

Habibi Asgarabad, M., Salehi Yegaei, P., Azami-Aghdash, S., Zaru, R., Hadi, N., Martin, N., & Trejos-Castillo, E. (2024). A Protocol Study for a Systematic Review and Meta-analysis of the Psychometric Properties of the Five Cs Model of Positive Youth Development Scales. In N. Wiium, D. Manrique-Millones, D. Miconi, & D. Stefenel (Eds.), *Addressing Social Justice: A Positive Youth Development Approach* (pp. 21–46). Fagbokforlaget.  
<https://doi.org/10.55669/oa311001>

## Chapter 1

# A Protocol Study for a Systematic Review and Meta-analysis of the Psychometric Properties of the Five Cs Model of Positive Youth Development Scales

Mojtaba Habibi Asgarabad,<sup>1\*</sup> Pardis Salehi Yegaei,<sup>2</sup> Saber Azami-Aghdash,<sup>3</sup> Rula Zaru,<sup>4</sup> Narges Hadi,<sup>4</sup> Nicole Martin, and Elizabeth Trejos-Castillo<sup>4</sup>

<sup>1</sup> Department of Psychology, Norwegian University of Science and Technology, Norway

<sup>2</sup> Health Promotion Research Center, Iran University of Medical Sciences, Tehran, Iran

<sup>3</sup> Tabriz Health Services Management Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

<sup>4</sup> Human Development & Family Sciences, Texas Tech University, Texas, USA

\* Corresponding author: [Mojtaba.h.asgarabad@ntnu.no](mailto:Mojtaba.h.asgarabad@ntnu.no)

**Abstract:** With the strength-based perspective of Positive Youth Development (PYD) surging ahead with great force in the field of developmental psychology, it is believed that PYD programs will facilitate the development of five PYD indicators: Competence, Confidence, Character, Connection, and Caring. To gauge the effectiveness of PYD programs on youth outcomes, the Five Cs Model of PYD Scale (PYDS) and its short versions, including the 34-item PYDS short form, the 17-item PYDS very short form, the 20-item PYDS-SF Shortened Version, the 78-item Five Cs of PYD scale, and the 40-item Bridge-PYD Instrument have been developed. However, no systematic review of their psychometric quality has been conducted. The objective of the present proposed systematic review and meta-analysis is to evaluate, summarize, and synthesize studies on the measurement properties (validity, reliability, responsiveness, and interpretability, along with the criteria of good measurement properties) of these versions. In compliance with the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P), the search will be carried out in PubMed, Scopus, Web of Science, and PsycINFO databases from the inception (2005) to 12<sup>th</sup> March 2024. Original articles published in English language on youth ages 10–29 years will be included if they report the reliability or validity, use quantitative methods, are systematic reviews on the Five Cs Model of PYD scales, or use them to assess study outcomes/predictors. Separate evaluators will assess research papers, extract data, and appraise their quality utilizing the CONsensus-based Standards for the selection of health Measurement INSTRUMENTS (COSMIN) tool. Any discrepancy will be resolved by three senior reviewers. The Grading of Recommendations, Assessment, Development and Evaluations (GRADE) guidelines will be applied to grade the total quality of evidence for the overall rating on each psychometric property. The critical and rigorous quality assessment of these scales is essential to clarify if the results of previous studies were biased and to guide scholars on what version of the Five Cs Model of PYD scales is the most valid, reliable, and appropriate for their research. Such evidence-based judgements on these scales' strengths and weaknesses are crucial to conduct future valid and rigorous research and suggest possible improvements on available measures.

*Keywords:* psychometric properties, reliability, systematic review protocol, the five Cs model of PYD scale, validity.

## Introduction

Throughout the past century, the “deficit perspective” described adolescence as a troublesome period (Erikson, 1968; Freud, 1969; Hall, 1905) that needs to be tamed (Roth & Brooks-Gunn, 2003). Therefore, more conservative developmental research and interventions in adolescence concentrated entirely on risks and vulnerabilities rather than characteristics and experiences that lead to youth thriving, positive developmental outcomes and well-being (Moore et al., 2004). From this deficit perspective, positive development of adolescence was characterized by the absence or reduction of problem behaviors (Crocetti et al., 2014). In response to such a problem-oriented approach, the ecological and asset-based perspective of Positive Youth Development (PYD; Lerner et al., 2005; Lerner et al., 2011) was developed with the purpose of comprehending, instructing, and engaging youth in productive activities. PYD particularly concentrates on empowering youth to take initiative in achieving important social, emotional, and moral competencies, a clear and positive sense of identity, self-efficacy, and self-determination. These are necessary steps toward flourishing and enhancing youths’ ability to contribute to their community (Catalano et al., 2002; Damon, 2004; Riggs et al., 2010; Schulman & Davies, 2007).

### **The Five Cs Model of Positive Youth Development**

One of the most prominent conceptualization models in the field of PYD is the Five Cs Model (Lerner & Lerner, 2013; Lerner et al., 2011), which posits the existence of five developmental indicators. First, Competence is defined as the positive image of individual’s activities in the domains of social skills (e.g., empathy), cognitive skills (e.g., decision-making), academic performance (e.g., test scores, school attendance, and school grades), and vocational competencies (work habits and career choice explorations, inclusive of entrepreneurship). Second, Confidence concerns a sense of self-efficacy, self-respect, and positive identity. Third, Character refers to possession of standard ideas for appropriate behaviors, showing respect to cultural and societal standards, and an overall sense of moral correctness. Fourth, Connection pertains to relationships with others, positive bidirectional interaction with family, peers, institutions, and community. Finally, Caring is defined as a sense of

empathy and sympathy for others. These domains have reciprocal influence on each other, and healthy development across all five indicators is crucial for adaptive development (Lerner et al., 2005).

