

## **Longitudinal Research About, and Program Evaluations of, Positive Youth Development in Low- and Middle-Income Countries: Methodological Issues and Options**

**Richard M. Lerner**

*Tufts University  
richard.lerner@tufts.edu*

**Jonathan M. Tirrell**

*Tufts University  
jonathan.tirrell@tufts.edu*

**Patricia K. Gansert**

*Boston College  
pgansert@gmail.com*

**Jacqueline V. Lerner**

*Boston College  
jacqueline.lerner@bc.edu*

**Pamela Ebstyne King**

*Fuller Theological Seminary  
pamking@fuller.edu*

**G. John Geldhof**

*Oregon State University  
john.geldhof@oregonstate.edu*

**Elizabeth M. Dowling**

*Tufts University  
elizabeth.dowling@tufts.edu*

**Alistair T. R. Sim**

*Compassion International  
alsim@us.ci.org*



New articles in this journal are licensed under a Creative Commons Attribution 4.0 License. This journal is published by the [University Library System](#), [University of Pittsburgh](#) and is cosponsored by the [University of Pittsburgh Press](#). The Journal of Youth Development is the official peer-reviewed publication of the [National Association of Extension 4-H Youth Development Professionals](#) and the [National AfterSchool Association](#).

## Abstract

*The study of positive youth development (PYD) requires theory-based methodological considerations pertinent to measurement, research and program design, and data analysis. We outline the appropriate steps that researchers and program evaluators must enact to address these methodological foci in their respective attempts to describe, explain, and optimize the course of positive development among diverse youth around the world. We focus on longitudinal (developmental) research designed to evaluate programs promoting PYD in low- and middle-income countries (LMICs), where life challenges are shaped by multiple adverse situations associated with racism, poverty, gender inequalities, political inequities, and the absence of adequate health and medical resources. Using the Compassion International Study of PYD as a sample case, we suggest how researchers and practitioners might collaborate to enact rigorous, theory-based research aimed at promoting PYD among youth living in LMICs and worldwide.*

Key words: positive youth development, methodology, longitudinal research designs, low- and middle-income countries, Compassion International

Interest in studying and enacting programs aimed at promoting positive youth development (PYD) has burgeoned across the first two decades of the 21<sup>st</sup> century (e.g., Arnold & Gagnon, 2020; Benson, 2008; Catalano et al., 2002; Damon, 2004, 2008; Eccles & Gootman, 2002; Larson, 2000; Lerner et al., 2015, 2018, 2021; Roth & Brooks-Gunn, 2003). This interest has been instantiated in the formulation of different models of the developmental process involved in youth thriving (see reviews in Arnold & Gagnon, 2020; Lerner et al., 2015, 2021). In addition, in the United States and internationally, and prominently in regard to youth development within low- and middle-income countries (LMICs), a range of measures, research designs, and data analysis methods have been used in the enactment and evaluation of thousands of community-based programs aimed at promoting PYD (e.g., see Banati, 2021; Lansford & Banati, 2018; Petersen et al., 2017).

The goal of this article is to advance research↔practice integration in the international study and promotion of PYD by presenting theory-based methodological considerations pertinent to measurement, the design of research, PYD programs, and program evaluations, and to data analyses required in research or program evaluations. Accordingly, we discuss the appropriate steps that researchers and program evaluators must enact to address these methodological foci in their respective attempts to describe, explain, and optimize the course of positive development among diverse youth around the world. Of course, space limitations preclude our going beyond initial discussion of the methodological points we make. Throughout the article, we provide references to more extended discussions of these ideas. As well, if there is an interest in having access to additional discussion of the methodological issues we introduce, and/or in having further examples of the implications of these issues for research↔practice integration, especially for programs or program evaluations in LMICs that are pertinent to PYD,

it is useful to access a December 2020 webinar presented by Richard M. Lerner, Marc H. Bornstein, and Elizabeth M. Dowling to grantees of the Templeton Word Charity Foundation.<sup>1</sup>

## The Concept of Development Within the Concept of PYD

The use of the term *development* in the concept of PYD, as compared to, for instance, the term behavior (i.e., positive youth behavior), is important to underscore. The concept of development is not equivalent to the concept of change (Lerner, 2018). Whereas development always involves change, not all changes are developmental ones. Overton (2015) explains that development involves transformational (qualitative) change and not only variational (quantitative) change. The concept of development invokes references to changes that involve the person being different in kind (e.g., as reflected in new ways of thinking, feeling, or behaving, alterations in the constitution of an individual that involve something different existing in the person and not just more or less of the same thing). For instance, Piaget (1970) described differences in the nature of thought (moving from sensorimotor intelligence to pre-operational, concrete operational, and finally formal operational thought), and Kohlberg (1978) described moral reasoning as changing from pre-conventional, to conventional, and finally to post-conventional reasoning. The new features of a person that are involved in the emergence of the qualitative differences that comprise development are always based on specific theoretical ideas about this variation (Lerner, 2018).

In addition, it is important to recognize that use of the term development in PYD is not only a theoretical nicety. It is also a conceptual step toward understanding that studying the development of youth has important methodological implications. That is, when the concern of researchers or program evaluators is understanding how to enhance or promote better outcomes associated with the development of specific facets of a young person (e.g., character, community contributions, positive peer relationships), longitudinal designs *must* be employed. Simply, repeated assessments of the young person must be taken across the specific times in development when theory suggests that developmental transformations will occur. As we shall

---

<sup>1</sup> Lerner, R. M., Bornstein, M. H., & Dowling, E. M. (2020). Building Programs, Evaluations, and Tools to Promote Human Flourishing in LMICs: Unmasked Questions and Welcomed Answers. (Video). [https://us02web.zoom.us/rec/share/uiKv9MZr0NFInIzhbp5YXYVKWht8RPIQGzJrt6Y4E5o22pjUvGoB-A\\_SJFKMWaww.PBR-IGW15R-J\\_IR](https://us02web.zoom.us/rec/share/uiKv9MZr0NFInIzhbp5YXYVKWht8RPIQGzJrt6Y4E5o22pjUvGoB-A_SJFKMWaww.PBR-IGW15R-J_IR) Passcode: !+Gg3AK3

### Longitudinal Research of PYD: Issues and Options

explain in more detail later in this article, the selection of when in life a young person is studied should not be a matter of convenience to a researcher or program evaluator.

