RESEARCH ARTICLE

Validation of a Chinese Positive Youth Development Scale: Dimensionality and factorial invariance

Xiaoqin Zhu[1], Carman K. M. Chu[1], Xue Wu[2], Daniel T. L. Shek[1]*

1 Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, PR China,
2 Office of Undergraduate Studies, The Hong Kong Polytechnic University, Hong Kong, PR China

* daniel.shek@polyu.edu.hk

Abstract

For a multi-dimensional measure of positive youth development (PYD), its factor structure should be invariant across groups and over time. This study examined the factorial validity of the 44-item short form of the "Chinese Positive Youth Development Scale" (CPYDS-SF) that assesses 15 dimensions of PYD attributes. Using two waves of longitudinal data with a one-year interval in between, this study examined the factor structure of the scale and whether the structure is invariant between gender groups and across time. The data were collected from 3,328 adolescents at Wave 1 and 3,638 adolescents at Wave 2, with a matched sample of 2,905 adolescents (mean age = 12.57 ± 0.72 at Wave 1; 49.54% girls). Confirmatory factor analysis revealed that the 15-factor structure fitted the data well. The findings of invariance tests further supported this structure's invariance across gender and time, indicating a stable factor structure of CPYDS-SF among Chinese adolescents. These findings suggest that CPYDS-SF can be used to examine gender differences and the longitudinal development of PYD qualities among Chinese adolescents.

Introduction

Adolescence is a critical period characterized by higher risks of mental illness and problematic behaviors [1–3]. The traditional pathology-based perspective attempts to "eliminate" adolescent developmental problems by dealing with specific risk factors of the respective problem [4]. This approach has been criticized for merely focusing on single problematic behavior while ignoring the co-occurrence of multiple problematic behaviors, and for the wrong assumption that the absence of problems means good development [1]. With a paradigm shift from focusing on pathologies to strengths, the positive youth development (PYD) approach posits that adolescents have potentials to thrive through building multi-faceted developmental assets (e.g., internal and external PYD qualities), which would reduce the likelihood of engaging in problem behaviors while simultaneously increasing the likelihood of having positive growth and well-being [1, 5, 6]. In line with this theoretical proposition, empirical findings have established the close associations between PYD qualities and better developmental

Competing interests: The authors have declared that no competing interests exist.

outcomes, such as better mental health, higher well-being, and fewer problematic behaviors [e.g., 7–9]. As Humphrey, Kalambouka [9] have argued, there has been a growing interest and investment in promoting PYD qualities, such as social and emotional skills, as a means to promote children's well-being, adjustment, and academic achievement. Indeed, rich evidence has shown the benefits of promoting PYD qualities through effective youth program in enhancing life satisfaction, thriving, learning engagement, academic achievement, and prosocial behavior as well as in reducing substance abuse, juvenile delinquency, aggressive behaviors, and academic failure [6, 10, 11].

Multiple frameworks have been proposed to conceptualize PYD qualities [1]. For example, the "Five Cs" model emphasizes "*Competence*," "*Confidence*," "*Character*," "*Connection*," and "*Caring*" as outcomes of positive individual adjustment to contextual factors, which further lead to the "six C," namely "*Contribution*" to self, others, groups, family, community, and society [5]. While achieving these "six Cs," adolescents are also simultaneously less likely to suffer from dysfunctional development [e.g., 12]. Another framework is "social and emotional learning" (SEL), which concentrates on five dimensions of SEL competencies, including "self-awareness," "social awareness," "self-management," "relationship skills," and "responsible decision making" [13]. Numerous evaluation studies have demonstrated the contribution of cultivating these qualities in promoting youths' healthy functioning and well-being [1, 10].

In addition, based on an inductive framework, Catalano et al. [14] identified 15 PYD constructs from effective youth prevention programs. The fundamental thesis is that enhancement programs cultivating 15 PYD qualities among adolescents are effective in promoting their positive outcomes while reducing negative outcomes. These qualities consist of "cognitive competence," "behavioral competence," "self-determination," "prosocial norms," "prosocial involvement," "belief in the future," "clear and positive identity," "bonding," "social competence," "emotional competence," "moral competence," "resilience," "self-efficacy," "spirituality," and "recognition for positive behavior" (see Table 1 for definitions of these qualities). This PYD framework covers the essential elements intrinsic to positive youth development that includes a broad array of qualities, including psychological, social, emotional, behavioral, moral, and cognitive competences. The benefits of PYD attributes have been supported by fruitful empirical evidence [11, 15]. More importantly, this framework integrates essential PYD concepts highlighted in the "Five Cs" (e.g., different psychosocial competencies including moral competence, self-worth, and prosocial involvement) and SEL (i.e., competencies in cognitive, emotional, social, and behavioral dimensions) models [16].

Researchers have adopted different measures to assess PYD qualities. For example, twelve different assessment tools were identified in Humphrey et al.'s [9] systematic review on measures of SEL skills in children and young people. Multiple scales have also been applied in measuring "Five Cs" in different youth samples [e.g., 17–19]. Obviously, valid and reliable assessment tools on PYD qualities are of paramount importance in the evaluation of PYD programs since it allows a more precise assessment of group differences or overtime changes in PYD qualities [20]. Unfortunately, existing literature regarding the assessment of PYD qualities showed several limitations. First, most of the extant PYD measures are developed in Western contexts based on Western samples, particularly North American young people. We argue that assessment tools with good reliability and validity among Western youth may not necessarily work effectively in non-Western populations, such as Chinese adolescents, due to the possible distinctive interpretations and structures of PYD qualities [21, 22]. Furthermore, the studies conducted in the Hong Kong sample did not report reliability values [23].

Second, although factor analyses have been used to examine the different dimensions underlying measures of PYD attributes, there are several problems. Primarily, some studies only used exploratory factor analyses [e.g., 24, 25]. Moreover, amongst studies using

Table 1. Definitions of the 15 PYD qualities and sample items.

Four higher-order dimensions	The 15 basic PYD qualities	A brief definition of the construct	Sample item in the scale
Cognitive-behavioral competence	1. Cognitive competence	Intellectual skills such as critical thinking and problem-solving.	I try new ways to solve my problems.
	2. Behavioral competence	Behavioral skills such as taking verbal and non-verbal actions.	I can express views that are different from others.
	3. Self-determination	Abilities to take age-appropriate actions with independent thinking	I am able to make wise choices.
Prosocial attribute	4. Prosocial norms	Clear standards for prosocial engagement.	I care about unfortunate people in society.
	5. Prosocial involvement	Engagement in activities to make positive contributions.	I will try my best to contribute to my school or society.
Positive identity	6. Beliefs in the future	Being optimistic about one's future and able to develop future goals and options.	I have the confidence to solve my future problems.
	7. Clear and positive identity	Having positive and healthy self-perceptions.	I am a person with self-confidence.
General PYD quality	8. Bonding	Being able to establish positive relationships with positive peers and healthy adults.	When I need help, I trust my teachers will help me.
	9. Social competence	Interpersonal skills such as communication and identification with groups.	I know how to communicate with others.
	10. Emotional competence	Abilities to recognize and manage emotions.	When I am unhappy, I can appropriately show my emotions.
	11. Moral competence	Abilities to make sound moral judgments and behave accordingly.	I have high moral expectations about my behavior.
	12. Resilience	Abilities to adapt to and deal with adversity in positive and effective ways.	When I face difficulty, I will not give up easily.
	13. Self-efficacy	A belief that one has abilities and can attain goals.	I can finish almost everything that I am determined to do.
	14. Spirituality	Sense of purpose and meaning in life and hope.	My life is exciting.
	15. Recognition for positive behavior	Having contexts that reward, recognize, or reinforce young people's positive behavior.	When I help others, my classmates will recognize my behavior.

