The Five Cs Model of Positive Youth Development: A Longitudinal Analysis of Confirmatory Factor Structure and Measurement Invariance[1]

Edmond P. Bowers, Yibing Li, Megan K. Kiely, Aerika Brittian

Tufts University

Jacqueline V. Lerner

Boston College

Richard M. Lerner

Tufts University

[1] Correspondence should be addressed to Edmond P. Bowers, Institute for Applied Research in Youth Development, 307 Lincoln-Filene Building, Tufts University, Medford MA, 02155. Email Ed.Bowers@tufts.edu

Abstract

The understanding of positive development across adolescence rests on having a valid and equivalent measure of this construct across the breadth of this period of life. Does the Positive Youth Development (PYD) construct based on the Five Cs model have satisfactory psychometric properties for such longitudinal measurement invariance? Using longitudinal data derived from the 4-H Study of PYD, we assessed 920 youth (61.6% female) from a racially and ethically diverse sample (67.3% European American) who participated in three waves (Grades 8 to 10) of data collection. Building on prior findings that the Five Cs (i.e., Competence, Confidence, Connection, Character, and Caring) model of PYD was a robust measure that could be assessed comparably during early adolescence, we tested a hierarchy of second-order confirmatory factor analysis models to assess the extent to which PYD can be measured equivalently across middle adolescence. Evidence was found for strict measurement invariance across three measurement occasions, including equivalence of first-order and second-order factor loadings, equality of intercepts of observed variables, and equality of item uniqueness and disturbances of the first-order factors. These results suggest that PYD can be measured in the same way across measurement occasions, a prerequisite for the study of development. Implications for research and application of being able to measure PYD equivalently across adolescence are discussed.

Keywords: Positive Youth Development (PYD), confirmatory factor analysis (CFA), longitudinal measurement invariance, adolescence.

Since its launching over a century ago, the scientific study of adolescence has largely been framed by a "deficit perspective" in which the second decade of life was considered a period of "storm and stress" (Hall, 1904), developmental disturbance (Freud, 1969), or crisis (Erikson, 1968). Adolescents, within this view, were problems to be managed (Roth & Brooks-Gunn, 2003), and if there was to be positive development during this period of life it involved absences of or decreases in problems. The pervasive influence of the "deficit perspective" on research aims, policy, and practice is reflected in the prevalence of measures of risk and problematic behaviors that are most often collected by researchers and by program and service organizations. It appears that it is much easier to determine what youth should avoid (violence, drugs) or not be marked by (mental health problems) than to agree on the characteristics and experiences that are either indicators of thriving or that could enhance adolescents' lives (Moore, Lippman, & Brown, 2004).

Partly in response to this focus on the problems and deficits among young people, a new approach to adolescent development has emerged over the past 20 years – the positive youth development (PYD) perspective. The PYD perspective moves beyond the negative, deficit view of youth that dominated developmental science, psychology, education, sociology, public health, and other fields through the twentieth century and towards a view of the strengths of youth and the positive qualities and outcomes we wish youth to develop.

PYD has been conceptualized in a number of ways, and several theoretical frameworks have been posited over the past few decades (for a review, see J. Lerner, Phelps, Forman, & Bowers, 2009). As these models become more popular with individuals working to enhance the positive growth of young people in homes, schools, and youth-serving organizations, it is important that these models are empirically valid, can be widely applied, and include constructs

that are specific and measurable. However, these models are only just beginning to be tested. Recent work has attempted to evaluate youth development frameworks (Heck & Subramaniam, 2009) and indicators of PYD (Dukakis, London, McLaughlin, & Williamson, 2009), but further investigation of suitable models is needed.

A recent review of PYD frameworks has indicated that the Five Cs Model of PYD is the most empirically supported framework to date (Heck & Subramaniam, 2009). Empirical evidence indicates that this construct has good psychometric properties. For instance, each of the Five Cs has good internal consistency (Lerner et al., 2005). Phelps et al. (2009) provided additional validity using data from early adolescence. However, it is important to determine whether the model is valid for a larger age range of adolescents. Therefore, to assess the structure and development of the Five Cs model of PYD in middle adolescence, the present research extended past findings from the 4-H Study of PYD (Lerner et al., 2005; Phelps et al., 2009) by using its longitudinal data from the middle years of adolescence (Grades 8-10, approximately 13-17 years old). By illuminating the developmental course of PYD across middle adolescence, we hoped to reinforce empirical support for a strength-based approach to youth, and to address the lack of valid and robust indicators of positive development in several domains of youth functioning. Achieving these ends would have both scientific and practical impact.

The Five Cs Model of PYD

The Five Cs model of PYD emphasizes the strengths of adolescents and, as a consequence, enables youth to be seen as resources to be developed (Roth & Brooks-Gunn, 2003). This model is framed by developmental systems theories and, more specifically, by the focus on the plasticity of development within such theories (Lerner, 2004). The model posits that positive development occurs if the strengths of youth (represented, for instance, by the enormous

potential for systematic growth, i.e., for plasticity, within the adolescent period) are aligned systematically with positive, growth promoting resources in the ecology of youth (resources that were termed "developmental assets;" Benson, Scales, Hamilton, & Sesma, 2006). The positive development that results from this alignment can be operationalized by "Five Cs" – Competence, Confidence, Connection, Character, and Caring. These Five Cs are based on both the experiences of practitioners and on reviews of the adolescent development literature (Eccles & Gootman, 2002; Lerner, 2004; Roth & Brooks-Gunn, 2003) and are linked to the positive outcomes of youth development programs (e.g., Roth and Brooks-Gunn, 2003). These domains are interactive and young people require healthy development in all of them (Dukakis et al., 2009). In addition, the 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 an adolescent manifests these Five Cs over time, he or she is more likely to be on a life trajectory marked by mutually beneficial person ←→ context relations that contribute to self, family, community, and civil society (i.e., Contribution – the sixth C; Lerner, 2004) and less likely to be on a trajectory of risk and problem behavior such as substance abuse, delinquency, and depression (e.g., Pittman et al., 2001; but see Phelps et al., 2007). Table 1 presents the definitions of the Cs as found in Lerner et al. (2005).

Insert Table 1 about here

Prior Measurement of the Five Cs Model of PYD

The 4-H Study of Positive Youth Development (Lerner et al., 2005), a longitudinal study spanning from Grade 5 to, at this writing, Grade 12, 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, which 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 (Jeličić et al., 2007). Results suggested that the Five Cs can be cast in terms of latent constructs which in turn, load on a second-order PYD latent construct. This second-order PYD construct has been found to be related to latent constructs for contribution, depression, and risk/problem behaviors (Jeličić, Bobek, Phelps, Lerner, & Lerner, 2007). 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.

As the 4-H Study of PYD continues, data have been collected from youth in Grades 8 to 10. Therefore, it is possible to determine whether a structurally stable PYD construct can be found as young people move from the period of early adolescence to middle adolescence. As adolescents change biologically, cognitively, and socially, so too may the nature of PYD. That is, measures of PYD may need to be updated to preserve their meanings as youth get older. Critical developmental changes occur over the middle adolescent years as the developmental tasks to be accomplished change; these tasks include the growth of the prefrontal cortex, the formulation of an adaptive identity, and the burgeoning needs for autonomy and intimacy (e.g., Susman & Dorn, 2009). At the same time, adolescents in the U.S. experience substantial contextual changes

over Grades 8 to 10, each providing new opportunities and risks. For example, youth proceed through the transition from middle school or junior high to high school, participate in more activities outside the home, and experience increased involvement with their peer group.

In addition, for youth living in the contemporary United States, the process of social redefinition begins in middle adolescence (Arnett, 1997); during this period, society recognizes that a young person's status has changed as he or she is now allowed to drive, work, or leave school, for example. As youth begin to experience increased power and freedom, they are also expected to be more adept at self-management, to be more personally responsible, and to participate more in society. As individuals and contexts change, the domains of self-concept that are most strongly related to overall self-worth and measures of achievement also undergo changes during adolescence (e.g., Harter, 1999; Shapka & Keating, 2005). Domains of self-perceived competence become more differentiated across ontogeny as cognitive competence eventually separates into scholastic competence, intellectual ability, and creativity. In a similar manner, social competence transforms into job competence, close friendships and romantic relationships over the adolescent period (Harter, 1999). Of particular note is that while perceptions of appearance are most closely tied to general self-worth in high-school students (Shapka & Keating, 2005), this indicator is rarely included in conceptions of confidence in PYD research. In short, the dynamic nature of adolescent development increases the necessity of validating acceptable measures of PYD during the middle adolescent years.