Various measurement tools have been designed, translated, and employed in different studies worldwide to capture different aspects of PYD (e.g., Klein et al., 2006; Sieng et al., 2018). The Five Cs Model of PYD Scale-Long Form (PYDS; 83-item; Lerner et al., 2005; Phelps et al., 2009) has gained ground in the PYD field and a series of short versions have been produced, including the 34-item PYDS short form (PYDS-SF; Geldhof et al., 2014), the 17-item PYDS very short form (PYDS-VSF; Geldhof et al., 2014), the 20-item PYDS-SF Portuguese shortened version (Gaspar de Matos et al., 2018), the 78-item Five Cs of PYD scale (Phelps et al., 2009), and the 40-item Bridge-PYD Instrument (Lopez et al., 2015). The original self-report PYDS measures five independent indicators of Competence (subscales = academic, social, and physical competence, and grades), Confidence (subscales = self-worth and positive identification), Connection (subscales = connection to family, neighborhood, school, and peers), Character (subscales = social conscience, values diversity, conduct behavior, and personal values), and Caring. To calculate the five Cs scores, first, items of each subscale are averaged, and then scores across subscales of each “C” are averaged to provide the scores of each C, except for Caring, which has no subscale; its score is equal to the mean of its items. In the last step, each C is multiplied by 8.33 to convert its score to a 100-point scale. Accordingly, PYDS-SF and PYDS-VSF capture the same characteristics, including Competence (6 and 3 items), Confidence (6 and 3 items), Connection (8 and 4 items), Character (8 and 4 items), and Caring (6 and 3 items), with separate forms for youths in grades 5–7 and youths in grades 8–12 (Geldhof et al., 2014; Geldhof et al., 2014). Items corresponding to the Competence and Confidence subscales are evaluated using a four-point Likert scale, while a five-point Likert scale is used for items corresponding to the Connection, Character, and Caring subscales. An overview of the characteristics of the Five Cs Model of PYD Scales is presented in Table 1.1.

Evidence has documented the psychometric soundness of the 83- and 78-item Five Cs Model of PYD Scale in terms of internal consistency ( $\alpha = .84-.91$ ; Gestsdottir et al., 2010; Lerner et al., 2005), structural validity (Bowers et al., 2010; Lerner et al., 2005; Phelps et al., 2009), and measurement invariance across time (Bowers et al., 2010; Phelps et al., 2009). The findings of Gaspar

de Matos et al. (2018) on the 20-item PYDS-SF Shortened Version supported its structural validity via exploratory and confirmatory factor analysis, and its internal consistency using Cronbach's alpha (.74–.87) and composite reliability (.74–.85). Additionally, Lopez et al. (2015) in their study revealed that the 40-item Bridge-PYD Instrument can be applied to youths with a diverse age range, that guarantees its generalizability and possesses excellent reliability ( $\alpha = .92$ ). The PYDS-SF and PYDS-VSF also retain acceptable psychometric properties of the long form, confirming the factorial structure of PYDS (Geldhof et al., 2014) with satisfactory reliability ( $\alpha = .73$ –.87; Babaei et al., 2018; Tomé et al., 2019). However, results in some countries have raised some concerns regarding the psychometric properties of PYDS-VSF, including weak internal consistency of its subscales in the Philippines (Buenconsejo et al., 2022), or lack of support for its construct validity in China (Wong et al., 2022). There have also been general methodological concerns regarding the abbreviated tools (Larsson & Larsson, 2002), e.g., their inability to meet reasonable standards of psychometric soundness (Smith & McCarthy, 1995; Smith et al., 2000). In essence, summarizing and comparing the existing findings on the short forms derived from the PYDS is warranted.

Robust evidence has documented the positive association of the Five Cs with adaptive outcomes, e.g., intentional self-regulation (Gestsdottir et al., 2010; Gestsdottir et al., 2017), psychological empowerment (Holsen et al., 2017), life satisfaction (Gaspar de Matos et al., 2018), and academic well-being (Shek & Chai, 2020), as well as their negative link to substance use, delinquent behavior, anxiety/depression, and emotion dysregulation (Conway et al., 2015; Dvorsky et al., 2018). Whether or not these results are reliable depends largely on the psychometric soundness of the instruments used. Thus, a systematic review of the methodological rigor of research papers on their validity is warranted for multiple reasons. First, it would provide researchers, clinicians, and health policy makers with critical and rigorous assessments of the scales to reduce a potential risk of bias in results (Terwee et al., 2016). Second, a meta-analysis would provide researchers with useful information to decide which version of the Five Cs Model of PYD measure is the most valid, reliable, and appropriate form to use in their studies. Third, a systematic review of the potential weaknesses of available measures is key to inform the development of new scales and/or the improvement of existing measures.

In our increasingly globalized and diverse world, ensuring the cross-cultural validity of PYD scales across different countries, ethno-cultural groups and languages is crucial; it is equally important to acknowledge the impact of power dynamics and structural barriers in societies within a social justice perspective. This involves the validation of scales across populations, particularly minority and vulnerable groups, and acknowledging variations in how measures work with different populations.

Such variations can impact the validity of findings in some marginalized or minority populations, and researchers should account for these factors when designing and validating scales to contribute to the promotion of inclusion and social justice. Factors such as systemic racism and discrimination can also impact the validity of PYD scales; thus, researchers must ensure cultural sensitivity and appropriateness across populations. Ultimately, taking these issues into account will enable us to better promote PYD and support adolescent development in diverse contexts.

Whereas there are systematic reviews available on the effectiveness of PYD intervention or prevention programs (e.g., Catalano et al., 2019; Shek & Yu, 2011; Tidmarsh et al., 2022), there is a dearth of reviews and meta-analyses that conduct comprehensive appraisal of the quality of the Five Cs Model of PYD Scales. According to our current understanding, the present systematic review and meta-analysis is the first of its kind that aims to critically evaluate, summarize, and synthesize the available body of research on the measurement properties of the Five Cs Model of PYD scales. As a prerequisite, such studies must possess high-quality methodology and minimal risk of bias to ensure the precision of drawn conclusions regarding the psychometric properties of the tools (de Vet et al., 2011; Terwee et al., 2009). For that purpose, we will follow COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN; Mokkink, de Vet, et al., 2018) criteria (see Tables 1.2 and 1.3). The estimates of the measurement properties will be synthesized to address the following research questions: (1) What sort of evidence is available on the methodological quality (as listed in Table 1.2) of the Five Cs Model of PYD Scales? (2) Which one of the reported psychometric properties of these tools meets the criteria for good measurement properties (as listed in Table 1.3)? and (3) Which of these tools, according to the feasibility of performing a meta-analysis, are the most appropriate for measuring the Five Cs Model of PYD?

