JOURNAL OF
Research on Adolescence

Society for Research on Adolescence
www.s-r-a.org

JOURNAL OF RESEARCH ON ADOLESCENCE, 30(4), 896–912

Culturally Relevant Resilience: A Psychometric Meta-Analysis of the Child and Youth Resilience Measure (CYRM)

Rachel L. Renbarger
Duke University

R. Noah Padgett
Baylor University

Richard G. Cowden
Harvard University

Kaymarlin Govender
University of KwaZulu-Natal

Münübe Z. Yilmaz
Istanbul Medeniyet University

Lakia M. Scott
Baylor University

Alexander V. Makhnach
Institute of Psychology of the Russian Academy of Sciences

Jan Sebastian Novotny
St. Anne's University Hospital Brno

Geoff Nugent
Nugent Family Counseling Center, Inc

Lacey Rosenbaum
National Council for Behavioral Health

Lucie Křeménková
Palacký University

Measuring key components of resilience is vital for understanding cross-cultural dynamics among youth and the environment. The Child and Youth Resilience Measure (CYRM-28) was developed as a cross-cultural measure of resilience and has been used globally. To examine the cross-cultural utility of the CYRM-28, we conducted a systematic review of the literature reporting on the psychometric properties of the measure. Using data representing six countries (N = 6,232) that were supplied from authors of the studies reviewed, a multilevel confirmatory factor analysis was also conducted to estimate the variability of the measurement properties among communities, ages, and sex. Results indicate that the literature generally did not include reliability and validity information for the instrument. From the multilevel confirmatory factor analysis, the measure was invariant between adolescent age-groups and sexes but not across communities.

INTRODUCTION

Research on resilience has early roots in efforts aimed at preventing the development of psychopathology among children (see Garmezy, 1981). The historical emphasis on understanding the processes involved in resistance to stress among children is not surprising given the biopsychosocial changes that occur during the first two decades of life. These changes are particularly prominent during adolescence, as young people navigate critical developmental tasks and milestones during their transition to adulthood (Linders, 2017). Children and adolescents also have heightened vulnerability to the effects of adverse environmental conditions and changes (Park & Schepp, 2015). Therefore, resilience in young people warrants special consideration, given the short- and long-term consequences that may develop out of maladaptive adjustment to adversity during this developmental period (Doyle & Cicchetti, 2017; McDougall & Vaillancourt, 2015). Although considerable strides have been made toward understanding resilience and its measurement in young people, research that comprehensively examines the psychometric properties of available instruments that purport to measure resilience is necessary (Ahern, Kiehl, Sole, & Byers, 2006). One measure, in particular, the Child and Youth Resilience Measure with 28 items (CYRM-28), is a promising tool used by clinicians and researchers to assess and compare resilience across

Requests for reprints should be sent to Rachel L. Renbarger, Duke University, 300 Fuller St, Durham, NC 27701. E-mail: rachelrenbarger@gmail.com

DOI: 10.1111/jora.12569

cultures and countries (Ungar & Liebenberg, 2009; Ungar & Liebenberg, 2011). A number of independent studies have examined the psychometrics of this measure, but a comprehensive assessment of the instrument does not exist. In the current study, a meta-analytic approach is used to examine the psychometric properties and cultural relevance of the CYRM-28.