confirmatory factor analyses, very few have further evaluated whether a scale's factorial structure is stable across subgroups (e.g., boys vs. girls) through invariance tests. For example, Phelps et al. [19] used a 78-item scale to measure "Five Cs" among participants attending a summer sports camp. The authors tested the factor structure of the scale without examining factorial invariance. Similarly, Lopez et al. [18] developed the "Bridge-Positive Youth Development Scale" based on the "Five Cs" model and examined the scale's structural properties without testing invariance. In developing measurement instruments, it is crucial to take into account respondent subgroups that exhibit distinct features, such as gender groups, as they may conceive a concept (e.g., PYD quality) in different ways [26]. In particular, gender differences in youth development have been widely studied and acknowledged in the field of Psychology and Development Science. Research has shown that males and females may experience distinct developmental trajectories, face different challenges, and exhibit varying patterns of strengths and vulnerabilities during adolescence [27–30]. For example, connection appeared to be of greater importance for girls than boys. Whereas girls' connection was associated with less somatic anxiety, fewer physical and psychological symptoms, and high well-being, boys' connection was only a predictor of high well-being [30]. However, boys who experienced higher ACEs were at a heightened risk of exhibiting elevated externalizing trajectories [28]. Nevertheless, most of the studies failed to establish the measurement equivalence of the measures across genders to rule out the possibility that gender differences are in fact an artifact of measurement differences. Thus, testing measurement invariance across groups (e.g., gender groups) plays an important role in ensuring the same structural interpretation between groups,

such that the scores derived from the assessment tools can be meaningfully compared between groups [31].

Third, while there is an increasing emphasis on longitudinal invariance of valid assessment tools [32, 33], very few studies have examined the longitudinal invariance of PYD measures. For example, the PYD measure developed by Wen et al. [34] based on the "Five Cs" model showed factorial invariance across gender (e.g., boys vs. girls) and geographical location (e.g., rural vs. urban) in Chinese contexts. However, its longitudinal invariance is unknown. Chai et al. [22] further examined the scale's psychometric properties and found that the "Four Cs" structure without the "*caring*" dimension better fit the data among Chinese adolescents. The authors argued that elements underlying "*caring*" can be accounted for by the love under "*character*" or the relationship with others under "*connection*" due to the collectivistic social structure in mainland China. Nevertheless, the authors did not examine longitudinal measurement invariance of the four-factor structure either.

While cross-group invariance enables cross-group comparisons of PYD qualities, longitudinal factorial invariance sets the foundation for over-time comparisons by ensuring that a given scale steadily measures the same construct with an identical structure on different occasions [35, 36]. This proposition has been increasingly acknowledged in recent years. For example, recent youth research has examined factorial invariance over time for different measurements, such as the "Problematic Trait Inventory" [37], the "Short Grit Scale" [38], and the "Center for the Epidemiological Studies Depression Scale" (CES-D) [32, 33]. Specifically, the longitudinal invariance of the CES-D ensures that any changes derived from the scale between different time points are attributable to over-time development rather than structural instability in measurement [33].

Given considerable developmental changes in physiological, intellectual, psychological, and social domains over the adolescent period, adolescents' interpretations and perceptions of their personal attributes, including PYD qualities, may vary over time. Noteworthy, age differences in PYD have also been widely acknowledged in previous studies [39–41]. This requires testing of longitudinal validity of a certain PYD scale to make sure that any longitudinal differences derived from the scale are due to life span development rather than measurement error over time [42, 43]. As emphasized by Putnick and Bornstein [44], analyses of measurement invariance over time become particularly important when researchers aim to evaluate program effectiveness by comparing pretest and posttest scores. The key is that the lack of longitudinal invariance will make the comparisons meaningless. Therefore, it is necessary to examine the longitudinal invariance of assessment tools on PYD attributes.

Within different Chinese contexts, Shek and his collaborators developed the 80-item "Chinese Positive Youth Development Scale" (long form, CPYDS-LF) based on the 15 PYD constructs identified by Catalano et al. [14]. This scale measures 15 PYD qualities (e.g., bonding and spirituality) using 15 subscales. According to Shek and Ma [45], the CPYDS-LF possessed adequate factorial validity and internal consistency and the 15 subscales can be further subsumed into four higher-order dimensions, including "cognitive-behavioral competencies," "prosocial attributes," "positive identity," and "general PYD qualities" (see Table 1). In addition, invariance tests indicated that the four-higher-order-factor structure was invariant across gender groups. Given its good validity and reliability, the scale has been widely applied in assessing the impacts of a multi-year PYD program entitled "P.A.T.H.S. Project" and investigating the relationship between PYD qualities and Chinese adolescents' developmental outcomes [15, 46].

To reduce scale length and minimize respondents' fatigue in completing the questionnaire, a 44-item version of CPYDS (i.e., short form, CPYDS-SF) has also been created by selecting the three items having the highest loadings in each construct (two items in the self-efficacy

subscale as there were in total two items in the original scale). Constructing CPYDS-SF as a short form the original CPYDS can contribute to youth research in several aspects. First, a short form allows for a more efficient and streamlined assessment process. Adolescents are often faced with various demands and time constraints in the classroom context, and a shorter assessment tool can minimize their burden and increase participation rates. This is particularly important in longitudinal studies where repeated measurements are required over time. Additionally, a short form can enhance the scalability and practicality of the PYD assessment. It can be easily administered in various settings, such as schools or community organizations, where resources and time may be limited. The availability of a concise measure facilitates the integration of PYD assessment into existing research or intervention programs, promoting the widespread implementation of PYD principles in practice. Furthermore, a short form can contribute to the dissemination of research findings and facilitate cross-study comparisons. When researchers adopt a standardized short form, it enhances the comparability of results across different studies [17, 23, 47–49]. This comparability strengthens the evidence base and allows for a more comprehensive understanding of PYD qualities among diverse populations and contexts. Lastly, the development of a short form reflects the ongoing refinement and advancement of measurement instruments in the field of PYD. By identifying the key dimensions and items that capture the essence of PYD, researchers can gain insights into the core components that contribute to positive youth outcomes in a more parsimonious manner. This knowledge can inform the design and implementation of targeted interventions and programs aimed at promoting PYD.

While the CPYDS-SF has also been widely adopted in youth research [46, 50, 51], no validation research has been conducted to examine the factorial validity and its invariance in different groups (e.g., gender) and over time. Given the above research gaps, the present study aimed to test the factor structure of the 44-item CPYDS-SF among Chinese adolescents. In addition, we also tested whether the factor structure was invariant across gender and over time. By validating and testing the measurement invariance of CPYDS-SF, we aimed to contribute to the existing literature by constructing a valid and reliable assessment tool for PYD qualities, which can be meaningfully adopted to investigate gender differences and longitudinal changes.