Testing Measurement Invariance of Second-Order Factor Models

Second-order factor models have been used in psychology over a diversity of areas such as the Big Five personality structure (DeYoung, Peterson, & Higgins, 2002), the Five Cs model of PYD (Lerner et al., 2005), and psychological well-being (Hills & Argyle, 2002). The Five Cs

second-order model represents the hypothesis that these five distinct, but related constructs can be accounted for by one common underlying higher order construct (i.e., PYD). This conceptualization allows for a more parsimonious and interpretable model of positive youth development (Chen, Sousa, & West, 2005). In order to determine whether PYD is a structurally stable construct across the middle adolescent period (Grades 8-10), it is appropriate to test whether the Five Cs second-order factor model exhibits measurement invariance over this three year span. In a seminal review, Vandenberg and Lance (2000) described a comprehensive paradigm for evaluating measurement invariance within a confirmatory factor analysis (CFA) framework. Measurement invariance is usually investigated by a series of nested models at the configural level (the same item must indicate the same latent factor), metric level (the factor loadings of like items are equal), and the scalar level (different groups have the same intervals and zero points). In the present work, we have applied this paradigm to data obtained from Grades 8-10 of the 4-H Study to establish invariance of the Five Cs model of PYD.

A measure is invariant when members from different populations who are identical on the construct of interest being measured obtain the same score on the measure (Schmitt & Kulijanin, 2008). Tests of measurement invariance are necessary given that, as adolescents develop, we want to be able to use the PYD construct to make group comparisons across gender, age, and ethnicity, among other characteristics. In order to be able to make meaningful statistical comparisons (e.g., involving means, regression coefficients) across different groups, we must determine that the same constructs are being assessed in each group. In the present research, we tested the measurement invariance of PYD across Grades 8, 9, and 10. We wanted to determine whether the Five Cs model of PYD was measuring the same construct at each grade. To accomplish this goal, we had to establish invariance across the grades at several levels of the

second-order factor model through a series of hierarchical series of structural models, as will be detailed more fully in the presentation of our analytic plan.

Study Aims and Hypotheses

Although instances of testing measurement equivalence are increasing, there are still few examples of testing measurement invariance in second-order models in the research literature (Byrne, 1995; Byrne & Campbell, 1999). Of those existing examples, most do not go beyond the examination of the covariance structure (for examples, see Chen et al., 2005). Prior work on the Five Cs model established configural invariance (Phelps et al., 2009). We planned to extend these findings and to establish measurement invariance of the Five Cs model by using the full mean and covariance structures in order to address the following questions. Does the Five Cs model of PYD identified by Phelps et al. (2009) among fifth through seventh graders describe the structure of PYD when considering youth as they progress from eighth through tenth grades? Does the constitutive structure of PYD remain as youth progress through middle adolescence, i.e., is PYD measured the same way over time? Based on theory and previous research, we hypothesized that the Five Cs model of PYD would continue to hold among middle adolescents and that strong longitudinal factorial invariance would exist, i.e., the latent construct of PYD would be measured the same way across measurement occasions.

Method

Full details of the methodology of the 4-H Study have been presented in prior reports (e.g., Lerner et al., 2005; Phelps et al., 2007, 2009). Accordingly, we present only those features of methodology pertinent to the present article.

Participants

Participants for the current study were 920 adolescents from the 4-H Study of PYD sample at the fourth through sixth years of data collection, i.e., Grades 8 to 10 (37.1% males, mean age in Grade 8 = 14.19 years, SD = .99 years). Youth were from 30 states. Youth were included only if they had participated in the study in at least two waves out of the three waves included in the analyses. The average family per capita income over the three year range was $16,543. The sample reflected racial and ethnic diversity. A majority of the youth were European American (67.3%), but African-American (7.5%), Latino American (8.0%), and Asian American (3.4%) youth were also represented, among others. Table 2 presents the sample characteristics. Comparing the present sample study to the sample used to study the structure and development of PYD in early adolescence (Phelps et al., 2009), there are several key demographic differences. In addition to sample size and age differences, the present sample is composed of a higher percentage of female participants and participants who reported higher per capita incomes. The present sample is also more likely to come from an urban community rather than a suburban community.

Insert Table 2 about here

Procedure

For Grades 8 through 10 of data collection, teachers or program staff gave each adolescent an envelope to take home to his or her parent/guardian, containing a letter explaining the study and a consent form. For those youth who received parental consent, data collection was conducted either in the school or program by trained researchers or hired assistants for remote

locations. Instructions were read by the research assistant or hired assistant before adolescents were permitted to begin the student questionnaire (SQ). Participants were instructed that they were allowed to skip any question(s) that is too personal or difficult to answer.

An online version of the SQ for participants to take the survey via the Internet was also used beginning in Grade 7. In most cases, schools or after-school programs allocated computers for participants to take the survey online. In cases where computers were not available, students were given a paper copy of the questionnaire to complete. The questionnaires took approximately one and a half hours to two hours to complete, and participants were encouraged to take short breaks at their leisure. Students, who were not present at their school or 4-H site, in that they were either absent during the day of testing or the school superintendent did not allow testing to occur in the school, received a document with login information for participants to take the survey online or were mailed a paper copy of the survey with a return prepaid envelope.

Measures

The measures used in the present study were relevant to the assessment of the latent structure of the Five Cs of PYD. The Five Cs are measured using subscales from the following measures: the *Profiles of Student Life-Attitudes and Behaviors Survey* (PSL-AB; Benson, Leffert, Scales, & Blyth, 1998), the *Self-Perception Profile for Adolescence* (SPPA; Harter, 1988), the *Peer Support Scale* (Armsden & Greenberger, 1987) from the *Teen Assessment Project Survey Question Bank* (Small & Rodgers, 1995), the *Eisenberg Sympathy Scale* (ESS; Eisenberg et al., 1996), and the Empathic Concern Subscale of the *Interpersonal Reactivity Index* (IRI; Davis, 1983). Detailed information regarding the measurement of each of the Cs is presented below.

Competence. In Grades 8, 9, and 10, Competence was defined as *Academic Competence,*

Social Competence, Athletic Competence, and *grades*. Six items from the *Self-Perception Profile for Adolescents* (SPPA, Harter, 1988) form the *academic competence* scale, six of the items form the *social competence* scale, and six items form the *athletic competence* scale. Harter (1988) developed a structured alternative response format to assess perceived competence in a domain. Participants are asked to choose between two types of teenagers. Once they have selected which person they are most like, they are asked to decide if it is "really true for me" or "sort of true for me." The items are counterbalanced so that half begin with a positive sentence, reflecting high competence, while half begin with a negative sentence, reflecting low competence. Each item is scored from 1-4, with 4 reflecting higher perceived competence.

An example of an item from the *academic competence* scale is "Some teenagers feel like they are *just as smart* as other teenagers their age BUT Other teenagers aren't so sure and *wonder* if they are as smart." Cronbach's alphas for Grades 8, 9, and 10 were .78, .78, and .81, respectively. An example of an item from the *social competence* scale is "Some teenagers have *a lot* of friends BUT Other teenagers *don't* have very many friends." Cronbach's alphas for Grades 8, 9, and 10 were .76, .79, and .80, respectively. An example of an item from the *athletic competence* scale is "Some teenagers do *very well* at all kinds of sports BUT Others *don't* feel that they are very good when it comes to sports." Cronbach's alphas for Grades 8, 9, and 10 were .82, .85, and .86, respectively.

Confidence. Confidence is defined by a composite of two subscales: *positive identity* and *self-worth*. Six items measure positive identity (Theokas et al., 2005) and come from the *Search Institute's Profile of Student Life – Attitudes and Behaviors Survey* (PSL-AB; Benson, Leffert, Scales, & Blyth, 1998). The response format for these six items ranged from 1 = *strongly agree* to 5 = *strongly disagree.* An example of an item used to measure *positive identity* is "On the

whole I like myself." Cronbach's alphas for Grades 8, 9, and 10 were .87, .87, and .88 respectively.

Six items from the *Self-Perception Profile for Adolescents* (SPPA; Harter, 1988) are used to measure *self-worth*. As mentioned above, the SPPS uses a structured alternative response format. Participants are asked to choose between two types of teenagers. Once they have selected which person they are most like, they are asked to decide if it is "really true for me" or "sort of true for me." The items are counterbalanced so that half begin with a positive sentence, reflecting high self-worth, while half begin with a negative sentence, reflecting low self-worth. Each item is scored from 1-4, with 4 reflecting higher perceived self-worth. An example of an item used to assess self worth is "Some teenagers don't like the way they are leading their lives BUT "Other teenagers do like they way they are leading their lives." Cronbach's alphas for Grades 8, 9, and 10 were .80, .74, and .82, respectively.

Connection. To index Connection, 22 of the items from the student questionnaire were used to measure the subscales of *connection to family* (six items), *school* (seven items), *peers* (four items), and *community* (five items). All of the items measuring connection to family, connection to school, and connection to community come from the *PSL-AB* (Benson, Leffert, Scales, & Blyth, 1998).

Five of the items measuring connection to family, six of the items used to measure *connection to school*, and all of the items used to measure connection to community use the forced choice response format ranging from 1=*strongly agree* to 5 = *strongly disagree.* An example of an item measuring *connection to family* is "My parents give me help and support when I need it." An example of an item measuring *connection to school* is "I get a lot of encouragement at my school." An example of an item measuring connection to community is

"Adults in my city or town make me feel important."