**Table 1.1** Overview of PYDS Long, Short, and Very Short Forms and Their Characteristics.

Scale	Developer	Target population	Administration method	Number of items	Subscales (number of items)	Scoring type
The PYDS Long Form	Lerner et al. (2005)	adolescents aged 10–18	self-report	83	1. Character (21 items) 2. Competence (19 items) 3. Caring (9 items) 4. Connection (22 items) 5. Confidence (12 items)	Likert scale
The PYDS Short Form	Geldhof et al. (2014)	separate forms for adolescents aged 11–14, and adolescents aged 14–21	self-report	34	1. Character (8 items) 2. Competence (6 items) 3. Caring (6 items) 4. Connection (8 items) 5. Confidence (6 items)	Likert scale
The PYDS Very Short Form	Geldhof et al. (2014)	separate forms for adolescents aged 11–14, and adolescents aged 14–21	self-report	17	1. Character (4 items) 2. Competence (3 items) 3. Caring (3 items) 4. Connection (4 items) 5. Confidence (3 items)	Likert scale
The PYDS-SF Portuguese Shortened Version	Gaspar de Matos et al. (2018)	youths aged 16 to 29	self-report	20	1. Character (4 items) 2. Competence (4 items) 3. Caring (4 items) 4. Connection (4 items) 5. Confidence (4 items)	Likert scale
The Five Cs of PYD	Phelps et al. (2009)	adolescents aged 10–12	self-report	78	1. Character (20 items) 2. Competence (11 items) 3. Caring (9 items) 4. Connection (22 items) 5. Confidence (16 items)	Likert scale
The Bridge-PYD Instrument	Lopez et al. (2015)	adolescents aged 7 to 18	self-report	40	1. Character (9 items) 2. Competence (8 items) 3. Caring and Compassion (7 items) 4. Connection (7 items) 5. Confidence (9 items)	Yes/no

Notes. PYDS = Positive Youth Development Scale,  $\alpha$  = Cronbach's Alpha.

**Table 1.2** *The Definitions of the Measurement Properties' Domains, Sub-domains, and Aspects Based on COSMIN (Mokkink et al., 2010a; Mokkink et al., 2010b).*

Domain	Measurement property	Aspects	Definition
Reliability			Presenting consistent scores in similar conditions
	1. Internal consistency		The extent the items of a tool are inter-correlated
	2. Measurement error		The dissimilarity between the unobserved true value of a variable and its observed value
Validity			The extent that a tool assesses the construct it is intended to assess
	1. Content validity		A tool's ability to adequately capture all aspects of the intended latent construct
		1.1 Face validity	The extent that a tool's items truly reflect the construct they were developed to measure
	2. Construct validity/ Hypotheses testing		The level that the tool's scores are compatible with hypotheses or theoretical model
		2.1 Structural validity	How well a tool can reflect the dimensionality of the latent construct it was developed to measure
		2.2 Cross-cultural validity/measurement invariance	How well a tool can function like the original tool in various cultural groups
	3. Criterion validity		The extent to which the tool's scores are correlated with a "gold standard" measure
	4. Known-group validity		How well the PROM can distinguish the subgroups of a sample
Responsiveness			The capability of a tool to discover the change in the scores of the aimed construct over time
	1. Longitudinal validity		How well a tool can discover the "expected changes" in the scores of the aimed construct over time
Interpretability			How well the qualitative meaning (i.e., everyday understanding or clinical meaning) can be assigned to a tool's scores

*Notes.* Interpretability is not one of the measurement properties, but a critical characteristic of a measure.



**Table 1.3** COSMIN Criteria for Good Measurement Properties (Mokkink et al., 2018b; Prinsen et al., 2018).

Measurement property	Rating	Criteria
Structural validity		1. CTT: One of the following criteria in CFA should be met: <ul style="list-style-type: none"> <li>• CFI/ TLI/ comparable measure &gt;.95</li> <li>• SRMR &lt;.08</li> <li>• RMSEA &lt;.06</li> </ul>
	+	2. IRT/Rasch: All the following criteria should be met: <ul style="list-style-type: none"> <li>• No violation of unidimensionality: CFI/ TLI/ comparable measure &gt;.95 OR SRMR &lt;.08 OR RMSEA &lt;.06</li> <li>• No violation of local independence: residual correlations among the items after controlling for the dominant factor &lt;.20 OR Q3's &lt;.37</li> <li>• No violation of monotonicity: adequate looking graphs/ item scalability &gt;.30 and adequate model fit: IRT: <math>\chi^2 &gt; .01</math>; Rasch: infit and outfit mean squares <math>\geq .5</math> and <math>\leq 1.5</math> OR Z standardized values &gt; -2 and &lt;2</li> </ul>
Internal consistency	?	1. CTT: Not reporting all information for “+” 2. IRT/Rasch: Not reporting model fit
	-	Not meeting criteria for “+”
	+	At least low evidence for sufficient structural validity (assessed via GRADE) and $\alpha(s) \geq 0.70$ for each unidimensional scale or subscale
	?	Not meeting criteria for low evidence or higher
	-	At least low evidence for sufficient structural validity and Cronbach's alpha(s) <.70 for each unidimensional scale or subscale
Reliability	+	ICC/ weighted Kappa $\geq .70$
	?	ICC/ weighted Kappa not reported
	-	ICC/ weighted Kappa <.70
Measurement error	+	SDC/ LoA < MIC
	?	MIC is not defined
	-	SDC/ LoA > MIC
Hypotheses testing/ construct validity	+	Most evidence in studies (75% or more) are in line with the hypotheses
	?	Hypotheses are not defined (by review team)
	-	Evidence is not in line with the hypotheses

Cross-cultural validity/measurement invariance	+	No important difference between group factors (e.g., age) in multiple group factor analysis OR no important DIF for group factors (McFadden's $R^2 < .02$ )
	?	The absence of multiple group factor analysis or DIF analysis
	-	Important differences between group factors or DIF
Criterion validity	+	Correlation with gold standard/ AUC $\geq .70$
	?	Not reporting all information for sufficient rating
	-	Correlation with gold standard/ AUC $< .70$
Responsiveness	+	Results are in line with the hypothesis or AUC $\geq .70$
	?	No hypotheses were defined (by review team)
	-	Results are not in line with the hypotheses or AUC $< .70$

Notes. “+” = Sufficient, “?” = Indeterminate, “-” = Insufficient, CTT = Classical test theory, CFA = Confirmatory factor analysis, IRT = Item response theory, CFI = Comparative fit index, TLI = Tucker-Lewis index, RMSEA = Root mean square error of approximation, SRMR = Standardized root mean residuals,  $\alpha$  = Cronbach's alpha, GRADE = Grading of Recommendations, Assessment, Development and Evaluations, ICC = Intra-class correlation coefficient, SDS = Smallest detectable change, LoA = Limits of agreement, MIC = Minimal important change, DIF = Differential item functioning, AUC = Area under curve.