Materials and methods

Participants and procedures

The current research used two waves of data collected in a longitudinal youth project in Hong Kong, which aimed to assess Hong Kong Chinese adolescents' developmental outcomes and related psychosocial correlates. A list of 399 public local secondary schools was formed based on information provided by the Education Bureau. A total of 30 schools in different districts were randomly selected from the school list and they were invited to participate in the longitudinal project. If one school rejected the invitation, an alternate school randomly selected from the same district was further invited. The process stopped until the selected school accepted the invitation. Finally, 28 schools were successfully recruited to participate in the project between 5 October 2009 and 9 February 2010. Grade 7 students in these schools were invited to complete a battery of questionnaires between 17 November 2009 and 8 July 2010 (i.e., Wave 1) and one year later (i.e., Wave 2). Before students' participation, consent of their parents was collected through schools via parental notice.

At the two waves, 3,328 (Mage = 12.59 years, SD = 0.74 at Wave 1; 47.60% girls) and 3,638 students (Mage = 13.64 years, SD = 0.75 at Wave 2; 47.84% girls) completed the same questionnaires in their classrooms, respectively, with the presence of two well-trained research staff. Among these students, 2,905 (Mage = 12.57, SD = 0.72 at Wave 1; 49.54% girls) had data at both waves. All the cases were included in the analysis.

The project was approved by the "Human Subjects Ethics Subcommittee" at the authors' university. Before data collection, the participating schools, adolescent participants, and their parents provided written consent.

Measures

This study used the 44-item CPYDS-SF. The 80-item long version of the scale included 15 subscales assessing the respective 15 PYD constructs proposed by Catalano et al. [14]. All items were rated from 1 ("*strongly disagree*") to 6 ("*strongly agree*"). This long version has been validated among Chinese adolescents in Hong Kong and a 4-higher-order-factor model was supported [45]. The short version was formed based on Shek and Ma's [45] study by selecting three items from each subscale that possessed the highest factor loadings, except for the "self-efficacy" subscale, in which the two items in the long version were included in the short version.

Data analysis

Data analysis was performed using Mplus 8.5 [52]. First, we checked whether the assumption of multivariate normality was fulfilled. Results revealed normal distributions (|skewness| < 2.0 and |kurtosis| < 7) of all item scores. Thus, the maximum likelihood (ML) estimation method can be correctly used in next steps [53]. Furthermore, the "full information maximum likelihood estimation" incorporated in Mplus was adopted to handle missing values in variables by making full use of all available data [54].

Second, confirmatory factor analysis (CFA) was used to test the conceptual structure of the 44-item CPYDS-SF. We compared two competing models at each wave including the 15-primary-factor structure and the 4-higher-order-factor structure identified by Shek and Ma [45]. Adequate absolute model fit was indicated by multiple indices and criteria [55]: "Comparative Fit Index" (CFI ≥ 0.90), "Tucker-Lewis Index" (TLI ≥ 0.90), "Root Mean Square Error of Approximation" (RMSEA ≤ 0.08), and "Standardized Root Mean Square Residual" (SRMR ≤ 0.08). As Chi-square difference tests may produce bias in a large sample [56], this study used differences in "Bayesian information criterion" (BIC) between the two competing models to decide which model was preferred [57]. The applied cut-off is 10 points (i.e., ΔBIC = 10). Specifically, ΔBIC = 10 indicates a 150:1 likelihood ($p < .05$) and ΔBIC greater than 10 represents "*very strong*" evidence favoring the model with a smaller BIC [58]. This method has been applied in previous research on scale validation [33].

Third, to determine whether the CPYDS-SF measures PYD qualities with the same structure between boys and girls (i.e., across gender) and across the two waves (i.e., over time), invariance tests across gender at each wave and across the two waves were further performed based on the favored factor structure identified in the second step. Following previous practices [33, 59], invariance across gender or overtime was tested sequentially, including (a) configural invariance (free estimation across gender or overtime), (b) metric invariance (equal factor loadings), (c) scalar invariance (equal factor loadings plus equal item intercepts), and (d) strict invariance (equal factor loadings and item intercepts plus equal residual variances). Changes in CFI and RMSEA (i.e., ΔCFI < 0.01 and ΔRMSEA < 0.015) were used to indicate model invariance [60, 61].

Results

Comparisons of the two competing structural models

As shown in Table 2, the 15-primary-factor structure (Model 1) demonstrated adequate model fit at the two assessment points (CFI and TLI > 0.90; SRMR and RMSEA < 0.08) while the

Table 2. Summary of model fit for different CFA models at Wave 1 and Wave 2.

Model	Description	χ^2	df	CFI	TLI	SRMR	RMSEA (90% CI)	BIC	$\Delta\chi^2$	ΔCFI	ΔBIC
Wave 1 (n = 3,328)											
1	15 primary factor model	6,817.09	793	.92	.91	.04	.054 (.047, .049)	381,726.95			
2	4 higher-order model	9,307.27	877	.89	.89	.05	.048 (.053, .055)	383,535.88	2,490.18	−.03	1,808.93
Wave 2 (n = 3,638)											
1	15 primary factor model	7,523.99	793	.93	.92	.04	.048 (.047, .049)	398,631.70			
2	4 higher-order model	11,200.36	877	.89	.88	.05	.057 (.056, .058)	401,619.35	3,676.37	−.03	2,987.66

Note. df = degree of freedom, CFI = comparative fit index; TLI = Tucker-Lewis index, SRMR = standardized root mean square residual, BIC = Bayesian information criterion, RMSEA = root mean square error of approximation, CI = confidence interval, $\Delta\chi^2$ = change in χ^2, ΔCFI = change in CFI, ΔBIC = change in BIC.

https://doi.org/10.1371/journal.pone.0303531.t002

4-higher-order-factor structure (Model 2) showed relatively worse model fit as far as CFI and TLI were concerned (i.e., lower than 0.90). In addition, Model 1 showed a lower value of BIC than Model 2 across waves (ΔBIC = 1808.93 and 2987.66 at the two waves, respectively). Thus, the 15-primary-factor structure was retained and used for the subsequent invariance tests. Factor loadings on the 15 dimensions ranged between 0.60 and 0.85 with an average factor loading of 0.72 at Wave 1. At Wave 2, factor loadings ranged between 0.61 and 0.87 with an average factor loading of 0.75. In addition, the 15 dimensions as well as the whole scale showed adequate internal consistency and medium to high inter-factor cross-sectional correlations at both waves (see Table 3). In addition, each of the 15 dimensions also showed a moderate cross-time corrlelation with the respective score after one year (r = 0.50–0.69, ps < .001).

Table 3. Reliability of the Chinese Positive Youth Development Scale and correlations among factors at two waves.