The sixth item measuring *connection to family*, "If you had an important concern about drugs, alcohol, or sex, or some other serious issue, would you talk to your parent(s) about it?" uses a forced choice response format ranging from 1 = *yes* to 5 = *no*. The seventh item measuring *connection to school*, "How often do you feel bored at school?" uses a forced choice response format ranging from 1 = *usually* to 3 = *never*. These responses were rescaled to a 1 – 5 point scale. Cronbach's alphas for *connection to family* in Grades 8, 9, and 10 were .88, .88, and .90, respectively; for *connection to school* in Grades 8, 9, and 10 were .84, .84, and .82, respectively; for *connection to the community* in Grades 8, 9, and 10 were .92, .92, and .90, respectively.

The four items used to measure connection to peers come from the *Peer Support Scale* (Armsden & Greenberger, 1987) and use a forced choice response format that ranges from 1 = *always true* to 5 = *almost never true or never true*. An example of an item is "My friends care about me." Cronbach's alphas for Grades 8, 9, and 10 were .95, .95, and .97, respectively.

Character. Eighteen items from the Search Institute's *Profile of Student Life – Attitudes and Behaviors Survey* were used to measure Character for Grades 8, 9, and 10 (PSL-AB; Benson, Leffert, Scales, & Blyth, 1998). These items index the subscales of *Behavioral Conduct*, *Valuing of Diversity*, *Personal Values*, and *Social Conscience*.

For the five items that measure *Personal Values* and the six items that measure *Social Conscience* participants are asked to rate how important each item is in their lives. Response formats range from 1 = *not important* to 5 = *extremely important*. An example of an item measuring *Personal Values* is "Telling the truth, even when it's not easy," while an example of an item measuring *Social Conscience* is "Helping other people." Cronbach's alphas for Grades 8, 9, and 10 were 86, .87, and .87, respectively for *Personal Values*, and .89, .87, and .87,

respectively, for *Social Conscience*.

One of the items used to measure *Valuing of Diversity*, "getting to know people who are of different race than I," uses the same response format as above for measuring importance. The remaining three items used to measure *Valuing of Diversity* and the three items used to measure *Interpersonal Skills* ask participants to think about the people who know them well and how they think they would rate them on each of the items. The response format ranges from 1 = *strongly agree* to 4 = *strongly disagree*. An example of an item used to measure *Valuing Diversity* is "Enjoying being with people who are a different race than I am." Cronbach's alphas for Grades 8, 9, and 10 were .79, .81, and .81, respectively.

Behavioral Conduct was measured by six items from the *Self-Perception Profile for Adolescence* (Harter, 1988). An example of an item from the *Behavioral Conduct* scale is "Some teenagers *usually* do the right think BUT Other teenagers often *don't* do the right thing." Cronbach's alphas for Grades 8, 9, and 10 were .59, .76, and .77, respectively.

Caring. At Grades 8, 9, and 10, five modified items from the *Eisenberg Sympathy Scale* (ESS; Eisenberg, et al., 1996) were used to measure Caring. The items measure the degree to which participants feel sorry for the distress of others. The response format for these items ranged from 1 = *really like you* through 3 = *not like you.* High scores indicate low levels of sympathy. An example of an item from ESS is "I feel sorry for people who don't have the things I have." In addition, four items were adapted from the Empathic Concern Subscale of the *Interpersonal Reactivity Index* (IRI; Davis, 1983). Example items include "It makes me sad to see a person who doesn't have friends" and have the same response format as described above for the *Eisenberg Sympathy Scale*. Davis (1983) reports adequate reliability and validity for this scale. Cronbach's alphas for Grades 8, 9, and 10 were .83, .85 and .84, respectively. For the

structural equation models to be discussed, individual items were randomly combined to form packets in order to enhance reliability. For the nine Caring items, the average of three sets of three items form packets 1, 2, and 3. These three packets contain the same items across Grades 8 to 10.

In short, these aggregated subscales, which were calculated based on the mean of the individual Likert-scale items were then used to represent observed scores. These scores were treated as continuous variables in our analysis.

Testing Invariance of the Measurement Model of PYD over Time

Our research question considers whether the measurement model of the PYD construct is invariant over time. The analysis was conducted by imputing a single augmented covariance matrix (Vandenberg & Lance, 2000). In order to test longitudinal factorial invariance, we followed the recommendations of Vandenberg and Lance (2000) and used the single augmented covariance matrix to input data. The single augmented covariance matrix allows us to merge Pearson's covariance matrices for each time point into a single augmented variance-covariance matrix. Enacting this procedure allows us to take full advantage of the time-structured nature of the data, to address the interdependency among like variables measured at different time points, and to deal with within-time covariances (Vandenberg & Lance, 2000).

A well-fitting configural invariance model suggests that a similar measurement model is plausible across Grades 8 to 10 and that additional measurement invariance tests may proceed. We first examined configural invariance, i.e., whether the same pattern of fixed and free factor loadings was specified for each measurement occasion (Models 1 and 2). Configural invariance needs to be established before conducting any subsequent tests. For each occasion, the first item's factor loading was fixed equal to 1 and its intercept was fixed equal to 0. The freely

estimated factor loadings, factor variances, covariances, and means were allowed to be heterogeneous across time points (Vandenberg & Lance, 2000).

After establishing configural invariance, we pursued metric invariance testing (Model 3). Specifically, we first constrained all the like first-order factor loadings to be equal. After establishing metric invariance at the first-order, we assessed metric invariance at the second-order by constraining corresponding factor loadings to be equal across three measurement occasions (Model 4). The next test represented scalar invariance testing, which involved constraining like observed items' intercepts to be equal across occasions (Model 5). Two more restrictive models were also tested to assess invariant uniqueness (the disturbances of the observed variables, Model 6) and invariant factor variances (the residuals of the first-order factors, Model 7).

Well-established standards for model comparison are lacking in the context of factorial invariance testing (Chen, et al., 2005). However, there are a few guidelines available. For instance, Cheung and Rensvold (2002) proposed using a difference of larger than 0.01 in the CFI to indicate a meaningful change in model fit for testing measurement invariance. In addition, chi-square change is arguably the most frequently used criterion in model selection.

Results

Descriptive statistics, including means and standard deviations are presented in Tables 3 and 4. A set of analyses was conducted to examine if the Five Cs model of PYD that was identified in previous research (e.g., Lerner, et al., 2005; Phelps, et al., 2009) held among eighth to tenth graders as well, and if not, how the model could be modified to accommodate developmental changes that might happen during the middle years of adolescence. The second goal of the current study was to investigate if the Five Cs model of PYD possessed longitudinal

measurement invariance at configural, metric, and scalar levels. In other words, the analyses were employed to determine whether the Five Cs model of PYD was measured the same way over time. Second-order confirmatory factor analyses using LISREL 8.80 were conducted.

We present, first, the descriptive statistics for the items and subscales that were used in the analyses, for both original and imputed data. Next, we report the results of confirmatory factor analysis in each Grade.

Insert Tables 3 and 4 about here

Treatment of Missing Data

As with most longitudinal studies, we experienced two types of missing data: across-occasion attrition and within-time-of-testing item non-response. A series of *t*-tests and Chi-Square tests were conducted to compare longitudinal participants and youth who only provided cross-sectional data. The results indicated that adolescents with complete data did not differ from cross-sectional participants on the demographic variables of sex, annual family per capita income, or mother's education. They did differ in race/ethnicity. Longitudinal participants were more likely to self identify as Asian American, but less likely to identify as European Americans or of multiracial background. In terms of location, longitudinal youth were less likely to be from suburban areas than youth who only provided cross-sectional data.

In longitudinal research, missing data are inevitable, particularly in studies that collect data across transitions in school, such as the shift from middle school to high school. Due to the cohort-sequential design of the 4-H study (Lerner et al., 2005), there are instances of planned missingness. In addition, our data are derived from participants who are in schools and youth

development programs from 30 states across the country. Thus, despite the work of staff to retain participants across all measurement occasions, we still had the problem of missing data. Instead of concentrating on the participants that completed all the three measurement occasions, we imputed missing data by using the expectation maximization (EM) algorithm implemented in PRELIS, that has been found superior to traditional methods such as pair-wise or list-wise deletion (Marsh, Gerlach, Trautwein, Ludtke, & Brettschneider, 2007). Auxiliary variables from each wave were used for imputation of missing subscale data at each wave and included measures of self-regulation and depression that are not included in the present set of analyses (Collins, Schafer, & Kam, 2001). All measurement models were computed using a complete data file of 920 participants across three grades.

Cross-Sectional Confirmatory Factor Analysis Models

To examine if the original PYD measure that was validated in early adolescence still held when youth were in the middle of the adolescent period, we first conducted cross-sectional factor analysis models separately for each of the three time points. Figure 1 provides a schematic representation of the PYD construct first identified by Lerner et al. (2005) and then confirmed by Phelps et al. (2009). According to Lerner et al. (2005), the PYD construct is a second-order factor model that was comprised of five first-order factors of the Cs (competence, confidence, character, caring, and connection). In other words, the Five Cs were hypothesized to be part of a hierarchical structure in which the second-order factor of PYD determines the common variance contained in all the first-order factors. Correlations among the measurement errors for subscales from the Harter SPSS scale (Harter, 1988) were allowed, because they may share method variance not accounted for by the model (Phelps, et al., 2009). Other measurement errors that were allowed to correlate in the Phelps, et al. (2009) model were also permitted.