## Materials and Methods

### Inclusion Criteria

The proposed systematic review and meta-analysis aligns with the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P; Page et al., 2021). Studies will be included in this study if they: 1) used the Five Cs Model of PYD Scales, 2) were published in or after 2005 (when the first version of PYDS was developed), 3) were published in the English language, and 4) focused on youth ages 10–29 years based on the suggestion of Catalano et al. (2019). They also must have either: 1) provided quantitative reports on the development and/or user rating (acceptability or appropriateness) of the Five Cs Model of PYD Scales, 2) reported the reliability or stability (test re-test, inter-rater, internal consistency, and multiple method reliability) of the Five Cs Model of PYD scales, 3) reported the validity (content, structural, construct, response process, and item functioning validity) of the Five Cs Model of PYD scales, 4) used quantitative methods and all types of study designs – namely, cohort studies, randomized/non-randomized trials, case-control studies, on-group, post-intervention, and cross-sectional studies, 5) used the Five Cs Model of PYD scales to assess study outcomes and/or risk factors (predictor variable), or 6) were systematic reviews on the Five Cs Model of PYD scales. The following studies will be excluded: 1) non-empirical research (editorials and theoretical framework discussions), 2) literature reviews, and 3) qualitative studies that aimed at developing, translating, adapting, and improving the Five Cs Model of PYD scales. Studies categorized as grey literature will be included if they can meet our eligibility criteria.

### Information Sources

By conducting a literature review using keywords, studies will be identified in three main domains: “Positive Youth Development,” “Measures,” and “Psychometric Properties” (Table 1.4). To perform the pilot searches, domains will be searched separately. Next, a thorough search will be run combining all domains to make sure a suitable search strategy is adjusted. PubMed, Scopus, Web of Science, and PsycINFO databases will be searched from the inception (2005) to 12<sup>th</sup> March 2024. Grey literature will be searched via the European Association for Grey Literature Exploitation (EAGLE). To cover more studies, we will search related scientific journals, as well as references of references. Experts in this field will be contacted to access under-print literature.



## **Search Strategy**

To develop the search strategy, a senior librarian from the Faculty of Medicine at Iran University of Medical Sciences will aid. All studies that have evaluated psychometric properties of the Five Cs Model of PYD scales in youths ages 10–29 will be included in this search strategy (see Table 1.4 for the sample search strategy). The search strategy will be adjusted based on a second version of the COSMIN initiative's search filter (Terwee et al., 2009).

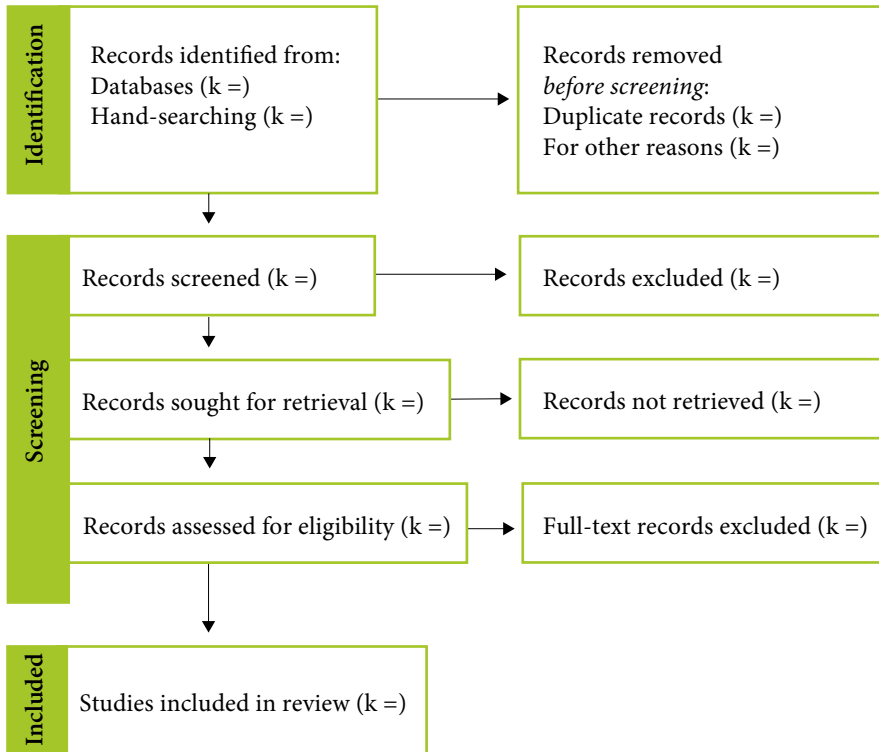
## **Data Screening**

The Rayyan QCRI online software (Ouzzani et al., 2016) will be employed to manage references, review titles and abstracts, and detect any duplicate entries. Two independent researchers will screen the literature (MHA and PSY). If their potential disagreements are not resolved through discussion, they will refer to a third researcher (SAA). In the screening step, the articles' titles and abstracts will be reviewed and studies not compatible with our study objectives will be excluded. The eligibility process involves reviewing the full texts of the articles and excluding the ones that fail to meet our eligibility criteria. Articles that meet our inclusion criteria will be included in the analysis. A template of the search and screening process has been visually depicted in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart (Figure 1.1).

## **Data Collection Process**

Prior to data charting and extraction, a data extraction form will be created (see Table 1.5 for a sample form). Subsequently, sample data extraction of five studies will be done to identify and fix the potential defects and shortcomings of the form. Data from the selected articles will be extracted separately by three researchers (PSY, RZ, and NH), and three senior researchers (MHA, SAA, and ETC) will be consulted in case of any ambiguities. For studies with missing data, we will request data from the corresponding author by contacting them via email.