Factor	Reliability		Descriptive		Correlations between factors at Wave 1 and Wave 2														
	α	MIIC	Mean/SD		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	W1/W2	W1/W2	W1	W2															
1. BO	.74/.77	.49/.53	4.70/.88	4.59/.90	–	.69	.63	.86	.63	.54	.59	.59	.55	.53	.56	.54	.78	.72	.59
2. RE	.79/.81	.55/.59	4.64/.91	4.57/.90	.65	–	.66	.62	.69	.73	.72	.64	.69	.65	.67	.71	.68	.66	.64
3. SC	.86/.88	.66/.71	4.74/.89	4.67/.89	.65	.60	–	.59	.70	.65	.67	.58	.61	.53	.70	.58	.63	.51	.53
4. PB	.76/.77	.52/.53	4.33/.98	4.20/.95	.84	.61	.65	–	.68	.58	.60	.62	.56	.57	.59	.54	.80	.72	.50
5. EC	.73/.74	.47/.49	4.26/.95	4.27/.91	.64	.73	.71	.70	–	.85	.86	.75	.71	.63	.70	.66	.70	.68	.51
6. CC	.81/.83	.59/.62	4.32/.91	4.34/.86	.52	.70	.64	.57	.87	–	.90	.73	.76	.68	.70	.70	.65	.60	.46
7. BC	.71/.76	.45/.52	4.54/.84	4.49/.82	.59	.72	.68	.65	.83	.85	–	.80	.81	.69	.72	.71	.70	.66	.50
8. MC	.73/.75	.48/.50	4.37/.91	4.38/.87	.60	.67	.57	.66	.75	.73	.80	–	.78	.67	.64	.67	.73	.80	.45
9. SD	.75/.78	.50/.55	4.47/.89	4.41/.89	.54	.67	.65	.57	.71	.76	.78	.77	–	.84	.84	.81	.69	.61	.50
10. SE	.65/.68	.48/.51	4.35/.93	4.33/.93	.51	.66	.55	.55	.64	.69	.67	.69	.82	–	.87	.81	.73	.63	.48
11. CPI	.78/.80	.54/.57	4.08/1.03	4.06/1.01	.56	.65	.71	.62	.68	.71	.71	.67	.83	.83	–	.96	.78	.57	.61
12. BF	.84/.85	.64/.66	4.38/1.06	4.26/1.05	.54	.71	.59	.55	.63	.66	.67	.68	.76	.73	.89	–	.82	.64	.59
13. PI	.80/.82	.57/.59	4.37/1.04	4.26/.99	.73	.65	.64	.74	.68	.60	.65	.65	.63	.61	.73	.71	–	.95	.64
14. PN	.72/.72	.46/.47	4.64/.96	4.50/.95	.59	.62	.47	.57	.57	.54	.59	.69	.53	.55	.54	.61	.75	–	.57
15. SP	.88/.89	.71/.73	4.45/1.10	4.32/1.08	.61	.65	.52	.54	.54	.49	.51	.50	.51	.48	.64	.60	.63	.51	–
Total	.96/.96	.37/.38	4.44/.70	4.38/.69															

Note. All parameters were significant in correlations (p < .001). Parameters below the diagonal represent correlations at Wave 1 and the above ones show correlations at Wave 2. MIIC = mean inter-item correlation, SD = standard deviation, W1 = Wave 1, W2 = Wave 2, BO = boding, RE = resilience, SC = social competence, PB = recognition for positive behavior, EC = emotional competence, CC = cognitive competence, BC = behavioral competence, MC = moral competence, SD = self-determination, SE = self-efficacy, CPI = clear and positive identity, BF = beliefs in future, PI = prosocial involvement, PN = prosocial norms, SP = spirituality.

https://doi.org/10.1371/journal.pone.0303531.t003

Table 4. Invariance tests across gender at two waves and longitudinal invariance tests over time.

	χ^2	df	CFI	TLI	SRMR	RMSEA (90% CI)	Compare	$\Delta\chi^2$	ΔCFI	Δdf	ΔRMSEA
Invariance tests across gender at Wave 1											
Boys (n = 1,740)	3,777.80	793	.93	.91	.03	.047 (.045, .048)					
Girls (n = 1,584)	4,334.25	793	.91	.89	.04	.053 (.052, .055)					
A. Configural	8,112.06	1,586	.92	.90	.04	.050 (.049, .051)					
B. Metric	8,170.18	1,615	.92	.90	.04	.049 (.048, .050)	B vs. A	58.12	.000	29	−.001
C. Scalar	8,433.21	1,644	.92	.90	.04	.050 (.049, .051)	C vs. B	263.03	−.003	29	.001
D. Strict	9,107.10	1,688	.91	.90	.05	.051 (.050, .052)	D vs. C	673.89	−.008	44	.001
Invariance tests across gender at Wave 2											
Boys (n = 1,895)	4,231.47	793	.93	.92	.03	.048 (.046, .049)					
Girls (n = 1,739)	4,347.04	793	.92	.90	.04	.051 (.049, .052)					
A. Configural	8,578.51	1,586	.93	.91	.04	.049 (.048, .050)					
B. Metric	8,697.74	1,615	.93	.91	.04	.049 (.048, .050)	B vs. A	119.23	−.001	29	.000
C. Scalar	8,970.22	1,644	.92	.91	.04	.050 (.049, .051)	C vs. B	272.49	−.003	29	.001
D. Strict	9,509.60	1,688	.92	.91	.05	.050 (.050, .051)	D vs. C	539.38	−.005	44	.000
Invariance tests over time (n = 4,061)											
Wave 1 (n = 3,328)	6,817.09	793	.92	.91	.04	.048 (.047, .049)					
Wave 2 (n = 3,638)	7,523.99	793	.93	.92	.04	.048 (.047, .049)					
A. Configural	17,028.77	3,253	.93	.91	.03	.032 (.032, .033)					
B. Metric	17,100.94	3,282	.93	.91	.03	.032 (.032, .033)	B vs. A	72.17	.000	29	.000
C. Scalar	17,453.52	3,311	.92	.91	.03	.032 (.032, .033)	C vs. B	352.58	−.002	29	.000
D. Strict	18,119.54	3,355	.92	.91	.04	.033 (.032, .033)	D vs. C	666.02	−.003	44	.001

Note. df = degree of freedom, CFI = comparative fit index; TLI = Tucker-Lewis index, SRMR = standardized root mean square residual, RMSEA = root mean square error of approximation, CI = confidence interval, $\Delta\chi^2$ = change in χ^2, ΔCFI = change in CFI, Δdf = change in df, ΔRMSEA = change in RMSEA.

https://doi.org/10.1371/journal.pone.0303531.t004

Invariance tests

Results of invariance tests across gender are presented in Table 4. The structure with 15 primary factors showed acceptable model fit in both gender groups at the two waves. Sequential invariance tests revealed that changes in CFI and RMSEA were less than the cut-offs (i.e., ΔCFI < 0.01 and ΔRMSEA < 0.015) in all pairs and comparisons. The results implied strict invariance characterized by equal factor loadings, item intercepts, and residual variances of the 15 factors across boys and girls with adequate model fit indices (Wave 1: $\chi^2_{(1688)}$ = 9107.10, CFI = 0.91, TLI = 0.90, RMSEA = 0.05; Wave 2: $\chi^2_{(1688)}$ = 9509.60, CFI = 0.92, TLI = 0.91, RMSEA = 0.05).

Results of longitudinal invariance tests over the two waves are also shown in Table 4. As changes in CFI and RMSEA were less than 0.01 and 0.015, respectively, in all pairs of model comparisons, longitudinal invariance of the 15-primary-factor structure was supported. The strict invariance model (i.e., equal factor loadings, item intercepts, residual variances over time) showed adequate model fit ($\chi^2_{(3355)}$ = 18119.54, CFI = 0.92, TLI = 0.91, RMSEA = 0.03).