In developmental research, the  $x$ -axis is most often some index of time, and the points in time used in a developmental study should be selected based on when transformations in the quality of thinking, feeling, behaving, etc., should occur. Demonstrating that a young person scores higher or lower on a measure of something that already existed in their repertoire of skills or attributes (e.g., changes in scores that reflect greater knowledge of geometry or algebra after the administration of an intervention program) may be important for managing behavior or to improving academic test scores; however, such quantitative change in and of itself says little, if anything, about whether development has occurred. Unless changes in such math scores were coupled with evidence of cognitive transformations in understanding math from a conceptual perspective, as compared to just a procedural or computational one, inferring that a developmental change had occurred would be questionable at best.

Therefore, whereas theoretically timed observations across the life of a youth are needed to conduct a useful longitudinal study of PYD, there are several other methodological issues that must be addressed to appropriately and productively use longitudinal designs to illuminate the developmental changes—the developmental process—involved in PYD. The Compassion International (CI) Study of PYD is a longitudinal research and program evaluation project that, at this writing, is ongoing in El Salvador, Rwanda, and—when international travel, now constrained by the COVID-19 pandemic, is able to again begin—will be extended to Uganda (e.g., Lerner et al., 2019; Tirrell et al., 2020, 2021; Tirrell, Gansert, et al., 2019; Tirrell, Geldhof, et al., 2019). Using the CI Study of PYD as a sample case, we will discuss these methodological issues and suggest options for appropriately addressing them. Our discussion of these issues begins with what, not surprisingly, given what we have already said about the theoretical nature of the concept of development, is the foundational tool of developmental methodology.

### **Theory is the Foundational Tool of Developmental Methodology**

Discussions of developmental methodology emphasize that theory should guide all decisions about methods to be used to study developmental change (e.g., Card, 2017; Collins, 2006; Little et al., 2009). Theory should guide the selection of what to measure, at what points in life (the  $x$ -axis points) to use these measures, and how to best analyze data collected from use of these measures to assess what, if any, developmental changes have occurred. Because the

Longitudinal Research of PYD: Issues and Options

study of development involves the assessment of changes within a person across life and, as well, differences between people in their respective courses of within-person change, development can be studied only when measures are used at repeated times across periods of life. However, as we have explained, theory must be used in regard to the selection of these time points.

As documented in the chapters of *Sustainable Development Across the Life Course: Evidence from Longitudinal Research*, edited by Prerna Banati (2021), repeated measures of youth must be taken at points in their lives during which there is a theoretical basis for expecting either changing features of the process of development to occur *and/or* optimal times in adolescence within which to enact interventions to enhance the course of positive developmental change. In both cases, theory instructs researchers to focus on periods within the lives of youth when there is reason to believe that the transformational changes (and not only variational change; Overton, 2015) involved in the process of positive development will or can be induced to occur.

However, there are still other reasons why theory is the primary tool for research and program evaluation pertinent to PYD. First, PYD research and PYD program evaluation have a common theoretical base. Both involve a strengths-based (non-deficit) conception of youth development (Lerner, 2018), one that is embedded within dynamic, relational developmental systems-based concepts (Overton, 2015). In other words, PYD emphasizes what can go right with young people (e.g., the development of valued academic and life skills, citizenship, employment), and not what can go wrong (e.g., substance abuse, depression, incarceration). The key idea in all models of PYD is that, when the strengths of youth are aligned across time and place with contextual resources (e.g., positive relationships with a caring, committed, competent, and consistently available adult; life-skills-development opportunities; and opportunities to use these skills in valued activities; Tirrell et al., 2020), mutually influential and supportive youth-context relations (represented as youth ↔ context relations) will develop and youth will then thrive (e.g., Lerner et al., 2015). This perspective emphasizes that the individual strengths of youth and the resources around them are both important but, what is even more important is that the individual strengths (e.g., interest in specific academic skills or subject areas, a sense of spirituality that motivates a youth to work to make positive contributions to the world) and community resources (e.g., the presence of a caring, competent, and committed adult who provides academic or spiritual support or guidance, respectively) are well aligned (i.e., create a goodness of fit) across time and place.

### Longitudinal Research of PYD: Issues and Options

Second, both researchers and youth-program practitioners recognize that assessing the relations between young people and their contexts, either to understand the dynamics of these relations or to evaluate if and how program practices are enhancing PYD, engages specific instantiations of the general issues of measurement, design, and data analysis to which we have already pointed. That is, how should PYD and the actual (ecologically valid) contexts (programs, families, peer-group relationships) influencing this development be measured? How should observations be designed to identify the crucial relations that occur across time between a young person and the context that promote PYD? What are the appropriate data analysis methods needed to ascertain if changes in positive functioning have developed, and if such changes are associated with changes in the contexts of youth? Indeed, because development is an instance of change in the individual *and* in the individual's relations with the context, measures, designs, and data analyses must all be able to detect changes if, in fact, they have occurred in youth, in contexts, and in youth↔context relations (e.g., see Cantor et al., 2021).

In short, using theory to illuminate the specific individual, context, and individual↔context relations that comprise the PYD process leads to sets of methodological issues involving all other facets of methodology. In past, ongoing, and planned work involved in the CI Study of PYD, we have addressed all these issues. Accordingly, in turning to a discussion of these methodological issues, we first delineate the nature of these issues and then discuss the ways in which the CI Study of PYD illustrates the use of one or more options available to address them.

### **Measuring PYD: Measuring Variables Versus Measuring People**

The study of human development entails describing, explaining, and optimizing changes within a person across the life span and differences between people in their individual changes (Baltes et al., 1977; Lerner, 2018). In studying the development of facets of PYD, for example, positive character or *character virtue* development, a developmental analysis should seek to ascertain changes within a specific individual young person on their scores for character virtue development across the second decade of life and, as well, differences between their pathway of changes in character and the pathways of other young people.

Of note in this example is that the focus of a developmental approach is on people and what transpires in them across time and place. Notice as well that another question can be asked about the character virtue component of PYD: How do scores for character virtues collected at one time in adolescence relate to scores for character virtues collected at a subsequent time of

Longitudinal Research of PYD: Issues and Options

adolescence? If a researcher or program evaluator was interested in this question, they might ask: Are the scores for character virtues among a group of adolescents at the beginning of an intervention program the same or different from the scores at the end of the program? These questions about character virtues necessarily involve the computation of group-based statistics (e.g., means [or averages] and standard deviations) and, as such, these statistics are not about individual people or their development. Indeed, an average score for a variable calculated from a group does not reflect the developmental status of any one specific individual within that group but, instead, reflects the status of a particular variable within the group (see Rose, 2016). Therefore, these questions are about a variable, in this case, the character virtue component of PYD, and not about the manifestation of character virtues within a specific young person.