Resilience: A Brief Overview

Scholars have long been interested in understanding the mechanisms that facilitate positive adaptation in the face of significant adversity. Resilience has emerged as a concept encapsulating such mechanisms, referring broadly to a dynamic process in which an individual emerges relatively unaffected despite exposure to ordinarily harmful circumstances (Infurna & Luthar, 2016; Luthar, Cicchetti, & Becker, 2000). Historically, research on resilience has been rooted in two overarching perspectives that differ on the property of the concept. Trait approaches assert resilience as a pattern of adaptation to adversity that is relatively stable across situations and time (Bonanno et al., 2011), whereas state perspectives emphasize the intraindividual variability of resilience (Masten, 2014a). These distinctions are often closely tied to variations in operationalizing the two core conditions on which resilience is contingent: adversity and adaptation. Descriptions of adversity tend to differ in the choice of terminology used to describe adversity (e.g., stressor and risk), along with distinctions based on type (i.e., positive versus negative), severity (i.e., acute versus chronic), and degree of specificity (i.e., broad versus narrow in scope). Diversity also exists in the definitions used to describe adaptation, ranging from the absence of negative outcomes (e.g., psychopathology) to the achievement of relevant developmental milestones or markers of performance (in a single or multiple domains of functioning) typically expected of a person (Masten, 2014b). Taken together, conceptual and operational discrepancies suggest there is considerable variability in both the contextual applicability and outcomes of resilience.

Developmental and Contextual Considerations

Successful adaptation in response to adversity is facilitated by the dynamic interaction between protective mechanisms (i.e., resources) available to a person (Ungar, 2011). These resources exist at multiple levels (e.g., psychological, family, and social) and vary in degree of proximity to the individual. Availability and interplay between such resources contribute to both between- and within-person variations in the capacity to navigate adversity and maintain adaptive functioning.

There are also distinctions in the salience of protective resources as young people transition from childhood to adulthood. A fundamental task of adolescence is establishing a distinct, individuated sense of self (Steinberg, 2014), a process that prompts reorganization of caregiver–child relationships. As adolescents acquire the capacity for autonomous selfhood and self-reliance, they become less dependent on primary childhood attachment figures (i.e., caregivers) for emotional and behavioral support (Ruhl, Dolan, & Buhrmester, 2014). Progression through adolescence is also associated with greater emphasis on forming and strengthening a broader range of social bonds with extrafamilial attachment figures, particularly peers (Blakemore & Mills, 2014). Some evidence suggests that peer support has a stronger effect on resilience in adolescence and early adulthood compared to family support (van Harmelen et al., 2017; van Hoorn et al., 2014). By mid-to-late adolescence, romantic relationships feature more prominently in the lives of young people (Furman & Winkles, 2012), ultimately reconfiguring adolescents' social landscape in ways that endure into adulthood (Madsen & Collins, 2011). These changes tend to differ based on sex; past research has found that males and females differ in terms of levels and types of protective resources throughout their adolescence (Hartman, Turner, Daigle, Exum, & Cullen, 2008; Sun & Stewart, 2007). These kinds of shifts in psychosocial processes dynamically reconstruct adolescents' self-position, the affiliations they have with others, and the relative priority of their relational bonds, thereby shaping the type of resilience resources that are available and drawn upon by young people when faced with adversity.

Additional variation may exist between and within communities or cultures (Wexler, DiFluvio, & Burke, 2009). For example, the celebration of finishing high school may differ between groups who have been traditionally marginalized in education (Wexler et al., 2009) or the promotion of individualism and independence in the United States differs from collectivistic societies. As such, groups may identify cultural-specific resources that sustain well-being rather than adhering to the settings of the researchers (Southwick et al., 2014). Despite the need for culturally relevant measures, most measures of resilience fail to take culture into account (Clauss-Ehlers, 2008).

Child and Youth Resilience Measure

Development. The CYRM was developed as a culturally sensitive resilience measure that aligns with social–ecological systems theory. Using a mixed-methods approach, the developers sampled from 11 countries: Canada, China, Colombia, the Gambia, India, Israel, Palestine, Russia, South Africa, Tanzania, and the United States (Ungar & Liebenberg, 2009). The authors utilized a qualitative approach to find common conceptions of resilience across the globe and identify the resources that allowed for individuals to succeed despite obstacles (Ungar & Liebenberg, 2011; Ungar & Liebenberg, 2013). Subsequently, 58 items were developed relating to four categories of resilience: individual, relationships, community, and culture. Data from pilot study groups were used to conduct quantitative analyses. Exploratory factor analysis and expert consensus resulted in a measure comprising of 28 items distributed across three categories of resilience at the individual, caregiver, and contextual levels.