To sum up, the above findings supported across-gender and longitudinal invariance of the 15-primary-factor structure of the 44-item CPYDS-SF.

Discussion and conclusions

The growing interest among researchers and practitioners in positive youth development (PYD) frameworks highlights the necessity for developing effective PYD assessment tools that are psychometrically sound, empirically helpful, and culturally applicable. Nevertheless,

findings from studies validating PYD measures among adolescents in Western contexts may not be entirely generalizable to non-Western settings, such as Chinese contexts. This research gap has been partially addressed by previous research that examined the validity and reliability of the 80-item long form of the "Chinese Positive Youth Development Scale" (CPYDS-LF). The research supported the scale's four-higher-order-factor structure and its invariance across gender groups [45]. The present study further addressed the aforementioned research gap by investigating factorial validity of the 44-item short form of the scale (CPYDS-SF) and measurement invariance across gender groups and over time among secondary school students in Hong Kong. Results showed that the CPYDS-SF is best represented by the 15-primary-factor structure (e.g., "social competence," "moral competence," and "resilience") rather than the four-higher-order-factor structure (e.g., "cognitive-behavioral competence" and "prosocial attributes"). Invariance tests provided support for factorial invariance of the 15-primary-factor structure across gender and over time. These findings provide support for the original conceptual model of the CPYDS.

The 80-item CPYDS-LF consists of 15 subscales that can be further subsumed into four higher-order factors. CPYDS-SF was formed by extracting items with the highest factor loadings in each subscale of CPYDS-LF. This approach has been commonly used in forming a short version of a measurement tool [62, 63]. This is because extracting items with the highest factor loadings can improve parsimoniousness while maintaining a high level of validity of the scale [64]. However, in the CPYDS-SF, the 15-primary-factor structure fitted the data better than the four-higher-order-factor structure. One explanation is that the retention of some specific items may reduce the common variance among items in different primary dimensions [65], thus making it difficult to give support for the higher-order factors. It is worth noting that items in the CPYDS-SF showed medium to high factor loadings and subscales showed adequate psychometric properties, which is consistent with features of other validated PYD measures [21, 22]. Therefore, CPYDS-SF is well represented by the 15-primary-factor structure.

Gender differences in PYD research fields have been a common topic of concern. For example, among the "Five Cs," Gomez-Baya et al. [40] found that boys scored higher in "*competence*" and "*confidence*", whereas girls scored higher in "*connection*," "*caring*," and "*character*". To correctly interpret such gender differences in terms of multiple possibilities, such as gender roles, educational systems, and cultural factors [66, 67], measurement invariance across gender should be first established to rule out the possibility of different measurement structures between girls and boys. Furthermore, measurement invariance of PYD assessment tools across gender is essential to correctly interpret gender differences in the effects of youth enhancement programs [67–69]. The present study found that the 15-primary-factor structure of the CPYDS-SF was invariant across gender groups. This is in line with the findings of measurement invariance for the CPYDS-LF and other PYD measures [e.g., 22, 26, 70] among Asian samples. These findings cumulatively indicate that boys and girls tend to interpret PYD measurement items in the same way. Thus, gender differences in PYD qualities and program effectiveness can be attributable to developmental differences rather than measurement differences between gender groups.

In addition to gender differences, it is also a common practice that PYD research investigates longitudinal development of adolescents' PYD attributes [71, 72]. Adolescence is associated with rapid biological, neurological, and cognitive development [35]. For example, enlargement of the ventromedial prefrontal cortex enables adolescents to have better abstract thinking and develop autonomy and independence [73, 74]. Due to these developmental changes, adolescents' perceptions and interpretations of themselves (e.g., their physical appearance, competence, and relationships) and things around them have been commonly found to

change during the adolescent period [75, 76]. However, as Hoyle and Smith [77] pointed out, the comparison across time points without longitudinal invariance measurement is a classic example of "comparing apples and oranges" (p. 433). In other words, we can only "promote a more thorough and unambiguous understanding of developmental stability and change" [78, p. 287] if "the invariance holds longitudinally" [79, p. 5]. As a result, examining longitudinal measurement invariance is a prerequisite to ensure that any difference shown in PYD qualities over time represents actual developmental changes rather than adolescents' different interpretations of assessment items [42, 80]. In this regard, the strict longitudinal invariance of the CPYDS-SF ensure that the scale measures the same PYD constructs on different occasions, enabling correct and meaningful investigations of longitudinal development or age group differences in PYD qualities among Chinese adolescents.

The present study has important implications in child and adolescent research. First, validation of the CPYDS-SF suggests that the short version can be effectively used in measuring Chinese adolescents' PYD qualities, shortening the time to fill in the questionnaire while also maintaining the validity and reliability of the measurement. This is much more likely to enhance schools' and adolescents' willingness to take part in and reduce participants' fatigue, especially considering that Chinese adolescents tend to have heavy schoolwork [81]. Second, the CPYDS-SF with strict gender and time invariance provides researchers with a valid and reliable assessment tool to investigate the development of PYD qualities, its antecedents (e.g., parental influence) as well as its contribution to adolescent developmental outcomes (e.g., academic and emotional well-being). Third, the CPYDS-SF also serves as a useful assessment tool for evaluating the effectiveness of PYD programs in Chinese communities. Specifically, scores obtained from the CPYDS-SF before and after program completion or between students participating in or not can be meaningfully compared to delineate potential impacts of program participation. Noteworthy, the long version of the scale has been successfully utilized in such program evaluation in both Hong Kong and mainland China [11, 82]. The current study further supports the similar employment of the short version in program evaluation.

Despite the pioneering nature of the study in the PYD literature, the present study has several limitations. First, only the data based on Grade 7 students over two waves were used. To portray a more holistic picture of the psychometric properties of the CPYDS-SF, adolescents from different age groups (e.g., students from senior high schools) can be further recruited in future studies. Second, the current sample only included adolescents in Hong Kong, thus making the present findings might not fully applicable to adolescent samples in other Chinese communities, such as different areas (e.g., rural and urban) in mainland China. Thus, future research is suggested to test the factor structure of the CPYDS-SF and its invariance in different adolescent samples. Third, the present study only validated the CPYDS-SF among adolescents. As PYD scales have also been widely utilized among emerging adults such as university students [69, 83], future studies will benefit from further validating CPYDS-SF among Chinese university students so that the usage of the scale can be extended. Fourth, the present study utilized data collected several years ago. Due to the rapid social and environmental changes, including the influence of COVID-19, adolescents' perceived PYD attributes may change as well. Thus, there is a need for replication using more recent data, which is important to ensure the stability and generalizability of the present findings and support the use of the CPYDS-SF in future studies. Nevertheless, the present studt can still give us a snapshot on the factorial invariance of the CPYDS-SF. Lastly, as we focused on examining the factorial validity and factorial invariance of the scale, we did not examine test-retest reliability, which is also an important factor of the scale's psychometric properties. While individual dimensions in the CPYDS-SF demonstrated moderate cross-time correlations after one year, which can be

regarded as support for test-retest reliability, futhre studies will certainly benefit from directly investigating the test-retest reliability.