*Care Chronicles*, 5(19), 34–36. <https://doi.org/10.3348/kjr.2004.5.1.11>



**Figure 1.1** Search and screening process.

## Data Extraction

The extracted data in the form will include: author(s), published year, country, language, sample size, number, age, and gender of participants, country group by income level (based on new World Bank country classifications by income level: 1- Low income, 2- Lower-middle income, 3- Upper-middle income, and 4- High income), family income level (1- Low income, 2- Lower-middle income, 3- Upper-middle income, and 4- High income), minority/majority, ethno-cultural group, immigration status, study setting (family/home, community, school, etc.), number of domains/subscales (excluding demographic information), number of questions, internal consistency, reliability, measurement error, face validity, content validity, structural validity, construct validity, cross-cultural validity, known-group validity, criterion validity, longitudinal validity, interpretability, and responsiveness.

**Table 1.5** *Data Extraction Form*

Author (s)	Published year	Country	Language	Participants (N/ age/ gender)	Country group by income level	Family income level	Minority/ majority	Ethno-cultural group	Immigration status	Study setting	Number of subscales	Number of questions	Reliability
1													
2													
3													
4													
5													

Internal consistency	Measurement error	Validity	Content validity	Face validity	Construct validity	Structural validity	Cross-cultural validity	Criterion validity	Known-group validity	Responsiveness	Longitudinal validity	Interpretability
1												
2												
3												
4												
5												

Notes. N = number.

## **Outcomes**

The most important outcomes of this study will include: 1) a comprehensive and clear picture of the psychometric properties (as listed in Tables 2 & 3) of the Five Cs Model of PYD scales, and 2) identification and comparison of the weaknesses and challenges of the Five Cs Model of PYD scales to inform on the most valid and reliable tools for measuring the Five Cs model of PYD under different circumstances.

## **Risk of Bias in Individual Studies**

The COSMIN tool (Mokkink, de Vet, et al., 2018) will be applied to measure risk of bias in each study. This tool evaluates the quality of study methodology and was specifically designed for studies on the psychometric properties of measurement tools. The 116-item COSMIN's Risk of Bias checklist consists of ten criteria (as listed in Table 1.2), including nine measurement properties along with detailed standards for questionnaire development (Prinsen et al., 2018) as follows: (1) validity (content, face, construct, structural, cross-cultural, criterion, known-group, and longitudinal validity) (Evans, 1996; Kerlinger, 1966; Mokkink et al., 2010; Prinsen et al., 2018; Teresi et al., 2009; Terwee et al., 2007), (2) reliability (measurement error, internal consistency, and reliability-stability) (de Vet et al., 2006; Mokkink et al., 2010; Rosner, 2015; Tavakol & Dennick, 2011), (3) responsiveness (Cohen, 2013; Guyatt et al., 1987; Yang & Berdine, 2017), and (4) interpretability (Prinsen et al., 2018).

## **Meta-analysis of Psychometric Properties**

Data pooling might not be doable in all meta-analyses; however, it is not impossible. Thus, if applicable, it will be included in our analysis ahead of data synthesis. Since there is controversy on preferred strategies and substantial discrepancies have been found when conducting two strategies for majority of cases (Borenstein et al., 2021), the conventional psychometric meta-analysis by Hunter and Schmidt (2004; 2014) will be applied. Their psychometric-principle-based approach argues that the large amount of observed variability in the association between two variables in primary evidence stems from the sources of artifact variability, such as measurement error (due to unreliable instruments), sampling error, and range restriction. Hence, meta-analyses need to calculate the potential



moderators in the relationships and rectify artifact variability across studies. To do so, investigators should adjust the artifact variability across studies through procedures such as deducting it from the total variance observed or study design.

Fisher's  $Z$  is opted for as the standardized effect size with a range from  $-\infty$  to  $+\infty$  and interpretation criteria resemble those for a correlation coefficient. In case intraclass, Pearson, or Spearman correlations are reported, we will utilize the Fisher's variance stabilizing transformation (Rosenberg et al., 2000; Rosenthal et al., 1994) to convert them to Fisher's  $Z$ . When unstandardized beta coefficients and  $F$ -ratios are reported, we will first convert them to  $r$  and afterwards to Fisher's  $Z$  (Rosenberg et al., 2000; Rosenthal et al., 1994). If no other statistics are reported but  $p$  values, we will first convert it to  $Z$ -score, after that to  $r$ , and finally to Fisher's  $Z$  (Rosenberg et al., 2000). If studies involve follow-up assessments, we will extract a) the total effect size for each psychometric property, and b) the effect sizes for each follow-up interval.

Comprehensive Meta Analysis V. 3.0 software (Borenstein, 2022) will be used to perform analyses. Random-effects models will be used with the assumption that the mean of effect size is unstable across studies and consequently, heterogeneity is likely. Heterogeneity will be assessed through the appraisal of the Cochran's  $Q$  test (heterogeneity is present if significant) and the  $I^2$  statistic (heterogeneity-related variability in effect estimates) (Higgins et al., 2003).  $I^2$  statistic is interpreted as not important (0–40%), moderate (30–60%), substantial (50–90%), and considerable (75–100%) (Higgins et al., 2019). In the presence of considerable heterogeneity, we will conduct sensitivity analysis to make sure that the results of meta-analysis are robust. In addition, funnel plots will be employed to discover the small study effects and reporting bias, where applicable (Higgins et al., 2019).

When pooling the data is not doable, we will make a qualitative abstract based on the studies' findings on measurement properties separately for each of the Five Cs Model of PYD scales. In case of inconsistencies across the results of different studies, we will provide potential justifications, or will sum up the results per subgroup of consistent results, if any pattern emerges. In the absence of any rational justification or clear pattern of findings, results will be rated based on the majority of findings.

### **Assessing Quality of Measurement Properties**

The criteria of COSMIN for good measurement properties will be used to measure the quality of each study's result on measurement properties of the Five

Cs Model of PYD scales, where each property will be rated on a scale of “*insufficient*, (-)” “*indeterminate*, (?)” or “*sufficient* (+)” (as defined in Table 1.3). Based on the COSMIN criteria, the lowest score of a property will be considered as its total quality. The total quality will be interpreted as high quality (50%), moderate quality (30–50%), and low quality (less than 30%) (Mokkink, de Vet, et al., 2018; Prinsen et al., 2018; Terwee et al., 2018). According to the taxonomy, there are four quality domains encompassed by psychometric properties : reliability, validity, responsiveness, and interpretability. As all articles might not assess all measurement properties, only those properties that were evaluated in each study will be completed. Where possible, all quality domains will be evaluated separately in the present systematic review and meta-analysis. Owing to the lack of criteria for rating EFA in COSMIN, our decision will be based on the Terwee et al.’s (2007) suggestion: EFA is proper if no theoretical or empirically emerged structural model is proposed in the literature. In case a theoretical hypothesis is present, or the empirical evidence has documented a stablished factorial model, confirmatory factor analysis should be conducted (Bollen, 1989; Vet et al., 2005). To rate the quality of EFA, we will use the following criteria: (+) = the selected factor can explain 50% or higher of variance OR elected factors explained less than 50% of variance but authors could justify their choice, (?) = unable to assign a score due to unclear or missing information, such as the absence of explained variance, and (-) = Criteria for “+” were not met (Terwee et al., 2007).