Psychometric Evidence. Since the initial validation of the CYRM-28, several studies have worked to provide additional evidence for drawing valid inferences from the scores obtained on the measure. Sanders, Munford, Thimasarn-Anwar, and Liebenberg (2017) conducted a large study assessing the factor structure of the CYRM-28 in youth from New Zealand. They found evidence in support of an alternative, four-factor structure to the three-factor structure reported in Liebenberg et al. (2012). Other studies (i.e., van Rensburg, Theron, & Ungar, 2017; Zand, Liebenberg, & Shamloo, 2017) have also found discrepancies in the structure of the CYRM-28. Although the CYRM-28 was developed based on research from countries on all inhabited continents for use across cultures, reported variations in the structure and composition of the scale suggest there may be challenges to making inferences about the cross-cultural relevance and stability of resilience resources.

The Present Study

Little research has reviewed psychometric properties of resilience scales within this cultural framework. In Windle, Bennett, and Noyes' (2011) qualitative review of existing resilience instruments, the authors found the CYRM-28 was the only instrument to mention the cultural and contextual nature of resilience. However, subsequent studies have overlooked variations in the measurement of resilience across sexes, ages, and cultures (e.g., Smith-Osborne & Whitehill Bolton, 2013; Pangallo, Zibarras, Lewis, & Flaxman, 2015). Given the relatively recent development of the CYRM and the dearth of resilience measures that are sensitive to culture, further evidence is necessary for drawing valid and reliable inferences about resilience from scores on the CYRM-28. Thus, the purpose of this study is to examine the collective evidence of the CYRM and evaluate the psychometric utility of the measure across several countries as well as between sexes and adolescent age-groups.

METHOD

Search Procedure

To find all articles that included data on the psychometric properties of the CYRM-28, we utilized a systematic review procedure. Keywords for the search included "Child and Youth Resilience Measure" and "CYRM" within the academic databases of Academic Search Complete, E-Journals, ERIC, PsycARTICLES, Psychology and Behavioral Sciences Collection, PsycINFO, and PsycTESTS. We also examined the official CYRM website as it contained some studies that had used the instrument. For all of the search results, we screened the abstract and title to determine eligibility. Articles were excluded if they (1) did not use the CYRM-28, (2) did not have CYRM information in English, or (3) did not have enough quantitative information for the analyses in this study.

Data Collection

Extraction. Information obtained from the retained articles included publication (title and authors) and sample information (percent female, sample size, and mean age), characteristics of the CYRM, and analysis information. Characteristics of the CYRM included the number of items, number of factors, which items loaded onto each factor, and reliability evidence. Analysis information included output from the exploratory and confirmatory factor analyses and correlations between the CYRM and other measures. The first two authors extracted the information, and data issues were discussed on a case-by-case basis.

Data received. Of the articles reviewed ($n = 50$), the contact information for corresponding authors of 27 articles was available on the manuscript or could be identified through an online search. Those authors were initially contacted in the spring of 2018, with relevant follow-up emails sent approximately 1 month later. More than half ($n = 15$) of the authors responded, and five were willing to share their deidentified, raw data. One portion of the data came from the first author on this manuscript. The raw data included age, sex, and responses to the 28 items of the CYRM. Data came from six countries: the Czech Republic, Costa Rica, Ghana, Russia, South Africa, and the United States. In some countries (i.e., Ghana, Russia, and the United States), data were collected from a single community. Data collection procedures for data that were supplied from other countries (i.e., Costa Rica, the Czech Republic, and South Africa) indicated that participants were sampled from multiple communities. For instance, Govender et al. (2017) sampled students from 12 schools located in distinct communities and at two different time points. When sufficient information was available to determine whether country-level samples were recruited from geographically unique locations, we disaggregated participants into community-level subsamples. Using this approach, we identified a total of 36 communities across the six countries (see Table 1). The full sample includes 6232 participants (female, 50.72%) ranging from 9 to 22 years old ($M_{\text{age}} = 16.51$, $SD_{\text{age}} = 2.39$). Given the focus of this paper was adolescents, we disaggregated the data into two age-groups—adolescents (10–18 years) and early emerging adults (19–22 years)—for invariance testing.