More specifically, we assessed the fit of the hypothesized model proposed earlier with Grades 8, 9, and 10 data. The results indicated that the theoretically-based second-order factor structure did not result in an adequate fit of the data. More specifically, with Grade 8 data, the hypothesized model provided an inadequate fit, resulting in a chi-square value of 938.65, *df* = 104, RMSEA = 0.093, NNFI = 0.92, CFI = 0.94, GFI = 0.89. The hypothesized model assessed with Grade 9 data also resulted in poor fit: χ^2 = 1090.15, *df* = 104, RMSEA = 0.102, NNFI = 0.94, CFI = 0.95, GFI = 0.88, as did the model assessed with Grade 10 data: χ^2 = 1359.84, *df* = 104, RMSEA = 0.115, NNFI = 0.89, CFI = 0.92, GFI = 0.85. Inspection of the first-order factor loadings of the CFA models identified several non-significant and low pathways. However, all the second order factor loadings were significant and above .30.[2]

Insert Figure 1.

Model Modification Procedure and Revised Model

Two stages of model modification were undertaken. First, we inspected the first- and second-order factor loadings and found that there were several low first-order factor loadings, for example, the factor loadings of athletic competence on the first-order latent factor were 0.08, 0.29, and 0.09 across the three grades, which were all lower than a 0.30 loading cutoff criterion. However, all the second-order factor loadings were significant and ranged from 0.59 to 0.88. Thus, the hypothesized higher-order Five Cs structure was retained. Modification then was focused on revising the measures of the first-order latent factors. More specifically, we examined modification indices and expected changes at the first-factor level. A few modifications were suggested by the program. For instance, modification indices suggested that conduct behavior

might load better on competence or connection, and that self-worth might load on competence. However, we implemented modifications indices for model development only if they met three conditions. First, suggestions made by modification indices were not implemented unless they were theoretically meaningful. Second, the modification indices must have been one of the larger indices. Third, the suggested paths must not have resulted in any change of the estimates in the structural part of the model. After identifying the indices that met these conditions, we pursued a second stage of model modification.

In the second stage of model modification, we created an item set that contains both the items that were originally specified to measure PYD, supplemented by additional items from the Student Questionnaire of the 4-H Study that were assessed by the present authors as theoretically relevant to one of the Five Cs. To make this assessment, we provided 13 current or past graduate student researchers involved in the 4-H Study with operational definitions of each of the Five Cs. These researchers independently rated the scales in regard to whether they reflected one of the Five C constructs. Overall, there were 16 substantive scales with greater than 80% agreement by expert raters as being relevant. The majority of the expert raters (more than 80%) considered self-perceptions of physical appearance as an important indicator of Confidence. Two scales, emotional regulation and job competence, were rated as indicators of competence. Based on the above results from the two stages of model modification, we revised the PYD measure and assessed the goodness of fit of the revised model with Grade 8 cross-sectional data. The revised hypothesized model is depicted in Figure 2.

Insert Figure 2 about here

Testing the Revised Model. The model resulted in fairly good fit with the data: χ^2 (104) = 1094.35, CFI = 0.95, RMSEA = 0.085. All the first- and second-order factor loadings were significant and were greater than 0.40. The revised PYD model was therefore retained and subjected to model-improving modifications. In the final revised model, the disturbance terms for indicators such as values diversity, personal values, and social conscience were allowed to be correlated. Taken together, these modifications significantly improved model fit, χ^2 (99) = 802.50, CFI = 0.97, RMSEA = 0.074. We also tested this final revised model with Grades 9 and 10 cross-sectional data. The results suggest that the revised model adequately fit with Grade 9 data: χ^2 (99) = 969.26, CFI = 0.96, RMSEA = 0.088; and Grade 10 data: χ^2 (99) = 1125.83, CFI = 0.97, RMSEA = 0.066 (Figure 3 presents the revised measurement model with Grade 10 cross-sectional data.).

Insert Figure 3 about here

Testing for Longitudinal Measurement Invariance of the PYD Model

Longitudinal invariance analysis was conducted in several steps. The structural part of the longitudinal model is displayed in Figure 4.

Configural invariance. Using the longitudinal data, we first estimated the revised final model (Model 1) described above, where only the covariances among the errors of the observed variables measured at the same time were released. This model fitted the data fairly well (χ^2 (1161) = 6351.10, CFI =0.94, RMSEA = 0.070. Next, 51 autocovariances between the errors of like variables across the three measurement occasions were estimated (χ^2 (1110) = 3876.75, CFI =0.97, RMSEA = 0.052). These estimates were meant to account for the unique aspect of each

measured variable that is correlated over time. The improvement of the model (Model 2) with the estimated autocovariances to its predecessor was statistically significant (χ^2decreased by 2474.35, CFI increased by 0.03). These results suggest that configural invariance, or weak factorial invariance, was acceptable. In other words, the items loaded on the same first- and second- order factors over time. Table 5 presents overall and comparative fit indexes.

Insert Figure 4 about here

Metric invariance. In order to test metric invariance of the Five Cs model of PYD, we tested whether the first- and second-order factor loadings of like items were invariant across the three measurement occasions. First, we constrained the first-order factor loadings of the like items across occasions to be equal (Model 3). This model fitted the data as well as the previous model, although producing a significant loss of fit in terms of Chi square (χ^2 (1134) = 4237.90. The decrease in CFI was negligible (0.0). Second, we set the like second-order factor loadings to be equal over time (Model 4). The results indicated that this model fitted the data equally as well as its predecessor: Chi square decreased by 205.99, and CFI did not change from the previous model. Overall the results suggest that the like first and second order factor loadings were invariant over time.

Scalar invariance. In the third step of factorial invariance testing, equality of intercepts of observed items across time (Model 5) was imposed on the model to test for scalar (strong) invariance. The analysis of Model 5 yielded satisfactory values on fit indices: χ^2 (1178) = 5008.16, CFI = 0.96, RMSEA = 0.059, CFI = 0.96. The goodness-of-fit of this model was not significantly different from Model 4: ΔCFI = 0.01. Consistent with Cheung and Rensvold's

(2002) recommendations, this model exhibited evidence of invariance across measurement occasions. From these findings, we conclude that the second-order PYD structure, as presented in Figure 3, was operating equivalently over time. The stability coefficients between the second-order factors across three measurement occasions were 0.66, 0.67, and 0.55, respectively.

Insert Table 5 about here

Strict factorial invariance. Finally, equality of disturbances of the observed variables (Model 6) and first-order factor variance (Model 7) across time was further imposed to test strict invariance in Models 6 and 7. Both models had satisfied values on fit indices: χ^2 (1202) = 5542.73, CFI = 0.95, RMSEA = 0.063; χ^2 (1208) = 5539.04, CFI = 0.95, RMSEA = 0.062. The results provide evidence of two strict forms of factorial invariance of the PYD measure across three measurement occasions. Thus, overall, in the present study the PYD measure demonstrated strict factorial invariance across measurement occasions (see Table 5).

Latent factor mean comparison. Because strict invariance was supported, latent factor means across time could be compared. In the strict invariance model, we set the first order factor latent means at Grade 8 to zero and freely estimated the factor means at Grades 9 and 10. The results indicated that relative to the scores at Grade 8, youth scores on competence, confidence, and connection were significantly lower in Grade 9 (-0.58, -0.45, and -0.15, respectively), whereas their scores on caring and character were significantly higher in Grade 10 than in Grade 8 (0.56 and 0.39, respectively). In addition, we tested for latent mean differences related to the higher order factor of PYD. Constraints, including first-order factor loadings, observed intercepts, second-order factor loadings, and first-order latent means, were imposed to test mean

differences related to PYD. These tests revealed statistically significant mean differences in PYD at Grades 8, 9 and 10: relative to Grade 8, youth reported lower levels of PYD in Grade 9, but higher levels of PYD in Grade 10.

Internal consistency and convergent and discriminant validity. The composite reliability and the variance that was associated with each of the five first-order factors and the overall higher-order factor underlying the PYD measure indicated that the PYD measure exhibited both convergent and discriminant validity. The results suggest that the internal consistency of the five first-order factors and the second-order factor were all greater than .70. Furthermore, the findings of the estimated second-order CFA model indicated that all factor loadings were statistically significant, with *t* values ranging from a low of 11.60 to a high of 26.47. These results provide direct evidence of convergent validity of the PYD measure. Most of the squared multiple correlations of the five Cs were greater than the generally accepted value of .50. Values ranged from .36 to .93 at Grade 8, .41 to .96 at Grade 9, and .28 to .92 at Grade 10. The Caring factor consistently explained the smallest proportion of variance over the three years and Connection explained the largest proportion. These results supported the discriminant validity of the PYD measure (Fornell & Larcker, 1981).