Despite these limitations, findings in this study suggest that Hong Kong Chinese adolescents' manifestation of PYD qualities assessed by the CPYDS-SF was well represented by the 15 primary factors proposed by Catalano et al. [14]. In addition, the current 15-primary-factor model was strictly invariant between male and female adolescents over one year, implying that the factor structure of PYD qualities was equivalent between gender groups and stable over time. To conclude, the presenting findings suggest that the CPYDS-SF can be reliably utilized in investigating gender differences and longitudinal changes in PYD qualities among Chinese adolescents.

Supporting information

S1 File. All data underlying the present findings can be found in S1 File.
(XLSX)

Acknowledgments

We thank all the participating schools and students as well their parents for their support and collaboration.

Author Contributions

Conceptualization: Xiaoqin Zhu, Daniel T. L. Shek.

Formal analysis: Xiaoqin Zhu.

Funding acquisition: Daniel T. L. Shek.

Methodology: Xiaoqin Zhu.

Supervision: Daniel T. L. Shek.

Writing – original draft: Xiaoqin Zhu, Carman K. M. Chu.

Writing – review & editing: Xiaoqin Zhu, Xue Wu, Daniel T. L. Shek.

References

1. Shek DTL, Dou D, Zhu X, Chai W. Positive youth development: Current perspectives. Adolesc Health Med Ther. 2019; 10:131–41. https://doi.org/10.2147/AHMT.S179946 PMID: 31572041

2. Epstein MK, Renk K, Duhig AM, Bosco GL, Phares V. Interparental conflict, adolescent behavioral problems, and adolescent competence: convergent and discriminant validity. Educ Psychol Meas. 2004; 64 (3):475–95. https://doi.org/10.1177/0013164403258462

3. Walters GD, Espelage DL. Latent structure of early adolescent bullying perpetration: A taxometric analysis of raw and ranked scores. Educ Psychol Meas. 2019; 79(4):754–72. https://doi.org/10.1177/0013164418824142 PMID: 32655182

4. Joseph S, Wood A. Assessment of positive functioning in clinical psychology: Theoretical and practical issues. Clin Psychol Rev. 2010; 30(7):830–8. https://doi.org/10.1016/j.cpr.2010.01.002 PMID: 20137841

5. Lerner JV, Bowers E, Minor K, Boyd M, Mueller M, Schmid K, et al. Positive youth development: Processes, philosophies, and programs. In: Lerner RM, Easterbrooks MA, Mistry J, editors. Comprehensive handbook of psychology. 6. New York: Wiley; 2013. p. 365–92.

6. Anderson-Butcher D, Conroy DE. Factorial and criterion validity of scores of a measure of belonging in youth development programs. Educ Psychol Meas. 2002; 62(5):857–76. https://doi.org/10.1177/001316402236882

7. Zhu X, Shek D. The predictive effect of depression on self-injury: Positive youth development as a moderator. Appl Res Qual Life. 2023. https://doi.org/10.1007/s11482-023-10211-x

8. Dew RE, Daniel SS, Armstrong TD, Goldston DB, Triplett MF, Koenig HG. Religion/spirituality and adolescent psychiatric symptoms: A review. Child Psychiatry Hum Dev. 2008; 39(4):381–98. https://doi.org/10.1007/s10578-007-0093-2 PMID: 18219572

9. Humphrey N, Kalambouka A, Wigelsworth M, Lendrum A, Deighton J, Wolpert M. Measures of social and emotional skills for children and young people: A systematic review. Educ Psychol Meas. 2011; 71 (4):617–37. https://doi.org/10.1177/0013164410382896

10. Taylor RD, Oberle E, Durlak JA, Weissberg RP. Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. Child Dev. 2017; 88(4):1156–71. https://doi.org/10.1111/cdev.12864 PMID: 28685826

11. Zhu X, Shek DTL. Impact of a positive youth development program on junior high school students in mainland China: A pioneer study. Child Youth Serv Rev. 2020; 114:105022. https://doi.org/10.1016/j.childyouth.2020.105022

12. Holsen I, Geldhof J, Larsen T, Aardal E. The five Cs of positive youth development in Norway: Assessment and associations with positive and negative outcomes. Int J Behav Dev. 2017; 41(5):559–69. https://doi.org/10.1177/0165025416645668

13. Weissberg RP. Promoting the social and emotional learning of millions of school children. Perspect Psychol Sci. 2019; 14(1):65–9. https://doi.org/10.1177/1745691618817756 PMID: 30799753

14. Catalano RF, Berglund ML, Ryan JAM, Lonczak HS, Hawkins JD. Positive youth development in the United States: Research findings on evaluations of positive youth development programs. Ann Am Acad Pol Soc Sci. 2004; 591(1):98–124. https://doi.org/10.1177/0002716203260102

15. Shek DTL, Zhu X, Leung JTY, Lee TY, Wu FKY. Evaluation of the Project P.A.T.H.S. in mainland China: Findings based on student diaries. Res Soc Work Pract. 2019; 29(4):410–9. https://doi.org/10.1177/1049731517745994

16. Li X, He S, Chen B-B. The five Cs model of positive youth development and the PATHS project in mainland China. In: Dimitrova R, Wiium N, editors. Handbook of positive youth development: Advancing research, policy, and practice in global contexts. Cham: Springer International Publishing; 2021. p. 63–79.

17. Geldhof GJ, Bowers EP, Boyd MJ, Mueller MK, Napolitano CM, Schmid KL, et al. Creation of short and very short measures of the five Cs of positive youth development. J Res Technol Educ. 2014; 24 (1):163–76. https://doi.org/10.1111/jora.12039

18. Lopez A, Yoder JR, Brisson D, Lechuga-Pena S, Jenson JM. Development and validation of a positive youth development measure: The Bridge-Positive Youth Development. Res Soc Work Pract. 2015; 25 (6):726–36. https://doi.org/10.1177/1049731514534899

19. Phelps E, Zimmerman S, Warren AEA, Jeličić H, von Eye A, Lerner RM. The structure and developmental course of Positive Youth Development (PYD) in early adolescence: Implications for theory and practice. J Appl Dev Psychol. 2009; 30(5):571–84. https://doi.org/10.1016/j.appdev.2009.06.003

20. Tolan P. Future directions for positive development intervention research. J Clin Child Adolesc Psychol. 2014; 43(4):686–94. https://doi.org/10.1080/15374416.2014.936604 PMID: 25023002

21. Chen B-B, Wiium N, Dimitrova R. Factor structure of positive youth development: Contributions of exploratory structural equation modeling. Pers Individ Dif. 2018; 124:12–5. https://doi.org/10.1016/j.paid.2017.11.039

22. Chai X, Wang J, Li X, Liu W, Zhao G, Lin D. Development and validation of the Chinese Positive Youth Development Scale. Appl Dev Sci. 2022; 26(1):127–40. https://doi.org/10.1080/10888691.2020.1712206

23. Wong PW-C, Kwok K-W, Chow S-L. Validation of positive youth development scale and implications for adolescent in Hong Kong community. Child Youth Care Forum. 2022; 51(5):901–19. https://doi.org/10.1007/s10566-021-09658-6

24. Garst BA, Gagnon RJ, Whittington A. A closer look at the camp experience: Examining relationships between life skills, elements of positive youth development, and antecedents of change among camp camp alumni. J Outdoor Recreat Educ Leadersh. 2016; 8(2):180–99. https://doi.org/10.18666/jorel-2016-v8-i2-7694