### **Confidence in Cumulative Evidence**

After summarizing the pieces of evidence utilized for general ratings on each one of psychometric properties, the trustworthiness of the summarized evidence will be rated on a scale from “*very low*” (e.g., we are skeptical about the measurement property’s estimate: it is highly likely that the actual measurement property is considerably different from the measurement property’s estimate) to “*high*” (e.g., we are assured that the estimate of the measurement property is quite similar to the real measurement property), according to the modified Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach (Balshem et al., 2011; Prinsen et al., 2018). Two independent reviewers (MHA and ETC) will assess the general quality of summarized results. A senior reviewer (SAA) will be consulted for discrepancies, and consensus will be reached.

## Recommendations for Selecting an Instrument

The evaluation of psychometric properties of the Five Cs Model of PYD scales, and recommendations for future applications, will be conducted according to the combination of the scores of general ratings for each one of the psychometric properties and the results of grading (Prinsen et al., 2018). The results of all the studies included for each measure will be classified into three recommendation categories (Mokkink, de Vet, et al., 2018; Mokkink, Prinsen, et al., 2018): (a) the Five Cs Model of PYD scales can be potentially introduced as the most proper instrument for the Five Cs Model of PYD in adolescents; (b) the Five Cs Model of PYD scales can be potentially recommended, though more studies on its validation are required; and (c) the Five Cs Model of PYD scales is not to be recommended. The reason for which each of the Five Cs Model of PYD scales is labeled as one of the above categories will be provided. Future directions for research will be presented, where applicable.

## Discussion

The primary objective of the proposed systematic review is to offer a summary statement on the suitability of using the Five Cs Model of PYD scales among youth aged 10 to 29. We seek to design a review protocol to critically evaluate, summarize, and synthesize research on the measurement properties (validity, reliability, responsiveness, and interpretability, as well as the criteria of good measurement properties) of the Five Cs Model of PYD scales.

The existing systematic reviews are concentrated mostly on the outcomes of PYD programs (e.g., Catalano et al., 2019; Shek & Yu, 2011; Tidmarsh et al., 2022), whereas no systematic review and meta-analysis was conducted to appraise the quality of the Five Cs Model of PYD scales. The contribution of this proposed systematic review to the existing literature is to primarily provide a comprehensive and clear picture of the psychometric properties of the Five Cs Model of PYD scales. Subsequently, we aim to identify and compare their weaknesses and challenges in terms of validity and reliability, and finally, to introduce the most valid, reliable, and appropriate tools for measuring the Five Cs Model of PYD. In light of the critical and rigorous quality assessments on these scales, the transparency, clarity, and reproducibility of the previous findings will be identified (Shamseer et al., 2015). Therefore, as

an example, we will be able to understand the extent to which the results of the previous studies were influenced by the risk of bias. Furthermore, such evidence-based judgments on their potential weaknesses may highlight the need for the improvement of the existing measures.

The present protocol has two major limitations. First, the heterogeneity may be high because of the dissimilarity of study designs and samples (e.g., general population, at-risk adolescents, disabled youth, etc.) and high level of heterogeneity may prevent us from performing data synthesis and meta-analysis. Second, the planned search process will only include studies in the English language. The exclusion of non-English studies may limit evaluation of tools' interpretability (Boström & Broberg, 2018). Additionally, it means that a significant body of evidence in other languages will be missed for assessing cross-cultural and cross-national validity/measurement invariance of the tools (e.g., Ghasemi et al., 2019).

### **Ethics and Dissemination**

Since this project does not include data collection directly from the participants, ethical approval is not required. The results of the planned review will be published in a peer-reviewed journal.

## References

- Babae, J., Najafi, M., & Rezaei, A. (2018). The psychometric properties of positive youth development scale in students. *Journal of Psychological Science, 16*(64), 540–553. <https://doi.org/20.1001.1.17357462.1396.16.64.6.9>
- Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., Vist, G. E., Falck-Ytter, Y., Meerpohl, J., & Norris, S. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology, 64*(4), 401–406. <https://doi.org/10.1016/j.jclinepi.2010.07.015>
- Borenstein, M. (2022). Comprehensive meta-analysis software. *Systematic Reviews in Health Research: Meta-Analysis in Context, 535–548*.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to Meta-analysis*. John Wiley & Sons. <https://doi.org/10.1002/9780470743386>
- Boström, P., & Broberg, M. (2018). Protection and restriction: A mixed-methods study of self-reported well-being among youth with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 31*(1), e164–e176. <https://doi.org/10.1111/jar.12364>
- Bowers, E. P., Li, Y., Kiely, M. K., Brittan, A., Lerner, J. V., & Lerner, R. M. (2010). The five Cs model of positive youth development: A longitudinal analysis of confirmatory factor structure and measurement invariance. *Journal of Youth and Adolescence, 39*(7), 720–735.
- Buenconsejo, J. U., Datu, J. A. D., Chiu, M. M., & Chan, R. C. H. (2022). Psychometric validity and measurement invariance of positive youth development in the Philippines during the COVID-19 pandemic. *Applied Developmental Science, 1–16*. <https://doi.org/10.1080/10888691.2022.2078719>
- Catalano, R. F., Berglund, M. L., Ryan, J. A., Lonczak, H. S., & Hawkins, J. D. (2002). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Prevention & Treatment, 5*(1), 15a. <https://doi.org/10.1037/1522-3736.5.1.151a>
- Catalano, R. F., Skinner, M. L., Alvarado, G., Kapungu, C., Reavley, N., Patton, G. C., Jessee, C., Plaut, D., Moss, C., & Bennett, K. (2019). Positive youth development programs in low-and middle-income countries: A conceptual framework and systematic review of efficacy. *Journal of Adolescent Health, 65*(1), 15–31. <https://doi.org/10.1016/j.jadohealth.2019.01.024>
- Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Conway, R. J., Heary, C., & Hogan, M. J. (2015). An evaluation of the measurement properties of the five Cs model of positive youth development. *Frontiers in Psychology, 6*, 1941. <https://doi.org/10.3389/fpsyg.2015.01941>
- Crocetti, E., Erentaitė, R., & Žukauskienė, R. (2014). Identity styles, positive youth development, and civic engagement in adolescence. *Journal of Youth and Adolescence, 43*(11), 1818–1828. <https://doi.org/10.1007/s10964-014-0100-4>
- Damon, W. (2004). What is positive youth development? *The ANNALS of the American Academy of Political and Social Science, 591*(1), 13–24. <https://doi.org/10.1177/0002716203260092>
- De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. (2011). Systematic reviews of measurement properties. In T. De Vet, L. Mokkink, & D. Knol (Eds.), *Measurement*

