. Steinberg, R. M. Lerner, & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 781–814). Hoboken: Wiley.
- Bowers, E. P., Geldhof, G. J., Johnson, S. K., Lerner, J. V., & Lerner, R. M. (in press). Thriving across the adolescent years: A view of the issues. *Journal of Youth and Adolescence*.
- Bowers, E. P., Gestsdottir, S., Geldhof, G. J., Nikitin, J., von Eye, A., & Lerner, R. M. (2011). Developmental trajectories of intentional self regulation in adolescence: The role of parenting and implications for positive and problematic outcomes among diverse youth. *Journal of Adolescence, 34*, 1193–1206.
- Bowers, E. P., Li, Y., Kiely, M. K., Brittan, 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.

- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- Davis, M. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology*, 10(4), 1–17.
- de Gruijter, D. N. M., & van der Kamp, L. J. Th. (2008). *Statistical test theory for the behavioral sciences*. New York: Chapman & Hall.
- 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, DC: National Academy Press.
- Eisenberg, N., Fabes, R. A., Murphy, B. C., Karbon, M., Smith, M., & Maszk, P. (1996). The relations of children's dispositional empathy-related responding to their emotionality, regulation, and social functioning. *Developmental Psychology*, 32, 195–209.
- Erikson, E. H. (1968). *Identity: youth and crisis*. Oxford: Norton & Co.
- Freud, A. (1969). Adolescence as a developmental disturbance. In G. Caplan & S. Lebovici (Eds.), *Adolescence* (pp. 5–10). New York: Basic Books.
- Geldhof, G. J., Bowers, E. P., Boyd, M. J., Mueller, M., Napolitano, C. M., Schmid, K. L., et al. (in press). The creation and validation of short and very short measures of PYD. *Journal of Research on Adolescence*.
- Geldhof, G. J., Little, T. D., & Colombo, J. (2010). Self-regulation across the lifespan. In M. E. Lamb & A. M. Freund (Vol. Eds.), and R. M. Lerner (Editor-in-Chief). *Social and emotional development*. Volume 2 of *The handbook of life-span development* (pp. 116–157). Hoboken, NJ: Wiley.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255–282.
- Hall, G. S. (1904). *Adolescence: Its psychology and its relations to physiology, anthropology, sociology, sex, crime, religion, and education* (Vols. 1 and 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. (1986). *Manual for the self-perception profile for adolescents*. Denver, CO: University of Denver.
- Harter, S. (1988). *Manual for the self-perception profile for adolescents*. Denver, CO: University of Denver.
- Harter, S., & Jackson, B. K. (1993). Young adolescents' perceptions of the link between low self-worth and depressed affect. *Journal of Early Adolescence*, 13(4), 383–407.
- 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., et al. (2005). Thriving in adolescence: The voices of youth-serving practitioners, parents, and early and late adolescents. *Journal of Early Adolescence*, 25(1), 94–112.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of estimation of test reliability. *Psychometrika*, 2, 151–160.
- 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, R. M. (2004). *Liberty: Thriving and civic engagement among American youth*. Thousand Oaks, CA: Sage.
- Lerner, J. V., Bowers, E. P., Minor, K., Boyd, M. J., Mueller, M. K., Schmid, K. L., et al. (2013). 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 edn). Editor-in-chief: I. B. Weiner. Hoboken, NJ: Wiley.
- Lerner, R. M., Lerner, J. V., Almerigi, J., Theokas, C., Phelps, E., Gestsdottir, S., et al. (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., Bowers, E., & Geldhof, G. J. (in press). Positive youth development. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of child psychology and developmental science*, Volume 1. *Theory and methodology* (7th edn). Richard M. Lerner: Editor-in-Chief. Hoboken, NJ: Wiley.
- Lerner, R. M., Lerner, J. V., Lewin-Bizan, S., Bowers, E. P., Boyd, M., Mueller, M. K., et al. (2011). Positive youth development: Processes, programs, and problematics. *Journal of Youth Development*, 6(3), 40–64.
- 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 edn., pp. 524–558). Hoboken, NJ: Wiley.
- Lerner, R. M., von Eye, A., Lerner, J. V., Lewin-Bizan, S., & Bowers, E. P. (Eds.). (2010). The meaning and measurement of thriving in adolescence: Findings from the 4-H study of positive youth development [special issue]. *Journal of Youth and Adolescence*, 39(7), 113–143.
- Lewin-Bizan, S., Doyle Lynch, A., Fay, K., Schmid, K., McPherran, C., Lerner, J. V., et al. (2010). Trajectories of positive and negative behaviors from early- to middle adolescence. *Journal of Youth and Adolescence*, 39(7), 751–763.
- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guilford.
- McClelland, M. M., Ponitz, C. C., Messersmith, E., & Tominey, S. (2010). Self-regulation: The integration of cognition and emotion. In R. Lerner (Series Ed.) & W. Overton (Vol. Ed.), *Handbook of life-span development* (Vol. 1: Cognition, biology and methods, pp. 509–553). Hoboken, NJ: Wiley.
- 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.
- Phelps, E., Balsano, A., Fay, K., Peltz, J., Zimmerman, S., Lerner, R. M., et al. (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., Jeličić, 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.
- 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.
- Small, S. A., & Rodgers, K. B. (1995). *Teen assessment project (TAP) survey question bank*. Madison: University of Wisconsin-Madison.
- von Eye, A., Martell, M. M., Lerner, R. M., Lerner, J. V., & Bowers, E. P. (2011). Interpreting 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.), *Positive youth development: Research and applications for promoting thriving in adolescence* (pp. 349–375). New York: Elsevier.

G. John Geldhof is an Assistant Professor at Oregon State University. He received his doctorate in Quantitative and Developmental Psychology from the University of Kansas. His research interests include quantitative methods, positive youth development, and models of self regulation.

Edmond P. Bowers is a Research Assistant Professor at Tufts University. He received his doctorate in Applied Developmental and Educational Psychology from Boston College. His research interests include a focus on important non-parental adults and the various ways that these adults can promote positive development in young people.

Megan Kiely Mueller is a research assistant professor in the Center for Animals and Public Policy and the Department of Clinical Sciences at the Tufts University Cummings School of Veterinary Medicine. Megan holds a Ph.D. in Applied Child Development from Tufts University, and her research focuses on the role of human-animal interaction in promoting positive youth development and health outcomes.

Christopher M. Napolitano is a post-doctoral fellow at the University of Zurich. His research interests include serendipity and intentional self regulation.

Kristina Schmid Callina is a post-doctoral fellow at the Institute for Applied Research in Youth Development at Tufts University. Her work integrates the study of positive youth development and the concept of hope.

Richard M. Lerner is the Bergstrom Chair in Applied Developmental Science and the Director of the Institute for Applied Research in Youth Development at Tufts University. He received his Ph.D. in developmental psychology from the City University of New York. His work integrates the study of public policies and community-based programs with the promotion of positive youth development and youth contributions to civil society.