Discussion

As the positive youth development perspective is becoming the primary framework for both researchers and practitioners in youth development, there is a pressing need to develop acceptable measures of positive youth development (PYD) (Dukakis et al., 2009). Establishing a valid framework for measuring PYD helps to give guidance to future empirical research and to the evaluation of youth-serving programs seeking to promote PYD. Recent reviews of the literature have presented evidence of empirical support for the validity and robustness of the Five

Cs Model of PYD (Heck & Subramaniam, 2009). This empirical work (Jeličić et al., 2007; Lerner et al., 2005; Phelps et al., 2009) provided information about the structural coherence and development of PYD across the early years of adolescence.

However, as young people enter the middle adolescent period there are shifts in essential developmental tasks and in the contexts in which to accomplish them. Therefore, it is important to determine whether the Five Cs model of PYD remains a useful index of thriving as youth undergo significant life changes. In addition, it is also critical not to assume that tools consistently measure the same constructs over time. In order to accurately measure change and interindividual differences, it is important to examine that the Five Cs Model exhibits measurement equivalence across age.

To address these two concerns, confirmatory factor analyses of PYD were conducted with 4-H Study data from Grades 8-10. We generated a version of the Cs that was slightly revised from the one used in previous work (Phelps et al., 2009). In addition, tests of measurement invariance were conducted via a series of hierarchically nested models to determine whether the measurement model of PYD was invariant over time. Results indicated that one index of Competence, athletic competence, was no longer relevant in assessing PYD during middle adolescence, but that perceptions of physical appearance had become an important indicator of Confidence from Grades 8 to 10. In addition, the overall Five Cs model of PYD can be defined comparably across Grades 8-10, as there was strong evidence of measurement invariance over the three years.

As young people proceed through middle adolescence, the developmental tasks to be achieved vary and measures of achievement and self-concept differ from earlier years. Consistent with the orthogenetic principle (Werner, 1957), the domains of self-concept become

more differentiated and hierarchically integrated across subsequent portions of development (Harter, 1999). Therefore, changes in the salience of different domains for overall health and well-being are expected. It is not surprising that physical appearance is a significant indicator of Confidence in this sample, as prior research has indicated that physical appearance is the self-perception measure most closely related to global self-worth (Harter, 1999; Shapka & Keating, 2005). The relative importance of physical appearance during the middle adolescent period studied in the present research could result from the confluence of physiological, cognitive, and social changes occurring during this period (e.g., the nearing completion of pubertal changes, increasing salience of dating and of romantic relations, and the presence of a cultural emphasis on physical fitness as a marker of attractiveness to others (e.g., Susman & Dorn, 2009). The identification of physical appearance as an indicator of Confidence was also consistent with the ratings of graduate-level researchers when assigning measures to one or more of the Five Cs.

In the Five Cs model, Competence has been defined as a positive view of one's actions in domain-specific areas including social, academic, and athletic. The present research, however, indicates that athletic competence does not load significantly on the Competence factor. This finding is consistent with work indicating that athletic competence is least likely to be correlated with one's overall sense of self (Harter, 1999; Shapka & Keating, 2005). A possible explanation for this finding is that in the large and diverse sample of youth participating in the 4-H Study, each young person is involved in various activities and hobbies outside of the classroom that define who he or she is as a person (Zarrett et al., 2007). Sports participation is common, but it is also usually combined with two or more non-sports activities (Zarrett et al., 2007). However, regardless of what a young person does outside the classroom, academic self concept, grades, and social competence are critical indicators of success to all youth who are still in school, as is

the case with an overwhelming majority of our sample. Therefore, it may be that many of the youth in our sample have devalued the importance of athletic competence in their lives (Harter, 1999; Wylie, 1979) as they become involved in a number of other after-school activities.

Establishing measurement invariance across Grades 8 to 10 allowed us to compare how adolescents fared on the Five Cs and the overall measure of PYD. Results indicated that, compared to Grade 8, youth scores on competence, confidence, and connection were all significantly lower in Grade 9 (and as a result so too were overall PYD scores). These results are consistent with Harter's findings that many of the domains in self concept decrease over the first stages of adolescence and eventually increase during later adolescence (Harter, 1998). As youth transition to new learning environments, experience new social situations, and autonomy becomes an important development goal, many adolescents may begin to doubt their academic and social abilities and, as well, may feel less connected to both parents, peers, and the larger ecological context.

When comparing Grade 10 scores to Grade 8 scores, however, results indicated that youth reported a significant rise in Caring and Character from Grades 8 to 10 (in addition to an overall rise in PYD scores). Similar results were found with an independent high school sample, as scores of character were skewed in the positive direction and adolescents scored higher than children on several indicators of character (Park, 2004). Character and Caring may grow during adolescence as youth begin to describe themselves in morally relevant terms, speaking of noble purposes, such as caring for others, as definitive to their identity (Damon, 2004). In future research, it will be important to consider the contextual-level factors (family, peers, school, and community) that may covary with changes in the Five Cs in order to identify the variables that

can be addressed in practice and policy (e.g., character education programs) to better foster healthy growth in all young people.

The present research addresses limitations present in both the empirical and applied realms of PYD. Although the PYD framework has become more popular among academics and practitioners, the field has yet to agree upon a set of indicators of positive development that spans research, policy, and practice (Moore et al., 2004). Earlier work (Lerner et al., 2005; Phelps et al., 2009) has already established the existence of a valid measure of PYD across early adolescence. The present work extends the validity and utility of the measure to an older range of adolescents. Researchers examining the point of development when young people transition to high school now have a valid measure to assess whether youth are developing positively. This work also affords researchers the ability to consider the relationships among PYD and other indicators of both positive and negative development during middle adolescence. In addition, this work moves the field closer to a shared vocabulary and to a widely-applied tool to measure PYD; our research meets the call for an effective PYD framework that is valid and useful for a diverse population of adolescents (Heck & Subramaniam, 2009) and that aids in the application of effective, age-appropriate, interventions and programs for youth.

The methodology of this study had, however, had its own limitations that require further investigation. First, although our results suggest 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 Phelps et al., (2009) sample in terms of sample size. It is possible that the fact that the original measurement structure does not fit middle adolescents is due to the fact that our model testing was based on a different sample.

It is useful to note that our revision of the indicators of two of the Five Cs (Competence and Confidence) reflects a strength of the Five Cs model and of the 4-H Study, as well as the dynamic nature of longitudinal research itself (Phelps & Colby, 2002). As individuals develop, changes within and beyond the individual lead to changes in how successful development is defined; changes in how to measure these definitions is self-evident. While our research is of course limited by our sample, measurement model, and methodology, the support we have provided for the present conceptualization of PYD enhances our understanding of positive development, and of how it might be studied longitudinally and used in application.

Establishing measurement invariance, as we have done in the present study, accomplishes a critical step for further research and practice using the Five Cs model of PYD. Unless measurement invariance is established, performing cross-group comparisons of mean difference, regression coefficients, or other parameters is meaningless (Schmitt & Kuljanin, 2008). As the PYD perspective is adopted in more youth-serving programs, the need for a measure that can be utilized in different contexts becomes paramount. Future work will continue to examine the individual and contextual bases of healthy development that form the crux of the developmental systems approach to PYD (Lerner 2004), as the youth in the 4-H Study progress through high school. The key hypothesis of the PYD perspective is that when the strengths of youth are aligned with the strengths of the context, positive and healthy development will occur; therefore, a valid and robust measure of PYD across early, middle, and late adolescence is needed in order to identify the factors that promote thriving in a diverse group of youth. Our goal is to continue evaluating and potentially revising the PYD measure across adolescence, and thereby continue to offer a tool for research and application.

References

Armsden, G., & Greenberg, M. (1987). The Inventory of Parent and Peer Attachment: Individual differences and their relationship to psychological well-being in adolescence. *Journal of Youth and Adolescence, 16,* 427-454.

Arnett, J.J. (1997). Young people's conceptions of the transition to adulthood', Youth *and Society 29(1)*, 3—21.

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*, 138–159.

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

Byrne, B. M. (1995). Strategies in testing for an invariant second-order factor structure: A comparison of EQS and LISREL. *Structural Equation Modeling, 2,* 53–72.

Byrne, B. M.,&Campbell, T. L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. *Journal of Cross-Cultural Psychology, 30,* 555–574.

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling, 12*(3), 471-492.

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

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*, 330–351.

Damon, W. (2004). What is positive youth development? *Annals of the American Academy of Political and Social Science, 591,* 13-24.

Davis, M. H. 1983. Measuring individual differences in empathy: Evidence for a multidisciplinary approach. *Journal of Personality and Social Psychology, 44,*113–126.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences, 33*(4), 533-553.

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.

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, England: Norton & Co.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research, 18*(3), 382-388.

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

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

Harter, S. (1998). The development of self-representations. In W. Damon (editor-in-chief) & N. Eisenberg (Ed.), *Handbook of Child Psychology: Vol 3. Social, emotional, and personality development* (5th ed., pp.1097-1149). New York: Wiley.

Harter, S. (1999). *The construction of self: A developmental perspective*. New York, NY: Guilford Press.