- in *Medicine: A Practical Guide* (pp. 275–314). Cambridge University Press. <https://doi.org/10.1080/010543406.2013.737220>
- De Vet, H. C., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, *59*(10), 1033–1039. <https://doi.org/10.1016/j.jclinepi.2005.10.015>
- Dvorsky, M. R., Kofler, M. J., Burns, G. L., Luebke, A. M., Garner, A. A., Jarrett, M. A., Soto, E. F., & Becker, S. P. (2018). Factor Structure and Criterion Validity of the Five Cs Model of Positive Youth Development in a Multi-University Sample of College Students. *Journal of Youth and Adolescence*, *48*(3), 537–553. <https://doi.org/10.1007/s10964-018-0938-y>
- Erikson, E. H. (1968). *Identity: Youth and crisis* (No. 7). WW Norton & company.
- Evans, J. D. (1996). *Straightforward Statistics for the Behavioral Sciences*. Thomson Brooks/Cole Publishing Co.
- Freud, A. (1969). Adolescence as a developmental disturbance. *Adolescence*, *20*(2), 5.
- Gaspar de Matos, M., Santos, T., Reis, M., Marques, A., Tomé, G., Simões, C., & Wiium, N. (2018). Portuguese validation of the positive youth development short form (PYD-SF): A shortened version. *Psicologia, Saúde & Doenças*, *19*(3), 477–489. <https://doi.org/10.15309/18psd190301>
- 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 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., Callina, K. S., & 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>
- Geldhof, G. J., Bowers, E. P., Mueller, M. K., Napolitano, C. M., Callina, K. S., Walsh, K. J., Lerner, J. V., & Lerner, R. M. (2015). The five Cs model of positive youth development. In *Promoting Positive Youth Development* (pp. 161–186). Springer. [https://doi.org/10.1007/978-3-319-17166-1\\_9](https://doi.org/10.1007/978-3-319-17166-1_9)
- Gestsdottir, S., Bowers, E., von Eye, A., Napolitano, C. M., & Lerner, R. M. (2010). Intentional self-regulation in middle adolescence: The emerging role of loss-based selection in positive youth development. *Journal of Youth and Adolescence*, *39*(7), 764–782. <https://doi.org/10.1007/s10964-010-9537-2>
- Gestsdottir, S., Geldhof, G. J., Lerner, J. V., & Lerner, R. M. (2017). What drives positive youth development? Assessing intentional self-regulation as a central adolescent asset. *International Journal of Developmental Science*, *11*(3–4), 69–79. <https://doi.org/10.3233/DEV-160207>
- Ghasemi, M., Farahani, M. T., & Abdolahi, M. H. (2019). Functional structure and presentation model for the positive youth development in Iranian adolescents. *Journal of Psychology*, *90*(2), 114–133.
- Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, *40*(2), 171–178. [https://doi.org/10.1016/0021-9681\(87\)90069-5](https://doi.org/10.1016/0021-9681(87)90069-5)
- Hall, G. S. (1905). *Adolescence: Its Psychology and its Relations to Physiology, Anthropology, Sociology, Sex, Crime, Religion and Education* (Vol. 2). D. Appleton.

- Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons.
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Holsen, I., Geldhof, J., Larsen, T., & Aardal, E. (2017). The five Cs of positive youth development in Norway: Assessment and associations with positive and negative outcomes. *International Journal of Behavioral Development*, *41*(5), 559–569.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.
- Kerlinger, F. N. (1966). *Foundations of Behavioral Research*. Holt, Rinehart and Winston.
- Klein, J. D., Sabaratnam, P., Auerbach, M. M., Smith, S. M., Kodjo, C., Lewis, K., Ryan, S., & Dandino, C. (2006). Development and factor structure of a brief instrument to assess the impact of community programs on positive youth development: The Rochester Evaluation of Asset Development for Youth (READY) tool. *Journal of Adolescent Health*, *39*(2), 252–260. <https://doi.org/10.1016/j.jadohealth.2005.12.004>
- Larsson, B. W., & Larsson, G. (2002). Development of a short form of the Quality from the Patient's Perspective (QPP) questionnaire. *Journal of Clinical Nursing*, *11*(5), 681–687.
- Lerner, R. M., & Lerner, J. V. (2013). The positive development of youth: Comprehensive findings from the 4-H study of positive youth development. *Washington, DC: National, 4*.
- Lerner, R. M., Lerner, J. V., Almerigi, J. B., Theokas, C., Phelps, E., Gestsdottir, S., Naudeau, S., Jelicic, H., Alberts, A., & Ma, L. (2005). Positive youth development, participation in community youth development programs, and community contributions of fifth-grade adolescents: Findings from the first wave of the 4-h study of positive youth development. *Journal of Early Adolescence*, *25*(1), 17–71.
- Lerner, R. M., Lerner, J. V., Lewin-Bizan, S., Bowers, E. P., Boyd, M. J., Mueller, M. K., Schmid, K. L., & Napolitano, C. M. (2011). Positive youth development: Processes, programs, and problematics. *Journal of Youth Development*, *6*(3), 38–62. <https://doi.org/10.5195/jyd.2011.174>
- Lopez, A., Yoder, J. R., Brisson, D., Lechuga-Pena, S., & Jenson, J. M. (2015). Development and validation of a positive youth development measure: The Bridge-Positive Youth Development. *Research on Social Work Practice*, *25*(6), 726–736. <https://doi.org/10.1177/10497315145348>
- Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1171–1179. <https://doi.org/10.1007/s11136-017-1765-4>
- Mokkink, L. B., Prinsen, C., Patrick, D. L., Alonso, J., Bouter, L., de Vet, H. C., Terwee, C. B., & Mokkink, L. (2018). COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). *User Manual*, *78*(1). <https://doi.org/10.1007/s11136-018-1798-3>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, *63*(7), 737–745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>