It is certainly fine to ask questions about variables. However, if a project is about development—if, say, a youth-serving professional wishes to change the course of character virtue development—then the researcher needs to study *people* and not variables. Studying variables provides averages or other group statistics that may not represent the developmental pathway of any of the individuals whose data are used to compute an average, a standard deviation, or a correlation (Molenaar & Nesselroade, 2012, 2014; Nesselroade & Molenaar, 2010; Rose, 2016). As such, if a developmental study focuses on variables, researchers have the challenge of ascertaining if and how their findings apply to the individuals in their samples (e.g., see Bornstein, 2017, 2019; Lerner & Bornstein, 2021).

Molenaar and Nesselroade (2015) have devised a methodological means of assessing whether, in a specific set of data, it is possible to move from *person-specific (idiographic)* information about the individuals in a sample to aggregates of individuals (i.e., subgroups within the sample or even to the overall sample). They term this method the idiographic filter (Molenaar & Nesselroade, 2015). We will have reason to return to this methodological tool later in this article.

Here, however, whether variables or individual people are the initial empirical focus of developmental research, there are several features of measurement that require attention. These features involve change sensitivity and the psychometric quality of measures.

Regarding change sensitivity, it may seem obvious that, because the study of development is the study of specific instances of change (theory-defined instances; Lerner, 2018), any measure of a person used in a developmental study of program evaluation should be capable of

Longitudinal Research of PYD: Issues and Options

detecting change if it, in fact, occurs. However, the use of change-sensitive measures has not always occurred in developmental research and program evaluation. Many researchers opt to use measures that have been designed to assess traits. A trait is a theoretical concept that is intended *not to change* across time or place (e.g., McCrae et al., 2000). If a measure of a trait has been developed appropriately, then, *by definition*, there will not be changes in scores across developmental periods or across the contexts of development. As such, to enact a study or program evaluation wherein the goal is to identify change if it has occurred, use of measures of traits makes no sense (Lerner, 2018; Lerner & Callina, 2014). Simply, then, in longitudinal research or program evaluation, change-sensitive measures must be used. Several publications explain how such sensitivity can be established (e.g., Campbell & Fiske, 1959; Card, 2017; Clifton, 2019; Rioux & Little, 2020).

In addition, such measures should be psychometrically sound. This soundness involves three psychometric properties: reliability, validity, and measurement invariance (also termed measurement equivalence; Card, 2017). Reliability is established when the scores derived from the use of a measure are consistent. There are many ways of assessing consistency. Consistency can be assessed across time, for instance, through repeated administrations of a measure (i.e., through test-retest consistency), by assessing if one form of a measure yields scores that are consistent with those yielded by another form of the measure (i.e., through alternate-form reliability), or by assessing if one portion of a measure yields scores consistent with another portion of the measure (i.e., through computing internal-consistency reliability, by computing split-half reliability or Cronbach's alpha coefficient).

Validity is established when scores on a measure are related to a measure of a phenomenon theoretically quite different than the measure itself. In other words, if reliability is the convergence of scores from measures designed to be maximally similar, validity is the convergence of scores from measures that are designed to be maximally different (Baltes et al., 1977; Campbell & Fiske, 1959). For instance, the measurement of growth mindset (Dweck, 2016) and the measurement of STEM skills are assessments of phenomena that are vastly different. However, if there is a theory that says that the presence of growth mindset at the beginning of an academic years should predict increases in STEM skills across the year, then the identification of such a relation would indicate the validity of a measure of growth mindset. There are also several ways of assessing validity (e.g., content-, construct-, convergent- and divergent-, factorial-, concurrent-, predictive-, and expert-rater validity); as with reliability,



Longitudinal Research of PYD: Issues and Options

there are many publications that explain these methods (e.g., Baltes et al., 1977; Campbell & Fiske, 1959; Card, 2017).

Although many researchers and program evaluators consider only reliability and validity in regard to psychometric quality, in the study of development it is essential to address the issue of measurement invariance (equivalence). This issue addresses the question of whether a measure has comparable meaning across time, place, and people (Card, 2017). Development must be studied across time points, for example, an  $x$ -axis divided by age of youth. Does a survey or behavioral measure used to index a construct (say, PYD, intentional self-regulation, spirituality, empathy, intelligence, resilience, etc.) at age 10 years mean the same thing when the measure is used to index this same construct at ages 15 years, 20 years, 25 years, etc.? Remembering that the study of development is the study of qualitative transformations, omission of information about measurement invariance across  $x$ -axis points means that the interpretation of one's findings about developmental change or about the outcomes of a program are indeterminate.

However, age invariance is not the only instance of measurement across which invariance must be established. Consider again the measurement of PYD. Does a measure of youth of the same age but of different gender, racial, ethnic, cultural, or national groups mean the same thing (e.g., Spencer & Spencer, 2014)? Is there invariance across these attributes of people and place or, in addition, rural versus urban locations, socioeconomic status, or faith groups? Moreover, is there invariance across historical time? Does the meaning of intelligence, social relationships, or civic engagement change across history? Is the nature of civic contributions, in a world where youth are increasingly connected to youth around the world through social media, so different than in prior generations that knowledge of youth activism in the 1960s and 1970s is qualitatively different from such active and engaged citizenship in the third decade of the 21<sup>st</sup> century?

Clearly, the need to establish measurement invariance across diverse youth, contexts, and times of life and history constitutes fundamental challenges for research and applications aimed at describing, explaining, or optimizing youth development around the world. As such, although the establishment of measurement invariance has not characterized the study of PYD, either in the United States or internationally (Lerner et al., 2019), we believe that such work is needed to understand the positive development of all youth, whether the focus of research is on young people within a specific nation or on cross-national or cross-cultural comparisons of youth.

There are different levels or degrees of invariance that can be established (e.g., Card, 2017) and there are different ways of describing these differences. The simplest instance of invariance is configural. Here the pattern of individual variables that collectively measure a facet of development (e.g., character virtues) appears to be the same across age, race, gender, etc. A second level of invariance is termed metric (or weak) invariance. Here the regression slopes of the components (i.e., the factor loadings) of the measure are demonstrated to be the same (the variation of corresponding factors are equal) across measures. Strong invariance (which may also be labeled scalar or intercept invariance) exists when the factor loadings and factor means are invariant across groups. Technically, strong invariance exists when the two lines that cross the y-axis are parallel; that is, the lines that are graphed to represent the scores of two groups are parallel. Finally, strict invariance exists when the degree to which measures have information that is unexplained by the measure (termed residual variance) are the same.