Analyses

In this study, we aggregated the estimates of reliability for the CYRM-28 based on reliability estimates reported from the studies retained. The aggregate estimate of reliability was calculated by using a sample size-adjusted average. We implemented these methods with the metaphor package (Viechtbauer, 2010) in R (R Core Team, 2018).

Evidence for convergent validity is typically reported as correlations among scores on the CYRM-28 and scores on other instruments conceptually related to resilience (e.g., psychological trauma and relationships with parents). Based on the reported correlations among scores on the CYRM-28 with other measures, we created conceptual groups to aggregate these reported correlations. We used the Hunter–Schmidt meta-analytic procedures to calculate the estimates of convergent validity for each group. The Hunter–-Schmidt method is advantageous because it pools the estimated correlations while accounting for sample size and measurement error (see Hunter & Schmidt, 2004; Kline, 2015, for technical details).

In addition to aggregating convergent validity evidence based on information available in each retained article, we used the data acquired by contacting authors to investigate measurement variability of resilience. Measurement investigations of resilience have utilized confirmatory factor analysis (CFA), an analytic technique for developing hypotheses about how items group together and the underlying processes that bring about the observed data. This process is accomplished by developing a theoretical model for why item responses tend to covary (i.e., relate to one another), which is why this type of analysis is sometimes called covariance structure analysis. One major assumption of this type analysis is that the observed cases (individual adolescent responses) are independent from the responses of others. This is more generally the assumption of independence of observations. However, in the study of resilience, seldom would such an assumption be tenable because data are collected across multiple communities that are expected to vary from one another. A more general method of investigating the measurement of resilience through a factor analytic framework is needed to account for the considerations for how data are gathered across communities and thus we conducted a multilevel confirmatory factor analysis (ML-CFA).

In ML-CFA, the covariance structure is decomposed into a level 1 and level 2 covariances. The level 1 model represents factors that influence individuals' responses or the factor(s) underlying individual differences in youth within a community. The level 2 model represents the factor(s) influencing between-community differences. By utilizing ML-CFA, we can explicitly model the mechanisms that influence resilience for an individual and what mechanisms influence differences in resilience between communities. A flexibility we are awarded based on the ML-CFA framework is that we can specify potentially unique influences at each level (see Muthén, 1994; Stapleton, Yang, & Hancock, 2016, for technical details).

The ML-CFA framework was chosen for this application to help measure the variability in resilience (as measured by the CYRM-28) across communities. This approach was selected over a

TABLE 1
Sample Characteristics for Multilevel Confirmatory Factor Analysis

Study	*Country*	N	*% Female*	*Mean Age (SD)*
Makhnach (2016)	Russia	188	46.81	16.56 (2.06)
Novotny (2016)	Czech Republic	45	37.78	16.33 (1.58)
		410	52.93	16.02 (5.89)
		499	53.31	16.70 (1.59)
Authors (in prep)	Costa Rica	23	43.48	14.39 (1.82)
		99	51.52	14.40 (0.95)
		16	75.00	22.19 (1.22)
		41	56.10	17.12 (1.49)
		134	49.25	14.26 (0.72)
Nugent (2014)	The United States	100	33.00	14.04 (1.41)
Rosenbaum (2017)	South Africa	7	42.86	17.86 (0.90)
Govender (2017)	South Africa	151	46.36	15.54 (2.01)
		37	72.97	15.02 (1.38)
		113	48.67	15.65 (1.72)
		150	39.33	15.85 (1.92)
		113	42.48	15.93 (1.97)
		260	48.85	14.91 (1.74)
		251	52.99	15.57 (1.94)
		237	53.59	14.67 (1.63)
		242	99.59	14.71 (1.42)
		419	55.85	14.55 (1.37)
		301	[a]	14.29 (1.21)
		287	54.70	14.42 (1.49)
		229	58.95	16.11 (1.68)
		213	100.00	15.85 (1.47)
		295	53.22	15.67 (1.42)
		237	[a]	15.51 (1.26)
		311	56.27	15.79 (1.51)
		118	46.61	16.87 (1.85)
		50	58.00	16.46 (1.37)
		81	45.68	16.40 (1.54)
		89	48.31	16.48 (1.71)
		99	39.39	16.62 (1.77)
		141	47.52	16.35 (1.69)
		150	54.00	15.70 (1.61)
Renbarger (in prep)	Ghana	96	60.42	12.60 (4.10)
Average		173.11	50.72	16.51 (2.39)
Total		6,232		