- Moore, K. A., Lippman, L., & Brown, B. (2004). Indicators of child well-being: The promise for positive youth development. *The Annals of the American Academy of Political and Social Science*, 591(1), 125–145. <https://doi.org/10.1177/000271620326010>
- Morris, S. L., Wagner, E. F., & Wales, E. (2018). Music education as a path to positive youth development: An El sistema-inspired program. *Journal of Youth Development*, 13(4), 149–163. <https://doi.org/10.5195/jyd.2018.572>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan – a web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 1–10. <https://doi.org/10.1186/s13643-016-0384-4>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews*, 10(1), 1–11. <https://doi.org/10.1136/bmj.n71>
- Phelps, E., Zimmerman, S., Warren, A., Jeličić, H., von Eye, A., & Lerner, R. M. (2009). The structure and developmental course of positive youth development (PYD) in early adolescence: Implications for theory and practice. *Journal of Applied Developmental Psychology*, 30(5), 571–584. <https://doi.org/10.1016/j.appdev.2009.06.003>
- Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Riggs, N. R., Bohnert, A. M., Guzman, M. D., & Davidson, D. (2010). Examining the potential of community-based after-school programs for Latino youth. *American Journal of Community Psychology*, 45(3), 417–429. <https://doi.org/10.1007/s10464-010-9313-1>
- Rosenberg, M. J., Rosenberg, M. S., Adams, D. C., & Gurevitch, J. (2000). *MetaWin 2.0 (User's Manual): Statistical Software for Meta-Analysis*. United Kingdom: Sinauer Associates, Incorporated.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The Handbook of Research Synthesis* (pp. 231–244). New York: Russell Sage Foundation.
- Rosner, B. (2015). *Fundamentals of Biostatistics*. Cengage learning.
- Roth, J. L., & Brooks-Gunn, J. (2003). Youth development programs: Risk, prevention and policy. *Journal of Adolescent Health*, 32(3), 170–182. [https://doi.org/10.1016/S1054-139X\(02\)00421-4](https://doi.org/10.1016/S1054-139X(02)00421-4)
- Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage publications.
- Schulman, S., & Davies, T. (2007). Evidence of the impact of the “youth development” model on outcomes for young people—A literature review. National Youth Agency, Leicester.
- Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *Bmj*, 349. <https://doi.org/10.1136/bmj.g7647>
- Shek, D. T., & Chai, W. (2020). The impact of positive youth development attributes and life satisfaction on academic well-being: A longitudinal mediation study. *Frontiers in Psychology*, 11, 2126. <https://doi.org/10.3389/fpsyg.2020.02126>



- Shek, D. T., & Yu, L. (2011). A review of validated youth prevention and positive youth development programs in Asia. *International Journal of Adolescent Medicine and Health*, 23(4), 317–324. <https://doi.org/10.1515/IJAMH.2011.028>
- Sieng, M., Cloutier, S., & Irinata, K. (2018). Positive youth development sustainability scale (pydss): The development of an assessment tool. *Journal of Social Change*, 10(1), 7. <https://doi.org/10.5590/JOSC.2018.10.1.07>
- Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 300–308. <https://doi.org/10.1037/1040-3590.7.3.300>
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12(1), 102–111. <https://doi.org/10.1037/1040-3590.12.1.102>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Teresi, J. A., Ocepek-Welikson, K., Kleinman, M., Eimicke, J. P., Crane, P. K., Jones, R. N., Lai, J.-s., Choi, S. W., Hays, R. D., & Reeve, B. B. (2009). Analysis of differential item functioning in the depression item bank from the Patient Reported Outcome Measurement Information System (PROMIS): An item response theory approach. *Psychology Science*, 51(2), 148–180.
- Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
- Terwee, C. B., Jansma, E. P., Riphagen, I. I., & de Vet, H. C. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*, 18(8), 1115–1123. <https://doi.org/10.1007/s11136-009-9528-5>
- Terwee, C. B., Prinsen, C., Ricci Garotti, M., Suman, A., De Vet, H., & Mokkink, L. B. (2016). The quality of systematic reviews of health-related outcome measurement instruments. *Quality of Life Research*, 25(4), 767–779. <https://doi.org/10.1007/s11136-015-1122-4>
- Terwee, C. B., Prinsen, C. A., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., De Vet, H. C., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research*, 27(5), 1159–1170. <https://doi.org/10.1007/s11136-018-1829-0>
- Tidmarsh, G., Thompson, J. L., Quinton, M. L., & Cumming, J. (2022). Process evaluations of positive youth development programmes for disadvantaged young people: A systematic review. *Journal of Youth Development*, 17(2), 106–140. <https://doi.org/10.5195/jyd.2022.1156>
- Tomé, G., Gaspar de Matos, M., Camacho, I., Gomes, P., Reis, M., Branquinho, C., Gomez-Baya, D., & Wiium, N. (2019). Positive youth development (pyd-sf): validação para os adolescentes portugueses. *Psicologia, Saúde and Doenças*, 20(3), 556–568. <https://doi.org/10.15309/19psd200301>
- Truskauskaitė-Kunevičienė, I., Romera, E., Ortega-Ruiz, R., & Žukauskienė, R. (2020). Promoting positive youth development through a school-based intervention program try volunteering. *Current Psychology*, 39, 705–719. <https://doi.org/10.1007/s12144-018-9790-1>

- Wong, P. W. C., Kwok, K. W., & Chow, S. L. (2022). Validation of positive youth development scale and implications for adolescent in Hong Kong Community. *Child & Youth Care Forum*, 51(5), 901–919. <https://doi.org/10.1007/s10566-021-09658-6>
- Yang, S., & Berdine, G. (2017). The receiver operating characteristic (ROC) curve. *The Southwest Respiratory and Critical Care Chronicles*, 5(19), 34-36. <https://doi.org/https://doi.org/10.3348/kjr.2004.5.1.11>