The computation of configural, weak, strong, and strict invariance requires fairly complex statistical procedures. However, these procedures are clearly presented in several publications and such computations are a useful arena for collaboration between researchers and practitioners (Card, 2017). This collaboration can be fruitful for both groups because developing a psychometrically sound measure is not a one-time project, and it must be enacted for each different application of that measure. Indeed, Nunnally (1973, 1978) has likened measure development to a life-span commitment.

In the CI Study of PYD, practitioners from Compassion International and researchers from Tufts University, Boston College, and Fuller Theological Seminary have collaborated to enact the very measure-development work described here (see Tirrell et al., 2020; Tirrell, Geldhof, et al., 2019). Tirrell, Geldhof, et al. (2019) established the reliability, validity, and invariance for measures to-be-used in a novel context—in this case, measures developed using samples from the United States and Mexico were then to be used in El Salvador. If the researchers had not established invariance, then the comparisons made across the two countries would not have been valid (see King et al., 2019).

Furthermore, when the collaborators of the CI Study wanted to assess context-level variables pertinent to the CI-supported youth programs, they created a measure for the “Big Three” features of effective youth programs (that is, positive and sustained adult-youth relationships; life-skill-building activities; and opportunities for youth contribution and leadership of valued

Longitudinal Research of PYD: Issues and Options

activities; see Tirrell et al., 2020) using data collected in Rwanda. In this example, the CI Study research team first developed items informed by theory to assess the Big Three, and then, when data were collected using those newly developed items, they used an iterative process of (a) refining the measure for *parsimony*, that is, efficiency in data collection through finding the smallest number of items that were able to measure the construct without compromising the levels of reliability, validity and invariance); and (b) *robustness*, that is ascertaining that reliability, validity, and invariance remained the same within subgroups of the sample (e.g., boys versus girls, or rural versus urban youth).

It is important to note, then, that *psychometric properties do not exist within a measure itself*—as if the words on a page possess reliability, validity, or invariance. Rather, *psychometric properties exist for a sample of individuals*, such that a measure found to be reliable, valid, and invariant in one context (say, with 18- to 25-year-old university students in the United States) may not demonstrate the same properties in another sample (say, with 9- to 16-year-old youth in El Salvador). As such, a “previously validated measure” must still be “validated” each time it is used in a new context and/or with a new sample to draw meaningful conclusions from data collected across diverse individuals, groups, and contexts. In other words, a researcher or program evaluator does not actually know how well a measure will work until data are collected. As such, this truism underscores the importance of using theory to guide the design of measures. This use of theory gives researchers the strongest chance of developing measures that will generate meaningful and useful data.

Developing and then refining a measure useful for indexing developmental change in a construct (e.g., PYD, intentional self-regulation, civic engagement) is, then, an iterative process; and researchers and practitioners, as well as youth, can collaborate in this work (e.g., see Tirrell et al., 2020, regarding the “Big Three” measure developed as part of the CI Study of PYD). Cognitive testing of the substance of measures, involving interrogation of the meaning of measure content to specific youth in specific community and cultural settings, is one instance of such collaboration. For example, in developing a survey measure, cognitive interviewing about both the stem and response scale of items would be useful. Iteration would then involve a mixed-method procedure wherein pilot testing of quantitative results might lead to further refinement through additional cognitive interviewing and, subsequently, to another quantitative test, for instance, assessing the measure for reliability, validity, and dimensions of invariance. As Nunnally (1973, 1978) intimated, this qualitative↔quantitative process might continue through several revisions of a measure.

Although the step of cognitive interviewing is costly both regarding time and other resources, it yields important information about how a sample is interpreting and responding to a measure. In the CI Study of PYD, although budget constraints prevented full cognitive interviews, measures were adapted for the context using a translation-back-translation procedure (in which the measures were translated to the local language and, then, the translated versions were independently translated back to English to check for accuracy) followed by further adaptation based on engagement with key local stakeholders during data collector training and pilot testing (in this case, program practitioners, data collectors, and the research team local to the context; see Tirrell, Geldhof, et al., 2019).

Across all the measure development steps we have described, there is another important point to keep in mind. It is important that the response options used with measures—whether surveys or behavioral options are used to assess development—maximize the variation in responding available to youth. Variability in responses to measures is needed to establish whether there are meaningful connections (covariation) among these responses (whether testing for reliability, validity, or invariance). When variability is constrained, the researcher or program evaluator is limiting the ability of data analysis to demonstrate psychometric quality.

For instance, in surveys, response options often involve Likert scales, where options may range from “1” to “5” or “7” (e.g., from strongly disagree to strongly agree). These scales substitute ordinal measurement for what should be true interval measurement and, mistakenly and statistically problematically, treat ordinal responses as if interval measurement has occurred (Rioux & Little, 2020). Historically, the use of ordinal, Likert-type response options was employed to make more feasible collection of survey data from large numbers of research participants. However, at this writing, the use of tablets, laptops, and other electronic devices (including smart phones) enable a return to true interval assessment (e.g., “0%” = disagree completely” to “100%” = agree completely).

In the CI Study of PYD, a pilot test was done in El Salvador using 5-point Likert scales (Tirrell, Geldhof, et al., 2019). However, as ceiling effects were noted in the pilot data (thus minimizing variance, as youth only used a relatively small portion of the response scale), and as the research team sought to maximize the opportunities for variance both within and between individuals, a 100-point response scale was then used when fully implementing the surveys in El Salvador and Rwanda (see Tirrell et al., 2020). This approach enhanced the quality of the data

### Longitudinal Research of PYD: Issues and Options

collected and analyzed, for example, in regard to response-option precision, not restricting variance, and correct use of statistical methods; that is, problems of interpretation of findings exist when statistical methods appropriate for use with interval data are used with ordinal data.

In short, developing theoretically-predicated measures that are change sensitive and psychometrically sound requires considerable effort if developmental change is to be adequately studied or promoted in youth programs. However, without such effort, neither change-sensitive developmental designs nor change-sensitive developmental analyses will be useful in the illumination of PYD processes. On the other hand, if such measures are used, then researchers and program evaluators can attend to issues pertinent to maximizing the chances that the design and analysis of developmental data advance understanding of how PYD may be promoted.