Note. [a]Sex not reported.

related method (i.e., multigroup CFA) that was unlikely to efficiently estimate between-community differences in measurement models for this study given the large number of communities in these data (36) and the high within-community variability in sample sizes (range = 7-499). ML-CFA was conducted to directly estimate the amount of heterogeneity that was present in the observed items and latent resilience factor(s). This was accomplished by estimating the level 2 variance in the resilience factor(s). For factors that were specified to be invariant across levels, we then estimated the intraclass correlations (ICC). The ICC is the proportion of variance of an outcome that is attributed to between-group differences. More specifically, $CC = \frac{\psi_B}{\psi_B + \psi_W}$, where ψ_B is the estimated variance of the outcome between groups or level 2 variance, while ψ_W is the estimated variance component within groups or level 1 variance (Heck & Thomas, 2015). This statistic was used to find evidence of the cross-cultural measurement variability of the items and the resilience factor(s) because the ICC describes how much groups are responding differently on average. The larger the ICC/level 2 variance is, the less evidence there is that the measure is operating consistently across cultures. Similarly, a low ICC/level 2 variance would provide evidence that the CYRM is providing consistent information across all communities.

Factor Structures

In this study, the factor structure of the CYRM-28 was investigated at the individual level and the community level. Five variations of published factor structures were investigated at both levels of the model. These five variations of the factor structure are (1) a single general resilience factor based on the original global sample, (2) the alternative Canada 1 model proposed by van Rensburg et al. (2017) which we simply called Canada, (3) the New Zealand model of van Rensburg et al. (2017) originally proposed by Sanders, Munford, Thimasarn-Anwar, Liebenberg, and Ungar (2015), (4) alternative New Zealand model of van Rensburg et al. (2017), and (5) a newly proposed bifactor model based on the alternative New Zealand model and a general resilience factor.

The last bifactor model is a conceptually logical model to test (at least on the level 1 model). The bifactor model proposed keeps the individual subscales of the CYRM-28 and adds this general resilience component which is conceptually distinct from other potential higher-order factor structures. Some higher-order factor structures of the CYRM-28 have been investigated in prior studies and were not retained (van Rensburg et al., 2017). The major conceptual difference between the bifactor model and a second-order factor is the distinction between each item's relationship with the global resilience factor. In the bifactor conceptualization, the general resilience factor directly influences item responses, whereas in a second-order factor, the items are only indirectly influenced by general resilience. This distinction allows for a deeper understanding of how participants respond to items about resilience. The specific breakdown of the mapping of items to factors is shown in Table 2.

The factor structures outlined in Table 2 were investigated for each level of modeling. Within the ML-CFA framework, we specified the factor structure of the level 1 and level 2 models separately which allowed for possibly unique factor structures at each level. To this end, we investigated the factor structures outlined in Table 2 for the level 1 and level 2 models to identify whether the factor structure is potentially invariant across levels.