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

Hills, P. & Argyle, M. (2002). The Oxford Happiness Questionnaire: A compact scale for the measurement of psychological well-being. *Personality and Individual Differences, 33*(7), 1071-1082.

Jeličić, H., Bobek, D., Phelps, E., 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.

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

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., Gestsdottir, 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.

Marsh, H. W., Gerlach, E., Trautwein, U., Ludtke, O., & Brettschneider, W. D. (2007). Longitudinal study of preadolescent sport self-concept and performance: Reciprocal effects and causal ordering. *Child Development, 78*(6), 1640-1656.

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.

Park, N. (2004). The role of subjective well-being in positive youth development. *The Annals of the American Academy of Political and Social Science, 591*, 25-39.

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

Phelps, E., & Colby, A. (2002). Introduction. In E. Phelps, R. R. Furstenberg, Jr., & A. Colby (Eds)., *Looking at Lives: American Longitudinal Studies of the Twentieth Century*. New York: Russell Sage Foundation, pp. 3-14.

Phelps, E., Zimmerman, S., Warren, A. 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 Public/Private Ventures (Ed.), *Youth development: Issues, challenges, and directions* (pp. 17–64). Philadelphia: Public/Private Ventures.

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

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

Shapka, J. D., & Keating, D. P. (2005). Structure and change in self-concept during adolescence. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 37*(2), 83-96.

Small, S. A. & Rodgers, K. B. (1995). *Teen Assessment Project (TAP) Survey Question Bank*. Madison, WI: University of Wisconsin-Madison.

Susman, E.J., & Dorn, L.D. (2009). Puberty: Its role in development. In R. M. Lerner, & L. Steinberg (Eds.), *Handbook of Adolescent Psychology: Vol 1. Individual bases of adolescent development,* (3rd ed., pp.116-151). Hoboken, NJ: Wiley.

Theokas, C., Almerigi, J. B., Lerner, R. M., Dowling, E. M., Benson, P. L., Scales, P. C., & von Eye, A. (2005). Conceptualizing and modeling individual and ecological asset components of thriving in early adolescence. *Journal of Early Adolescence, 25*(1), 113-143.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-69.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D. Harris (Ed.), The concept of development: An issue in the study of human behavior (pp. 125–148). Minneapolis: University of Minnesota Press.

Wylie, R. C. (1979). *The self-concept: Vol II. Theory and research on selected topics* (Rev. ed.). Lincoln, NB: University of Nebraska Press.

Zarrett, N., Peltz, J.S., Fay, K., Li, Y., Lerner, R. M., & Lerner, J. V. (2007). Sports and youth development programs: Theoretical and practical implications of early adolescent participation in multiple instances of structured out-of-school (OST) activity. *Journal of Youth Development, 2*(1), 7-20.

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.

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

Table 2

Sample characteristics for Longitudinal Youth from Grades 8, 9, and 10 of the 4-H Study of PYD

	Grade 8	Grade 9	Grade 10	Total Sample
Number of students	736	750	611	920
Age (mean, sd)	14.19 (0.99)	15.04 (1.04)	16.05 (0.95)	--
Female (n, %)	463 (62.9%)	457 (60.9%)	390 (63.8%)	567 (61.6%)
Annual family per capita income (mean, sd)	$16812 (10416)	$16540(10844)	$16989(11049)	$16543(10486)
Mothers' education, in years (mean, sd)	14.14 (2.45)	14.00 (2.47)	14.31 (2.36)	14.12 (2.43)
Participant's geographic location	Grade 8	Grade 9	Grade 10	Total Sample
Northeast: CT,DE,MA,MD,NH,NJ,NY,PA,RI	148 (20.1%)	143 (19.1%)	80 (13.1%)	157 (17.1%)
North Central: IL, KS, MI, MN, MO, OH, WI	230 (31.2%)	298 (39.7%)	283 (46.3%)	362 (39.3%)
Southeast: AL, FL, KY, NC, SC, TN, TX,VA	192 (26.1%)	179 (23.9%)	110 (18.0%)	210 (22.8%)
West: AZ, CO, ID, MT, OR, WA	166 (22.6%)	130 (17.3%)	138 (22.6%)	191 (20.8%)
Race/Ethnicity	Grade 8	Grade 9	Grade 10	Total Sample
African American	53 (7.2%)	63 (8.4%)	50 (8.2%)	69 (7.5%)
Asian American	27 (3.7%)	28 (3.7%)	17 (2.8%)	31 (3.4%)
American Indian	9 (1.2%)	12 (1.6%)	4 (0.6%)	12 (1.3%)
European American	494 (67.1%)	480 (64.0%)	424 (69.4%)	619 (67.3%)
Latino/a American	49 (6.7%)	67 (8.9%)	47 (7.7%)	74 (8.0%)

Multiracial	9 (1.2%)	9 (1.2%)	9 (1.5%)	11 (1.2%)
Inconsistently reported	94 (12.8%)	89 (11.9%)	59 (9.6%)	102 (11.1%)
Missing	1 (0.1%)	2 (0.3%)	1 (0.2%)	2 (0.2%)
Location	Grade 8	Grade 9	Grade 10	Total Sample
Urban	354 (48.1%)	382 (50.9%)	268 (43.9%)	429 (46.6%)
Suburban	95 (12.9%)	81 (10.8%)	66 (10.8%)	113 (12.3%)
Rural	281 (38.2%)	282 (37.6%)	275 (45.1%)	372 (40.4%)
Missing	6 (0.8%)	5 (0.7%)	1 (0.2%)	6 (0.7%)

Table 3

Descriptive Statistics for Subscales Before Imputation

	Grade 8		*Grade 9*		*Grade 10*		*Possible Range*
	N	Mean (sd)	N	Mean (sd)	N	Mean (sd)	
Confidence							
Positive identity (6 items)	736	8.92 (2.19)	750	8.92 (2.23)	611	8.96 (2.20)	0-12
Self-worth (6 items)	736	9.02 (2.62)	750	7.86 (2.72)	611	8.60 (2.50)	0-12
Competence							
Academic (6 items)	736	8.52 (2.68)	750	7.52 (2.86)	611	8.46 (2.66)	0-12
Grades (1 item)	736	9.86 (2.28)	750	9.28 (2.60)	611	9.56 (2.65)	0-12
Social (6 items)	736	8.74 (2.58)	750	7.84 (2.72)	611	8.60 (2.59)	0-12
Athletic (6 items)	736	6.83 (2.98)	750	6.25 (2.80)	611	7.07 (2.81)	0-12
Character							
Personal values (5 items)	736	9.14 (2.30)	750	9.21 (2.26)	611	9.55 (2.31)	0-12
Social conscience (6 items)	736	8.77 (2.49)	750	8.78 (2.44)	611	9.07 (2.45)	0-12
Values diversity (4 items)	736	7.93 (2.46)	750	7.90 (2.51)	611	7.99 (2.57)	0-12
Conduct behavior (6 items)	736	7.71 (2.15)	750	7.50 (2.74)	611	8.16 (2.52)	0-12
Connection							
Family (6 items)	736	9.31 (2.38)	750	8.84 (2.59)	611	9.28 (2.59)	0-12
Community (5 items)	736	6.56 (2.60)	750	6.41 (2.63)	611	6.52 (2.71)	0-12
Peer (4 items)	736	9.46 (2.54)	750	10.75 (1.99)	611	9.67 (2.35)	0-12
School (7 items)	736	7.82 (2.07)	750	7.69 (2.08)	611	7.96 (2.00)	0-12
Caring							

Caring 1	736	9.09 (2.37)	750	9.05 (2.37)	611	9.64 (2.22)	0-12
Caring 2	736	8.21 (2.35)	750	8.19 (2.28)	611	8.49 (2.41)	0-12
Caring 3	736	8.74 (3.07)	750	8.69 (2.94)	611	8.95 (2.88)	0-12

Table 4

Descriptive Statistics for Subscales after Imputation

Imputed samples	*Grade 8*		*Grade 9*		*Grade 10*		*Possible Range*
	N	Mean (sd)	N	Mean (sd)	N	Mean (sd)	
Confidence							
Positive identity (6 items)	920	8.90 (1.97)	920	8.90 (2.08)	920	9.13 (1.82)	0-12
Self-worth (6 items)	920	8.99 (2.35)	920	7.51 (2.60)	920	8.35 (2.08)	0-12
Competence							
Academic (6 items)	920	8.20 (2.50)	920	7.12 (2.74)	920	8.19 (2.23)	0-12
Grades (1 item)	920	9.63 (2.10)	920	9.09 (2.47)	920	9.56 (2.18)	0-12
Social (6 items)	920	8.68 (2.31)	920	7.49 (2.59)	920	8.48 (2.13)	0-12
Athletic (6 items)	920	6.96 (2.68)	920	6.02 (2.60)	920	7.24 (2.31)	0-12
Character							
Personal values (5 items)	920	9.32 (2.10)	920	9.28 (2.12)	920	9.88 (1.95)	0-12
Social conscience (6 items)	920	8.98 (2.28)	920	8.89 (2.30)	920	9.40 (2.06)	0-12
Values diversity (4 items)	920	7.93 (2.24)	920	8.05 (2.35)	920	8.38 (2.17)	0-12
Conduct behavior (6 items)	920	7.68 (1.93)	920	7.15 (2.62)	920	7.83 (2.12)	0-12
Connection							
Family (6 items)	920	9.32 (2.16)	920	8.66 (2.48)	920	9.37 (2.12)	0-12
Community (5 items)	920	6.48 (2.39)	920	6.37 (2.41)	920	6.51 (2.22)	0-12
Peer (4 items)	920	9.31 (2.33)	920	10.45 (2.01)	920	9.30 (2.01)	0-12
School (7 items)	920	7.80 (1.86)	920	7.67 (1.94)	920	7.72 (1.68)	0-12
Caring							