### **Designing Longitudinal PYD Research and Program Evaluation**

There are many types of longitudinal designs (e.g., Baltes et al., 1977; Collins, 2006). For instance, some longitudinal designs focus on groups, or “cohorts” (e.g., all individuals born in the United States in the 1940s), and address questions about whether scores for the group (e.g., probability of voting in national elections) change across the group when they are 20 years of age (in the 1960s), when they are in their 30s (in the 1970s), when they are in their 40s (in the 1980s), etc. Such research can also ask questions about whether one cohort (all individuals born in the United States in the 1940s) differs from another cohort (e.g., all individuals born in the United States in the 1970s). For example, the groups can be compared to their probability of voting in national elections when they were in their 20s (in the 1960s and 1990s, respectively), in their 30s (in the 1970s and the 2000s, respectively), in their 40s (in the 1980s and the 2010s, respectively), etc. Note, however, that these repeated measurements of cohorts can illuminate group constancies or changes in scores, but they cannot say anything about whether any specific individual in a group changed or did not change in their voting behavior (or in any other facet of their behavior).

In the CI Study of PYD in El Salvador and Rwanda, for instance, two major groups of interest are assessed: Those youth enrolled in CI-supported programs, and comparison youth not enrolled but living in similar conditions of poverty. Repeated measures of these groups can demonstrate, on average, whether differences exist between CI-supported and non-CI-

Longitudinal Research of PYD: Issues and Options

supported youth. However, such a data set cannot assess whether a specific individual is changing because of CI enrollment.

Only longitudinal studies of individuals can discern whether within-person change occurred. There are also many different types of longitudinal designs that involve the repeated measurement of individuals, and there is a rich literature describing these designs (e.g., Baltes et al., 1977; Collins, 2006; Hamaker et al., 2018; Ram & Grimm, 2015). We have already noted some important features of such within-person longitudinal research that are important to attend to, no matter the specific design used to study the PYD process; that is, all designs should focus on people and not variables, and  $x$ -axis points should be selected to illuminate the transformation in the changes in attributes that are theoretically believed to represent the PYD process. However, some additional issues must be addressed in longitudinal designs intended to document within-person change.

We have noted that not all changes are developmental ones. Importantly, change may be an artifact of statistical phenomena not related to an individual's development. If an initial observation of a person's attribute (e.g., the person's score on a measure of PYD; Geldhof, Bowers, Boyd, et al., 2014; Geldhof, Bowers, Mueller, et al., 2014) is "by chance" discrepant from the true mean of the distribution of scores for the attribute in a population, then a second observation of the person's score will likely be closer to the true mean. This phenomenon is termed *regression to the mean*. If there are only two times of measurement involved in a repeated measurement design, then researchers or program evaluators will be challenged in attempting to accurately discriminate within their sample of youth between actual developmental change versus regression to the mean.

Observing and identifying or creating the basis of true developmental change is obviously the goal of the researcher or program evaluator, respectively. Therefore, distinguishing such change from a statistical artifact of repeated measurement is vital. Nesselroade et al. (1980) addressed this problem, and they discovered that if a third time of testing was included in a longitudinal design, effects of regression to the mean diminished substantially. In addition, the greater the number of occasions of repeated measurement, the less substantial were any changes due to regression to the mean effects. Accordingly, *at least three times of testing* is the minimum number of  $x$ -axis points used in a longitudinal study of within-person change. If resources of time and money exist, more than three times of testing should be used.

Longitudinal Research of PYD: Issues and Options

This point raises another key issue in the design of longitudinal research aimed at studying within-person change: if at least three times of testing are a minimum number of observation points to include in a design, how many times of testing—how many *x*-axis points—should be used to determine the precise nature of a specific person’s changes across periods of development? How many times of testing are needed to assess a specific young person’s specific PYD pathway across adolescence? How many times of testing are needed to ascertain if the PYD pathway of one youth is the same as the pathway of another youth?

Answering these questions is complicated by the fact that, although each young person has unique attributes, each young person also has attributes shared with some but not all people and, as well, attributes possessed by all people (Kluckhohn & Murray, 1948). For instance, each youth possesses a complement of DNA that is specific (i.e., the complement of DNA received at conception and located in the cell nucleus and in other cell structures, such as the mitochondria); as such, each youth, including monozygotic (identical) twins, is in this sense a specific individual, and this individuality is enhanced as the person coacts with the specific contexts encountered across life and a specific *epigenetic* history is therefore generated (Lerner, 2018; Moore, 2015; Slavich & Cole, 2013). However, this same biologically unique individual also has attributes that are shared with only some other people; examples here are the reproductive system of males being different from the reproductive system of females, or the culture of youth in LMICs being different than the cultures of youth developing within economically prosperous nations. Moreover, each individual, whether male or female, whether from LMICs or not, has attributes that are possessed by all humans. Examples are the common respiratory, circulatory, digestive, and nervous systems of people and, as well, the presence in all individuals of developmental processes that create regularities in sensorimotor, language, emotional, and cognitive development.

Therefore, to address questions about the number of *x*-axis points needed to understand the course of PYD of a specific young person, decisions must be made by researchers and program evaluators about whether the youth-specific (idiographic) aspects of development are of concern and/or whether there is interest in understanding the group (differential) or general (nomothetic) aspects of the young person’s development.

Depending on whether idiographic, differential, or nomothetic facets of the young person are of interest, decisions about research design must be integrated with a focus on data analysis. We discuss such integration next.

## Analysis of PYD Data

We have noted that there are many publications about developmental measurement and research design. There are at least comparable numbers of publications pertinent to methods of analyzing developmental data (e.g., Bolger & Laurenceau, 2013; Hamaker et al., 2018; Little et al., 2009; Molenaar et al., 2014; Ram & Grimm, 2015; von Eye et al., 2015). In the present article, we will certainly not attempt to provide a thorough or even introductory discussion of this literature. Instead, we will point to some of the data analysis issues the PYD researcher or program evaluator should consider in selecting an approach to data analysis.

As we have emphasized in this article, the first step to be taken is to return to the theoretical basis of one's work. Depending on the theory-based question being asked, statistical procedures that focus on nomothetic, differential, or idiographic data analyses may be used. For instance, researchers or program evaluators may want to ascertain if there is evidence that one or more youth show developmental changes that are consistent with general (nomothetic) sequences of cognitive, moral reasoning, identity, or academic development. Analyses that involve at least three data points may be sufficient for such analyses, which would ascertain the extent to which the youth in a sample varied from normative, population data or, at least, from the data descriptive of the sample under study. Such comparisons are common in educational assessments (Cantor et al., 2021). Such assessments are also used in the CI Study of PYD in El Salvador and Rwanda, in which groups of CI-supported and non-CI-supported youth are compared to ascertain whether developmental trajectories differ based on program enrollment.