25. Oman RF, Vesely SK, McLeroy KR, Harris-Wyatt V, Aspy CB, Rodine S, et al. Reliability and validity of the youth asset survey (YAS). J Adolesc Health. 2002; 31(3):247–55. https://doi.org/10.1016/s1054-139x(02)00363-4 PMID: 12225737

26. Yıldırım S, Arslan Y. Psychometric properties and measurement invariance of the positive youth development scale–Turkish version. Percept Mot Skills. 2023:00315125231206311. https://doi.org/10.1177/00315125231206311 PMID: 37847854

27. Chen HY, Cheng CL. Developmental trajectory of purpose identification during adolescence: Links to life satisfaction and depressive symptoms. Journal of Adolescence. 2020; 80:10–8. https://doi.org/10.1016/j.adolescence.2020.01.013 PMID: 32045759

28. Fernandez Castelao C, Kröner-Herwig B. Developmental trajectories and predictors of externalizing behavior: A comparison of girls and boys. Journal of Youth and Adolescence. 2014; 43:775–89. https://doi.org/10.1007/s10964-013-0011-9 PMID: 24002677

29. Leban L. The effects of adverse childhood experiences and gender on developmental trajectories of internalizing and externalizing outcomes. Crime & Delinquency. 2021; 67(5):631–61. https://doi.org/10.1177/0011128721989059

30. Tomé G, Gaspar de Matos M, Reis M, Gomez-Baya D, Coelhoso F, Wiium N. Positive youth development and wellbeing: Gender differences. Frontiers in Psychology. 2021; 12. https://doi.org/10.3389/fpsyg.2021.641647 PMID: 34354625

31. Seddig D, Leitgöb H. Approximate measurement invariance and longitudinal confirmatory factor analysis: Concept and application with panel data. Surv Res Methods. 2018; 10(1):29–41. https://doi.org/10.18148/srm/2018

32. Dou D, Shek DTL, Zhu X, Zhao L. Dimensionality of the Chinese CES-D: Is it stable across gender, time, and samples? Int J Environ Res Public Health. 2021; 18(22):11818. https://doi.org/10.3390/ijerph182211818 PMID: 34831573

33. Zhu X, Shek DTL, Dou D. Factor structure of the Chinese CES-D and invariance analyses across gender and over time among Chinese adolescents. J Affect Disord 2021; 295:639–46. https://doi.org/10.1016/j.jad.2021.08.122 PMID: 34509779

34. Wen M, Ye Z, Lin D, Wang W. Preliminary development of a multidimensional positive youth development scale for young rural and urban adolescents in China. PloS one. 2022; 17(7):e0270974. https://doi.org/10.1371/journal.pone.0270974 PMID: 35901101

35. Widaman KF, Ferrer E, Conger RD. Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. Child Dev Perspect. 2010; 4(1):10–8. https://doi.org/10.1111/j.1750-8606.2009.00110.x PMID: 20369028

36. Millsap RE, Cham H. Investigating factorial invariance in longitudinal data. In: Laursen B, Little TD, Card NA, editors. Handbook of developmental research methods: The Guilford Press; 2012. p. 109–26.

37. Luo J, Wang X, Wang M-C, Zhang X, Deng J, Zhong C, et al. Longitudinal measurement invariance of the Child Problematic Trait Inventory in older Chinese children. PLoS One. 2019; 14(7):e0219136. https://doi.org/10.1371/journal.pone.0219136 PMID: 31269064

38. Luo J, Wang M-C, Ge Y, Chen W, Xu S. Longitudinal invariance analysis of the Short Grit Scale in Chinese young adults. Front Psychol. 2020; 11:466. https://doi.org/10.3389/fpsyg.2020.00466 PMID: 32265788

39. Bowers EP, Li Y, Kiely MK, Brittian A, Lerner JV, Lerner RM. The five Cs model of positive youth development: A longitudinal analysis of confirmatory factor structure and measurement invariance. Journal of Youth and Adolescence. 2010; 39(7):720–35. https://doi.org/10.1007/s10964-010-9530-9 PMID: 20397040

40. Gomez-Baya D, Reis M, Gaspar de Matos M. Positive youth development, thriving and social engagement: An analysis of gender differences in Spanish youth. Scand J Psychol. 2019; 60(6):559–68. Epub 2019/09/20. https://doi.org/10.1111/sjop.12577 PMID: 31535399.

41. Pivec T, Kozina A. Anxiety and COVID-19 anxiety in positive youth development: A latent profile analysis study. Journal of Youth and Adolescence. 2023; 52(11):2328–43. https://doi.org/10.1007/s10964-023-01829-z PMID: 37500813

42. Chan MHM, Gerhardt M, Feng X. Measurement invariance across age groups and over 20 years' time of the Negative and Positive Affect Scale (NAPAS). Eur J Psychol Assess. 2020; 36(4):537–44. https://doi.org/10.1027/1015-5759/a000529

43. Brown TA. Confirmatory factor analysis for applied research. Second ed. New York: Guilford Publications; 2015.

44. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Dev Rev. 2016; 41:71–90. https://doi.org/10.1016/j.dr.2016.06.004 PMID: 27942093

45. Shek DTL, Ma CMS. Dimensionality of the Chinese Positive Youth Development Scale: Confirmatory factor analyses. Soc Indic Res. 2010; 98(1):41–59. https://doi.org/10.1007/s11205-009-9515-9

46. Sun RCF, Shek DTL. Positive youth development, life satisfaction and problem behaviour among Chinese adolescents in Hong Kong: A replication. Soc Indic Res. 2012; 105(3):541−59. https://doi.org/10.1007/s11205-011-9786-9 PMID: 22247583

47. Ahn JS, Plamondon A, Bouchard M, Denault A-S. Validation of the French-Canadian Version of the Positive Youth Development Scale-Short Form (PYD-SF). Applied Developmental Science.1–14. https://doi.org/10.1080/10888691.2023.2215988

48. Bello B, Flynn E, Geldhof GJ, Yu D, Mueller MK, Licardi K, et al. Measuring social and emotional functioning as a facet of positive youth development among children and adolescents in special education and mental health treatment. Social Development.e12738. https://doi.org/10.1111/sode.12738

49. Stopa L, Thorne P, Waters A, Preston J. Are the short and long forms of the Young Schema Questionnaire comparable and how well does each version predict psychopathology scores?. Journal of Cognitive Psychotherapy. 2001; 15(3):253–61.

50. Lin L, Shek DTL, Li X. Who benefits and appreciates more? An evaluation of Online Service-Learning Projects in Mainland China during the COVID-19 pandemic. Appl Res Qual Life. 2023; 18:625–46. https://doi.org/10.1007/s11482-022-10081-9 PMID: 35873305

51. Zhu X, Shek DTL. Promoting service leadership qualities and well-being among university students: Evaluation findings in Hong Kong. Appl Res Qual Life. 2021; 16:2475–95. https://doi.org/10.1007/s11482-021-09922-w

52. Muthén LK, Muthén BO. Mplus user's guide. 8th ed. Los Angeles, CA: Muthén & Muthén; 2018.

53. Finney SJ, DiStefano C. Non-normal and categorical data in structural equation modeling. In: Hancock GR, Mueller RO, editors. Structural equation modeling: A second course. Greenwich, CT: Information Age Publishing; 2006. p. 269–314.