Caring 1	920	9.10 (2.14)	920	9.04 (2.20)	920	9.96 (1.89)	0-12
Caring 2	920	8.04 (2.14)	920	8.21 (2.10)	920	8.64 (1.98)	0-12
Caring 3	920	8.74 (2.77)	920	8.77 (2.72)	920	8.97 (2.35)	0-12

Table 5

Fit Indices and Fit Change for Measurement Model 1 Through 7

	χ^2	df	RMSEA	GFI	CFI	$\Delta\chi^2(\Delta df)$	ΔCFI
M1: Base model	6351.10	1161	0.070	0.79	0.94	--	
M2: Correlated errors of like items	3876.75	1110	0.052	0.86	0.97	2474.35	0.03
M3: 1st-order loading invariant	4237.90	1134	0.055	0.85	0.97	361.15	0.00
M4: 2nd-order loading invariant	4031.91	1144	0.052	0.85	0.97	-205.99	0.00
M5: 1st-order observed intercept invariant	5008.16	1178	0.059	0.84	0.96	976.25	-0.01
M6: observed disturbance invariant	5542.73	1202	0.063	0.84	0.96	534.57	0.00
M7: first-order factor variance invariant	5539.04	1208	0.062	0.83	0.95	-3.69	-0.01

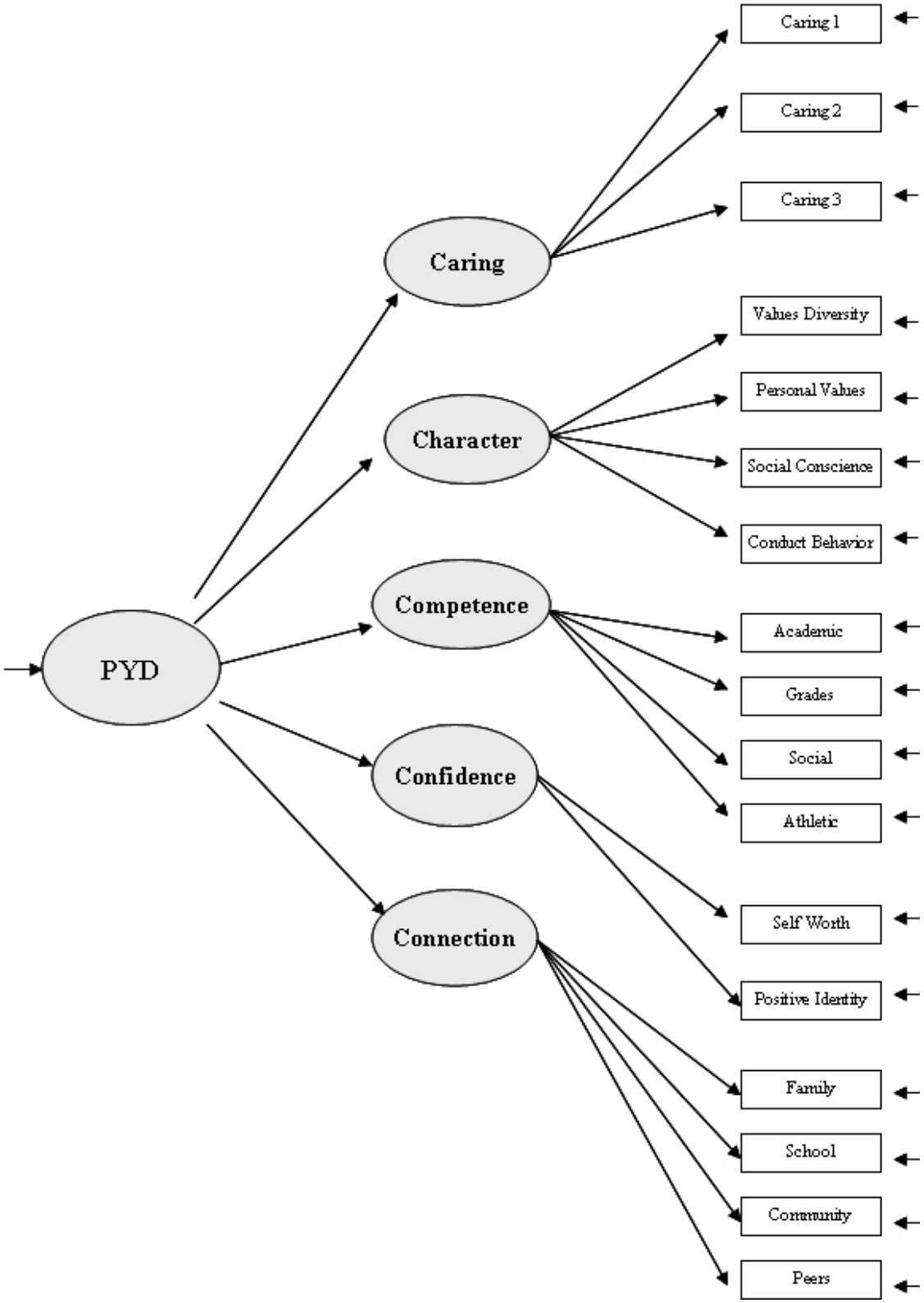


Figure 1. Hypothesized model of Five Cs model of PYD proposed in Phelps, et al. (2009)