Analogously, researchers or program evaluators may want to ascertain if there is evidence that specific youth have differential developmental pathways (or trajectories) of PYD; if so, questions may arise about whether there are youth whose pathways are consistent with some youth in a sample but different from other youth in a sample. Such interests may be addressed by several types of analyses that fall under the label of person-centered analyses (e.g., latent trajectory analyses, growth mixture modeling). Once again, at least three data points may be sufficient for such analyses.

However, what would be the needed methodological steps to take if researchers or program evaluators were interested in ascertaining if each young person in their study or program showed evidence of meaningful individual facets of PYD? Such interests pertain to person-



Longitudinal Research of PYD: Issues and Options

specific designs and data analyses. These analyses fall within the categories of intensive longitudinal research designs (Bolger & Laurenceau, 2013) and involve variants of what are termed time-series analyses.

To analyze a quantitative data set correctly, statisticians advise researchers to ascertain if the data set has sufficient *power* to use a specific data analytic method. Power is simply a computation undertaken to ascertain if the data set can be used to answer questions with specific statistical methods. One way to increase the power of a data set is to increase the number of observations included in it. Typically, this step involves increasing the number of participants that are observed. If complex statistics (e.g., latent trajectory analyses) are needed to address a specific question (e.g., are there groups of youth with common PYD trajectories?) and power is insufficient, an appropriate step to take would be to increase the number of youth being studied.

However, what if only one young person is the subject of one's interest, either across a portion of adolescence or across the time in which the young person participated in a PYD-promotion program? What if one wanted to see if this one young person had a meaningful PYD trajectory? Where would power come from? The answer here is to increase the number of *observations* of the youth (as compared to increasing the number of *youth* being observed, as in a variable-centered design). Increase the number of *x-axis* points so that statistical procedures could be used to ascertain if the young person showed a meaningful PYD trajectory.

Moreover, if a series of observations for one youth could have a sufficient number of *x-axis* points (or, in other words, if a time series had enough points to provide power for one youth), then comparable time series could be enacted for other youth. If so, then we may return to the statistical tool noted earlier, that is, the idiographic filter developed by Molenaar & Nesselroade (2012, 2014, 2015). The use of this statistical tool in sufficiently powerful time series with several youth enables, on the one hand, youth-specific meaningful trajectories to be identified if, in fact, they exist and, on the other hand, analyses to be undertaken to see if there are also differential or even nomothetic facets of these youth-specific pathways.

Through this procedure, Molenaar and Nesselroade (2012, 2014, 2015) have given PYD researchers and program evaluators a powerful tool for bringing PYD research to the point wherein all facets of a young person's development—the nomothetic, the differential, and the idiographic—can be studied simultaneously and integratively within one data set. The specificity

### Longitudinal Research of PYD: Issues and Options

of each young person's development (Bornstein, 2017, 2019) can now not be lost if there is also interest in the differential or nomothetic facets of development that are also involved in the development of all youth (Kluckhohn & Murray, 1948). Indeed, the generative influence of the integrative approach to developmental data analysis stimulated by Molenaar and Nesselroade has led to new research designs in developmental science. One example involves a "burst" of person-specific, intensive *x*-axis assessments of a subsample of youth who are part of a longitudinal study involving a large sample of youth tested at three or so times of testing.

Thanks to funding from the Templeton World Charity Foundation, such an integrative design is to-be-implemented in Uganda as part of the CI Study of PYD, involving both variable-centered assessments across three times of testing, and a person-specific "burst" design with a subsample of the youth being assessed. Indeed, such an innovative design enables researchers and practitioners to assess both variable-centered questions (e.g., "is enrollment in CI-supported programs associated with increased thriving and PYD compared to youth not enrolled?") as well as person-specific questions (e.g., "is one specific youth developing life skills needed to thrive and is that development linked to her enrollment in CI-supported programs?"). Building such a data set capable of assessing both group-based differences (i.e., a variable-centered design) as well as individual trajectories (i.e., a person-specific design) gives researchers and practitioners powerful tools for conducting program evaluations aimed at describing, explaining, and, in turn, optimizing PYD among youth living in LMICs.

## Conclusions

As exemplified by Banati (2021) and Petersen et al. (2017), and by the special section of *Child Development* edited by Leman et al. (2017), interest in longitudinal studies of PYD in settings outside the United States has burgeoned across the second decade of the 21<sup>st</sup> century. As we have explained in this article, the developmental portion of the concept of PYD places a set of theory-based methodological requirements pertinent to measurement, research and program design, and data analysis. All researchers and program evaluators must enact the appropriate steps to address these requirements in their respective attempts to describe, explain, and optimize the course of positive development among diverse youth around the world.

Arguably, these requirements are needed most in the work of researchers and practitioners aimed at understanding and promoting PYD among youth in LMICs, young people whose life challenges are shaped by multiple adverse situations associated with racism, poverty, gender

### Longitudinal Research of PYD: Issues and Options

inequalities, political inequities, and the absence of adequate health and medical resources. The need for actionable research and cutting-edge evaluation is greatest for youth whose adversities are most pressing and whose opportunities for thriving are most constrained.

The examples that we have provided from our work on the CI Study of PYD serve only as one example of the important ongoing efforts of governmental and non-governmental organizations around the globe to enact applied developmental science projects that can contribute to a knowledge base that will positively impact the lives of the millions of youth developing in LMICs and, as well, youth living in more prosperous nations. We have suggested the importance of practitioner-researcher collaborations in these instances of the application of applied developmental science. We underscore here both the importance of such collaborations and, as well, the need for researchers and program evaluators to remain humble in such collaborative activities.

The practitioners and the youth, families, and communities they are serving are the actual experts in development in the LMIC contexts within which they live. Simply, productive collaborations can exist when there is respect for and humility about this knowledge by members of the researchers engaged in collaboration. Our hope is that the methodological ideas that we have presented in this article will help the researchers and practitioners serving all these youth to engage in better and more effective collaborative efforts to promote PYD for all youth in all nations around the world.

### Acknowledgements

Preparation of this article was supported in part by grants from the National 4-H Council, Compassion International, the Templeton Religion Trust, the Templeton World Charity Foundation, and the Chan Zuckerberg Initiative.

### References

- Arnold, M. E., & Gagnon, R. J. (2020). Positive youth development theory in practice: An update on the 4-H Thriving Model. *Journal of Youth Development, 15*(5), <https://doi.org/10.5195/jyd.2020.954>
- Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods*. Monterey, CA: Brooks/Cole.