54. Cham H, Reshetnyak E, Rosenfeld B, Breitbart W. Full information maximum likelihood estimation for latent variable interactions with incomplete indicators. Multivariate Behav Res. 2017; 52(1):12–30. https://doi.org/10.1080/00273171.2016.1245600 PMID: 27834491

55. Kline RB. Principles and practice of structural equation modeling. Fourth ed. New York: The Guilford Press; 2015.

56. Meade AW, Johnson EC, Braddy PW. Power and sensitivity of alternative fit indices in tests of measurement invariance. J Appl Psychol. 2008; 93(3):568–92. https://doi.org/10.1037/0021-9010.93.3.568 PMID: 18457487

57. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. Methods Psychol Rea Online. 2003; 8(2):23–74.

58. Raftery AE. Bayesian model selection in social research. Sociol Methodol. 1995; 25:111–63. https://doi.org/10.2307/271063

59. Svetina D, Rutkowski L, Rutkowski D. Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using Mplus and the lavaan/semtools packages. Struct Equ Modeling. 2020; 27(1):111–30. https://doi.org/10.1080/10705511.2019.1602776

60. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling. 2002; 9(2):233–55. https://doi.org/10.1207/S15328007SEM0902_5

61. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct Equ Modeling. 2007; 14(3):464–504. https://doi.org/10.1080/10705510701301834

62. Shek DTL, Zhu X, Chan K-M. Development of service leadership behavior scale: Background and conceptual model. Int J Child Adolesc health. 2018; 11(4):415–24.

63. Shek DTL, Chai WY. Psychometric properties of the service leadership attitude scale in Hong Kong. Front Psychol. 2019; 10:1070. https://doi.org/10.3389/fpsyg.2019.01070 PMID: 31133950

64. Thornton A, Binstock G, Young-DeMarco L, Mitchell C, Yount KM, Xie Y. Evaluating the measurement reliabilities and dimensionality of developmental idealism measures. Chin J Sociol. 2016; 2(4):609–35. Epub 2016/10/01. https://doi.org/10.1177/2057150X16672813 PMID: 29560276; PubMed Central PMCID: PMC5857385.

65. Yong AG, Pearce S. A beginner's guide to factor analysis: Focusing on exploratory factor analysis. Tutor Quant Methods Psychol. 2013; 9(2):79–94. https://doi.org/10.20982/tqmp.09.2.p079

66. Viner RM, Ozer EM, Denny S, Marmot M, Resnick M, Fatusi A, et al. Adolescence and the social determinants of health. Lancet. 2012; 379(9826):1641–52. https://doi.org/10.1016/S0140-6736(12)60149-4 PMID: 22538179

67. Årdal E, Holsen I, Diseth Å, Larsen T. The five Cs of positive youth development in a school context; gender and mediator effects. Sch Psychol Int. 2018; 39(1):3–21. https://doi.org/10.1177/0143034317734416

68. Muturi N, Kidd T, Daniels AM, Kattelmann KK, Khan T, Lindshield E, et al. Examining the role of youth empowerment in preventing adolescence obesity in low-income communities. J Adolesc. 2018; 68 (1):242–51. https://doi.org/10.1016/j.adolescence.2018.08.001 PMID: 30212755

69. Gomez-Baya D, Santos T, Gaspar de Matos M. Developmental assets and positive youth development: An examination of gender differences in Spain. Appl Dev Sci. 2022; 26(3):516–31. https://doi.org/10.1080/10888691.2021.1906676

70. Buenconsejo JU, Datu JAD, Chiu MM, Chan RCH. Psychometric validity and measurement invariance of positive youth development in the Philippines during the COVID-19 pandemic. Appl Dev Sci. 2023; 27(4):336–51. https://doi.org/10.1080/10888691.2022.2078719

71. Hull DM, Ferguson SL, Fagan MA, Brown EA. Positive youth development in Jamaica: Latent growth of self-efficacy and youth assets. In: Dimitrova R, Wiium N, editors. Handbook of Positive Youth Development. Cham: Springer; 2021. p. 299–313.

72. Kaniušonytė G, Leggett-James MP, Laursen B. A longitudinal examination of perceived parent behavior and positive youth development: Child-driven effects. Soc Sci. 2021; 10(10):369. https://doi.org/10.3390/socsci10100369

73. Branje S. Development of parent-adolescent relationships: Conflict interactions as a mechanism of change. Child Dev Perspect. 2018; 12(3):171–6. https://doi.org/10.1111/cdep.12278

74. Pfeifer JH, Kahn LE, Merchant JS, Peake SJ, Veroude K, Masten CL, et al. Longitudinal change in the neural bases of adolescent social self-evaluations: Effects of age and pubertal development. J Neurosci. 2013; 33(17):7415–9. Epub 2013/04/26. https://doi.org/10.1523/JNEUROSCI.4074-12.2013 PMID: 23616547; PubMed Central PMCID: PMC3809090.

75. Marsh HW, O'Mara A. Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. Pers Soc Psychol Bull. 2008; 34(4):542–52. Epub 2008/03/15. https://doi.org/10.1177/0146167207312313 PMID: 18340036.

76. Kuzucu Y, Bontempo DE, Hofer SM, Stallings MC, Piccinin AM. Developmental change and time-specific variation in global and specific aspects of self-concept in adolescence and association with depressive symptoms. J Early Adolesc. 2014; 34(5):638–66. Epub 2014/08/22. https://doi.org/10.1177/0272431613507498 PMID: 25143664; PubMed Central PMCID: PMC4136385.

77. Hoyle RH, Smith GT. Formulating clinical research hypotheses as structural equation models: A conceptual overview. J Consult Clin Psychol. 1994; 62(3):429–40. https://doi.org/10.1037//0022-006x.62.3.429 PMID: 8063970

78. Obradović J, Pardini DA, Long JD, Loeber R. Measuring interpersonal callousness in boys from childhood to adolescence: An examination of longitudinal invariance and temporal stability. J Clin Child Adolesc Psychol. 2007; 36(3):276–92. https://doi.org/10.1080/15374410701441633 PMID: 17658974

79. Millsap RE. Testing measurement invariance using item response theory in longitudinal data: An introduction. Child Dev Perspect. 2010; 4(1):5–9. https://doi.org/10.1111/j.1750-8606.2009.00109.x

80. Van De Schoot R, Schmidt P, De Beuckelaer A, Lek K, Zondervan-Zwijnenburg M. Editorial: Measurement Invariance. Front Psychol. 2015; 6:1064. Epub 2015/08/19. https://doi.org/10.3389/fpsyg.2015.01064 PMID: 26283995; PubMed Central PMCID: PMC4516821.

81. Salili F, Lai MK, Leung SSK. The consequences of pressure on adolescent students to perform well in school. Hong Kong J Paediatr. 2004; 9(4):329–36.

82. Shek DTL, Ma CMS. Impact of the Project P.A.T.H.S. in the junior secondary school years: Individual growth curve analyses. Sci World J. 2011; 11:253–66. https://doi.org/10.1100/tsw.2011.6 PMID: 21298216

83. Abdul Kadir NBy, Mohd RH. The 5Cs of positive youth development, purpose in life, hope, and well-being among emerging adults in Malaysia. Front Psychol. 2021; 12:641876. https://doi.org/10.3389/fpsyg.2021.641876 PMID: 34335359