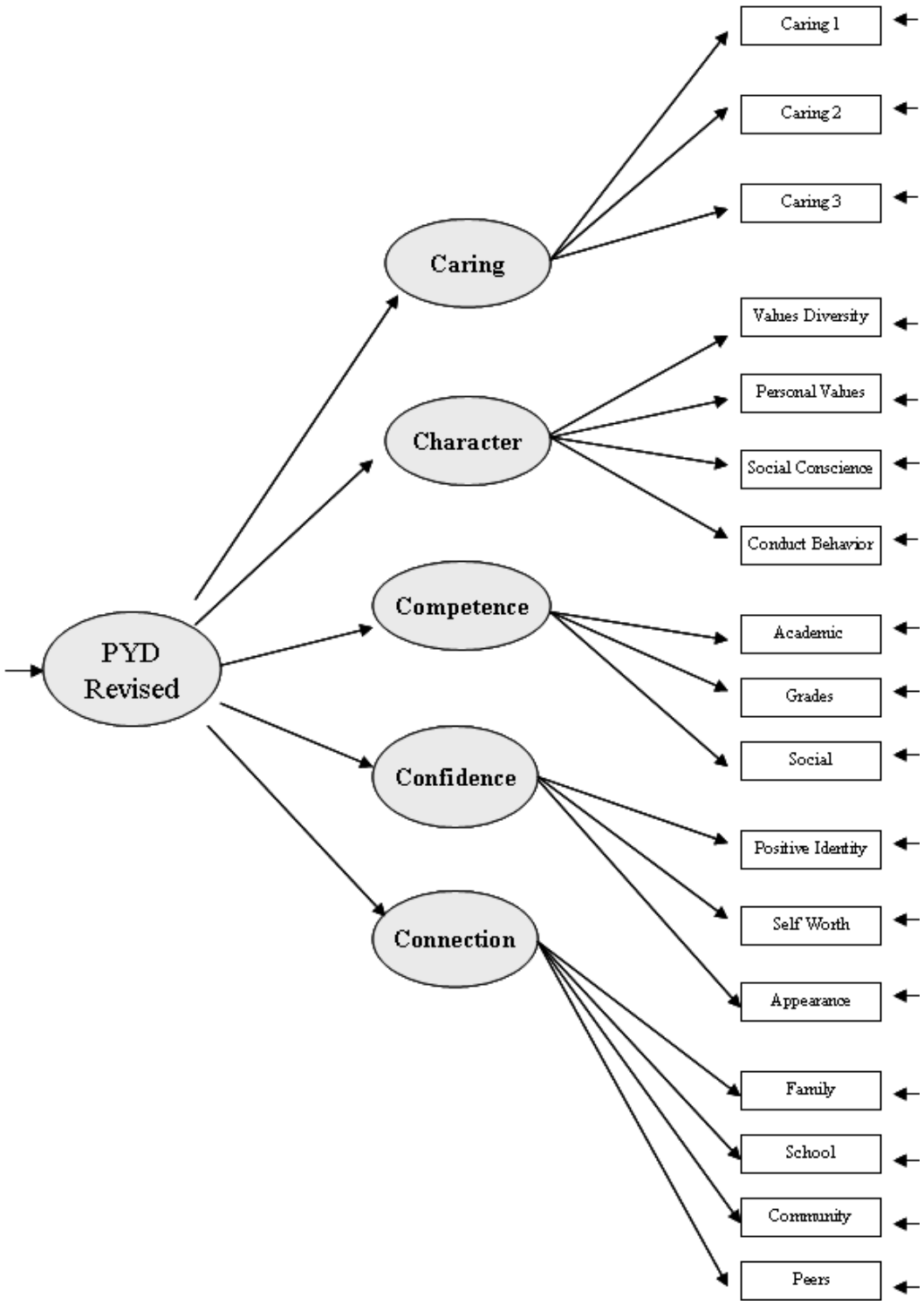
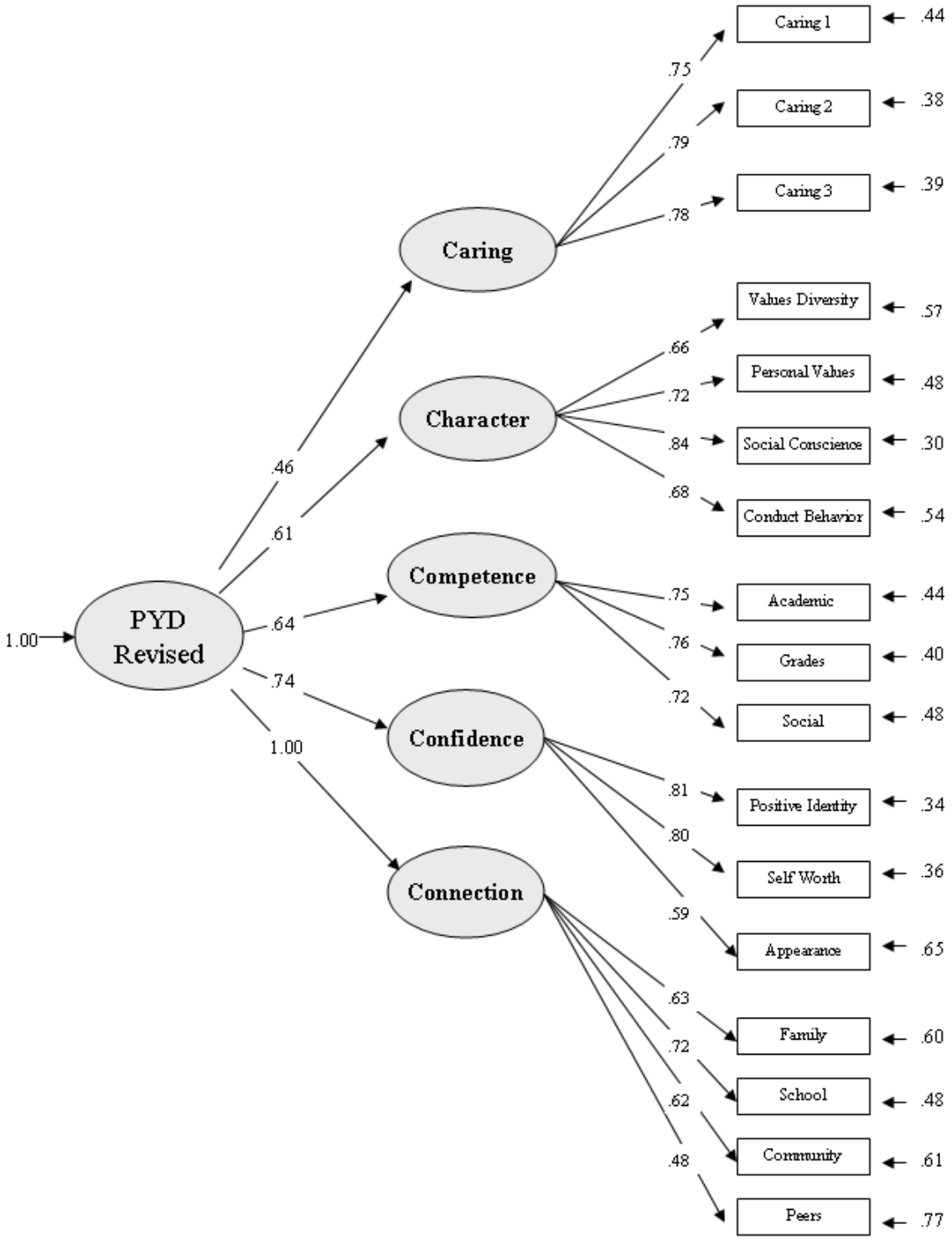


Figure 2. Revised conceptual model of Five Cs of PYD for Grade 10 youth
Note. Athletic Competence was removed from revised model and Physical Appearance now loads on Confidence.



$\chi 2$ = 1125.83, df = 99, p-value < .001, RMSEA = 0.066

Figure 3. Grade 10 PYD measurement model with standardized parameter estimates
Notes: Several residual variances were correlated, but not included in the figure for clarity. Model parameter estimates for Grades 8 and 9 data are available upon request.

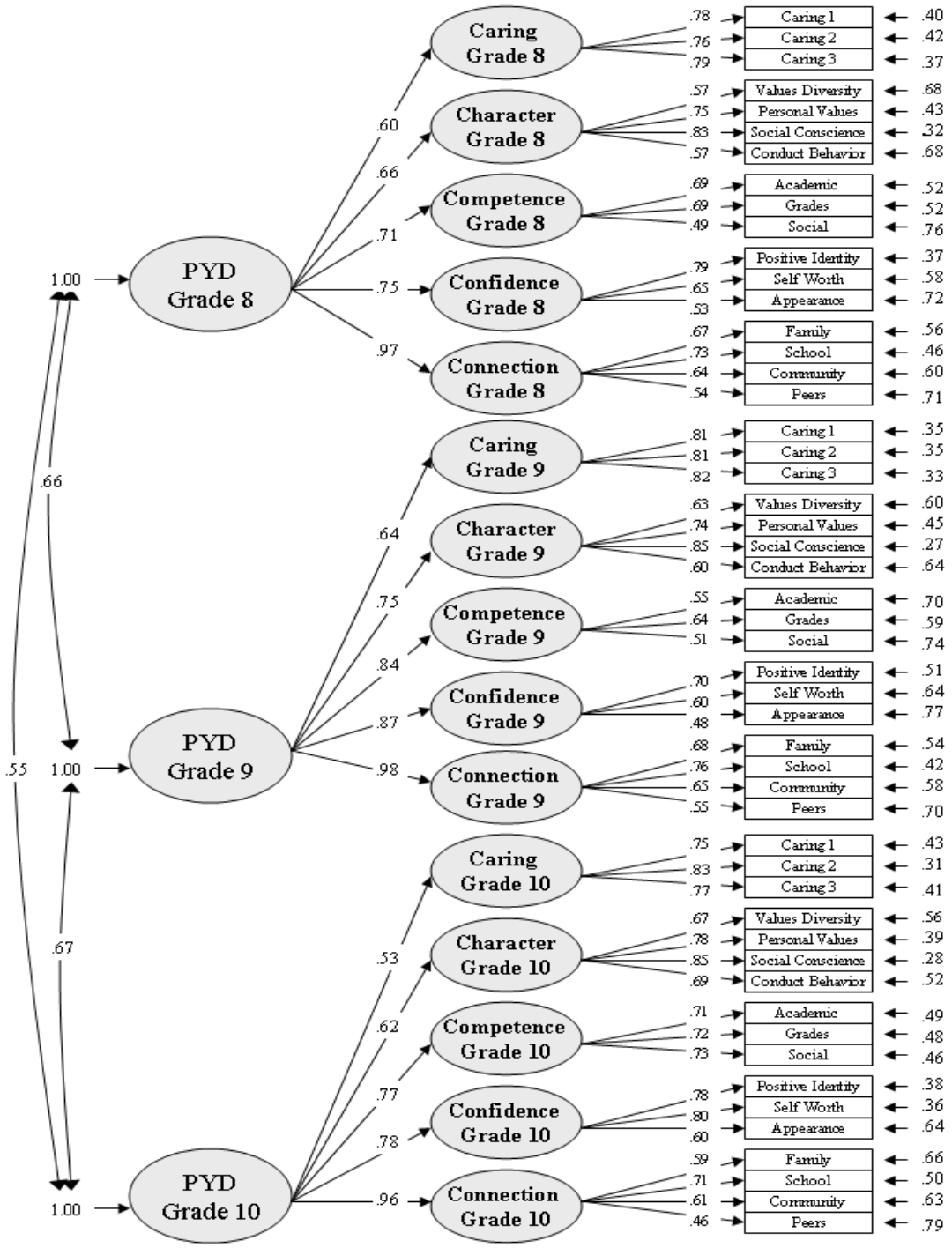


Figure 4. Longitudinal PYD measurement model for Grades 8, 9, and 10, with standardized parameter estimates.
Note: Several residual variances within time and all across-time residual variances for like items were correlated, but not included in the figure for clarity.