Longitudinal Research of PYD: Issues and Options

- Banati, P. (Ed.). (2021). *Sustainable development across the life course: evidence from longitudinal research*. Bristol, UK: Policy Press/Bristol University Press.
- Benson, P. L. (2008). *Sparks: How parents can help ignite the hidden strengths of teenagers*. San Francisco, CA: Jossey-Bass.
- Bolger, N., & Laurenceau, J-P. (2013). *Intensive longitudinal methods: An introduction to diary and experiences sampling research*. New York, NY: The Guildford Press.
- Bornstein, M. H. (2017). The specificity principle in acculturation science. *Perspectives in Psychological Science*, 12(1), 3-45. <https://doi.org/10.1177/1745691616655997>
- Bornstein, M. H. (2019). Fostering optimal development and averting detrimental development: Prescriptions, proscriptions, and specificity. *Applied Developmental Science*, 23(4), 340-345. <https://doi.org/10.1080/10888691.2017.1421424>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105. <https://doi.org/10.1037/h0046016>
- Cantor, P., Lerner, R. M., Pittman, K., Chase, P. A., & Gomperts, M. (2021). *Whole-child development, learning, and thriving: A dynamic systems approach*. New York, NY: Cambridge University Press.
- Card, N. A. (2017). Methodological issues in measuring the development of character. *Journal of Character Education*, 13(2), 29-45.
- Catalano, R. P., Hawkins, J. D., Berglund, M. L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health*, 31, 230-239. [https://doi.org/10.1016/s1054-139x\(02\)00496-2](https://doi.org/10.1016/s1054-139x(02)00496-2)
- Clifton, J. D. W. (2019). Managing validity versus reliability trade-offs in scale-building decisions. *Psychological Methods*. <http://dx.doi.org/10.1037/met0000236>
- Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, 57, 505-528. <https://doi.org/10.1146/annurev.psych.57.102904.190146>
- Damon, W. (2004). What is positive youth development? *Annals of the American Academy of Political and Social Science*, 591, 13-24. <https://doi.org/10.1177/0002716203260092>
- Damon, W. (2008). *The path to purpose: Helping our children find their calling in life*. New York, NY: Simon and Schuster.
- Dweck, C. (2016). *Mindset: The new psychology of success*. New York, NY: Ballantine.
- Eccles, J. S., & Gootman, J. A. (Eds.). (2002). *Community programs to promote youth development/committee on community-level programs for youth*. Washington, D. C.: National Academy Press.
- Geldhof, G. J., Bowers, E. P., Boyd, M. J., Mueller, M. K., Napolitano, C. M., Schmid, K. L., Lerner, J. V., & Lerner, R. M. (2014). Creation of short and very short measures of the Five Cs of Positive Youth

Longitudinal Research of PYD: Issues and Options

- Development. *Journal of Research on Adolescence*, 24(1), 163-176.  
<https://doi.org/10.1111/jora.12039>
- Geldhof, G. J., Bowers, E. P., Mueller, M. K., Napolitano, C. M., Schmid, K. L., & Lerner, R. M. (2014). Longitudinal analysis of a very short measure of positive youth development. *Journal of Youth and Adolescence*, 43(6), 933-949. <https://doi.org/10.1007/s10964-014-0093-z>
- Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, 53, 820-841.  
<https://doi.org/10.1080/00273171.2018.1446819>
- King, P. E., Yoo, Y., Vaughn, J. M., Tirrell, J. M., Geldhof, G. J., Iraheta, G., Williams, K., Sim, A., Stephenson, P., Dowling, E., Lerner, R. M., & Lerner, J. V. (2019). Evaluating the measure of diverse adolescent spirituality in samples of Mexican and Salvadoran youth. *Psychology of Religion and Spirituality*. <https://doi.org/10.1037/rel0000279>
- Kluckhohn, C., & Murray, H. (1948). Personality formation: The determinants. In C. Kluckhohn & H. Murray (Eds.), *Personality in nature, society, and culture* (pp. 3-32). New York, NY: Knopf.
- Kohlberg, L. (1978). Revisions in the theory and practice of moral development. *New directions for child development*, 2, (93-120). <https://doi.org/10.1002/cd.23219780207>
- Lansford, J. E., & Banati, P. (Eds.). (2018). *Handbook of adolescent development research and its impact on global policy*. New York, NY: Oxford University Press.
- Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist*, 55(1), 170-183. <https://doi.org/10.1037//0003-066x.55.1.170>
- Leman, P. J., Smith, E. P., & Petersen, A. C. (2017). Introduction to the special section of child development on positive youth development in diverse and global contexts. *Child Development*, 88(4), 1039-1044. <https://doi.org/10.1111/cdev.12860>
- Lerner, R. M. (2018). *Concepts and theories of human development* (4<sup>th</sup> ed.). New York, NY: Routledge.
- Lerner, R. M., & Bornstein, M. H. (2021). Contributions of the specificity principle to theory, research, and application in the study of human development: A view of the issues. *Journal of Applied Developmental Psychology*, 75(July-August 2021), 101294.  
<https://doi.org/10.1016/j.appdev.2021.101294>
- Lerner, R. M., & Callina, K. S. (2014). The study of character development: Towards tests of a relational developmental systems model. *Human Development*, 57(6), 322-346.  
<https://doi.org/10.1159/000368784>
- Lerner, R. M., Lerner, J. V., Bowers, E., & Geldhof, G. J. (2015). Positive youth development and relational developmental systems. In W. F. Overton & P. C. Molenaar (Eds.), *Handbook of child*

Longitudinal Research of PYD: Issues and Options

- psychology and developmental science: Theory and method* (7<sup>th</sup> ed., Vol. 1, pp. 607-651). Editor-in-chief: R. M. Lerner. Hoboken, NJ: Wiley.
- Lerner, R. M., Lerner, J. V., Murry, V. M., Smith, E. P., Bowers, E. P., Geldhof, G. J., & Buckingham, M. H. (2021). Positive youth development in 2020: Theory, research, programs, and the promotion of social justice. *Journal of Research on Adolescence*.
- Lerner, R. M., Tirrell, J. M., Dowling, E. M., Geldhof, J., Gestsdóttir, S., Lerner, J. V., King, P. E., Williams, K., & Sim, A. T. R. (2019). The end of the beginning: Evidence and absences studying PYD in a global context. *Adolescent Research Review*, 4(1), 1-14. <https://doi.org/10.1007/s40894-018-0093-4>
- Little, T. D., Card, N. A., Preacher, K. J., & McConnell, E. (2009). Modeling longitudinal data from research on adolescence. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology: Individual bases of adolescent development* (3<sup>rd</sup> ed., Vol. 1, pp. 15-54). Hoboken, NJ: Wiley.
- McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Hrebícková, M., Avia, M. D., Sanz, J., Sánchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R., Saunders, P. R., & Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, 78, 173-186. <https://doi.org/10.1037/0022-3514.78.1.173>
- Molenaar, P. C. M., Lerner, R. M., & Newell, K. (Eds.). (2014). *Handbook of developmental systems theory and methodology*. New York, NY: Guilford.
- Molenaar, P. C. M., & Nesselroade, J. R. (2012). Merging the idiographic filter with dynamic factor analysis to model process. *Applied Developmental Science*, 16, 210-219. <https://doi.org/10.1080/10888691.2012.722884>
- Molenaar, P. C. M., & Nesselroade, J. R. (2014). New trends in the inductive use of relation developmental systems theory: Ergodicity, nonstationarity, and heterogeneity. In P. C. Molenaar, R. M. Lerner, and K. M. Newell (Eds.), *Handbook of developmental systems and methodology*. (pp. 442-462). New York, NY: Guilford.
- Molenaar, P. C. M., & Nesselroade, J. R. (2015). Systems methods for developmental research. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of child psychology and developmental science: Theory and method* (7<sup>th</sup> ed., Vol. 1, pp. 652-682). Editor-in-chief: R. M. Lerner. Hoboken, NJ: Wiley.
- Moore, D. S. (2015). *The developing genome: An introduction to behavioral epigenetics*. New York, NY: Oxford University Press.
- Nesselroade, J. R., & Molenaar, P. C. M. (2010). Emphasizing intraindividual variability in the study of development over the life span. In W. F. Overton (Ed.), *Handbook of life-span development*.

Longitudinal Research of PYD: Issues and Options

- Cognition, biology, methods* (Vol. 1, pp. 30-54). Editor-in-chief: R. M. Lerner. Hoboken, NJ: Wiley.
- Nesselroade, J. R., Stigler, S. M., & Baltes, P. B. (1980). Regression toward the mean and the study of change. *Psychological Bulletin*, *88*(3), 622-637. <https://doi.org/10.1037/0033-2909.88.3.622>
- Nunnally, J. C. (1973). Research strategies and measurement methods for investigating human development. In J. R. Nesselroade & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological issues* (pp. 87-109). New York, NY: Academic Press.
- Nunnally, J. C. (1978). *Psychometric theory*, 2<sup>nd</sup> ed. New York, NY: McGraw-Hill.
- Overton, W. F. (2015). Process and relational developmental systems. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of child psychology and developmental science: Theory and Method* (7<sup>th</sup> ed., Vol. 1, pp. 9-62). Editor-in-chief: R. M. Lerner. Hoboken, NJ: Wiley.
- Petersen, A. C., Koller, S. H., Motti-Stefanidi, F., & Verma, S. (Eds.). (2017). *Positive youth development in global contexts and social and economic change*. New York, NY: Routledge.
- Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (3<sup>rd</sup> ed., Vol. 1, pp. 703-723). New York, NY: Wiley.
- Ram, N., & Grimm, K. J. (2015). Growth curve modeling and longitudinal factor analysis. In W. F. Overton, P. C. M. Molenaar, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Theory and method* (7<sup>th</sup> ed., Vol. 1, pp. 758-788). Hoboken, NJ: Wiley.
- Rioux, C., & Little, T. D. (2020). Underused methods in developmental science to inform policy and practice. *Child Development Perspectives*, *14*(2), 97-103. <https://doi.org/10.1111/cdep.12364>
- Rose, T. (2016). *The end of average: How we succeed in a world that values sameness*. New York, NY: HarperCollins.
- 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. [https://doi.org/10.1207/S1532480XADS0702\\_6](https://doi.org/10.1207/S1532480XADS0702_6)
- Slavich, G. M., & Cole, S. W. (2013). The emerging field of human social genomics. *Clinical Psychological Science*, *1*, 331-348. <https://doi.org/10.1177/2167702613478594>
- Spencer, M. B., & Spencer, T. R. (2014). Exploring the promises, intricacies, and challenges to positive youth development. *Journal of Youth and Adolescence*, *43*, 1027-1035. <https://doi.org/10.1007/s10964-014-0125-8>
- Tirrell, J. M., Dowling, E. M., Gansert, P., Buckingham, M., Wong, C. A., Suzuki, S., Naliaka, C., Kibbedi, P., Namurinda, E., Williams, K., Geldhof, G. J., Lerner, J. V., King, P. E., Sim, A. T. R., & Lerner, R. M. (2020). Toward a measure for assessing features of effective youth development programs: Contextual safety and the "big three" components of positive youth development

Longitudinal Research of PYD: Issues and Options

programs in Rwanda. *Child & Youth Care Forum*, 49, 201-220. <https://doi.org/10.1007/s10566-019-09524-6>

Tirrell, J. M., Gansert, P. K., Dowling, E. M., Geldhof, G. J., Lerner, J. V., King, P. E., Iraheta, G., Williams, K., Sim, A.T.R., & Lerner, R. M. (2019). Illuminating the use of the specificity principle to go inside the black box of programs: The sample case of an El Salvador PYD program. In S. Verma, A. Petersen, & J. Lansford (Eds.), *Sustainable human development: Challenges and solutions for implementing the United Nations' Goals* [Special issue]: *Zeitschrift für Psychologie*, 227(2), 121-128. <https://doi.org/10.1027/2151-2604/a000363>

Tirrell, J. M., Gansert, P. K., Dowling, E. M., Williams, K., Iraheta, Lerner, J. V., King, P. E., Sim, A. T. R., & Lerner, R. M. (2021). Interrogating ergodicity and specificity in youth development programs in El Salvador. *Journal of Applied Developmental Psychology*. <https://doi.org/10.1016/j.appdev.2021.101243>

Tirrell, J. M., Geldhof, G. J., King, P. E., Dowling, E., Sim, A., Williams, K., Iraheta, G., Lerner, J. V., & Lerner, R. M. (2019). Measuring spirituality, hope, and thriving among Salvadoran youth: Initial findings from the Compassion International study of positive youth development. *Child & Youth Care Forum*, 48, 241-268. <https://doi.org/10.1007/s10566-018-9454-1>

von Eye, A., Bergman, L. R., & Hsieh, C. (2015). Person-oriented methodological approaches. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of Child Psychology and Developmental Science: Theory and Method*, (7<sup>th</sup> ed., Vol. 1., pp. 789-842). Editor-in-chief: R. M. Lerner. Hoboken, NJ: Wiley.