Imputation, Estimation, and Fit Evaluation

Data were examined for miscoding and missingness with R (R Core Team, 2018) using the mice package (van Buuren & Groothuis-Oudshoorn, 2011) and naniar package (Tierney et al, 2018). The ML-CFAs were estimated in M*plus* v8 (Muthén & Muthén, 1998–2018). Following the recommendations of Muthén (1994), Dyer, Hanges, and Hall (2005), and Stapleton (2013), the ML-CFA was fit in five steps. In step one, we estimated a single-level CFA ignoring the clustering and dependence among these data. In step two, the corrected level 1 (within) covariance matrix and the corrected level 2 (between) covariance matrix were estimated. In step three, we fit each factor structure (see Table 2) to the corrected level 1 covariance matrix and then to the corrected level 2 covariance matrix. In step four, the level specific fit indices were estimated based on the off level being saturated (Ryu & West, 2009; Ryu, 2014). Lastly, in step five, we estimated the full ML-CFA based on the top two level 1 and level 2 factor structures that fit best in steps 3–4 (so four ML-CFA models in total). While fitting the full ML-CFA, we found that a level 2 being invariant tended to fail to converge or result in poor fit so we decided to also fit a single factor at level 2 with all specifications described in Table 2 at level 1 for consideration.

Because of the complexities that arise when fitting these complex models, we combined two methods of estimation: a robust diagonally weighted least squares estimator called weighted least squares mean and variance-adjusted chi-square (WLSMV) and maximum likelihood with robust standard errors (MLR). Using WLSMV and treating the data as categorical are recommended for estimating CFA models when observed data are ordered with fewer than six categories (Finney & DiStefano, 2013, p. 475). However, simulation studies that examine fit of ML-CFA have so far only investigated MLR (Hsu et al., 2015), and we decided to fit the ML-CFA models with both to gain converging evidence of model fit. When the models were fit with WLSMV, information-based indices such as Akaike's information criteria (AIC) and Bayesian information criterion (BIC) were unavailable. When the full ML-CFA models were estimated with MLR, the optimal fitting model based on the AIC and BIC is whichever model has the lowest value. For both estimators, fit was determined by the model χ^2 and degrees of freedom, comparative fit index (CFI $\geq$ .97), Tucker–Lewis index (TLI $\geq$ .97), root-mean-square error of approximation (RMSEA $\leq$ .03), and the standardized root-mean-square residual within/between (SRMRW $\leq$ .04, SRMRB $\leq$ .10) values according to commonly accepted guidelines (e.g., Hu & Bentler, 1998). The chi-square test of model fit is known to be sensitive to sample size, and not much emphasis

TABLE 2
Summary of Published Factor Structures of the CYRM-28

Item	*Content*	*One-Factor*	*Canada*[a]	*New Zealand*	*Alternative New Zealand*
1.	I cooperate with people around me	Res.	Ind.	Ind.	Ind.
2.	I try to finish what I start	Res.	Ind.	Ind.	Ind.
3.	People think that I am fun to be with	Res.	Ind.	Ind.	Ind.
4.	I am able to solve problems without hurting myself or others	Res.	Ind.	Ind.	Ind.
5.	I know my own strengths	Res.	Ind.	Ind.	Ind.
6.	Spiritual beliefs make me strong	Res.	Con.	Spirit/Com.	Comp. Context
7.	I think it is important to serve my community	Res.	Con.	Spirit/Com.	Comp. Context
8.	My friends are on my side	Res.	Ind.	Ind.	Ind.
9.	My friends stand by me during difficult times	Res.	Ind.	Ind.	Ind.
10.	My caregiver(s) watch me closely	Res.	Rel.	Rel./Family	Rel./Family
11.	My caregiver(s) know a lot about me	Res.	Rel.	Rel./Family	Rel./Family
12.	If I am hungry, there is something to eat	Res.	Rel.	Rel./Family	Rel./Family
13.	I talk to my caregiver(s) about how I feel	Res.	Rel.	Rel./Family	Rel./Family
14.	My caregiver(s) stand by me during difficult times	Res.	Rel.	Rel./Family	Rel./Family
15.	I feel safe when I am with my caregiver(s)	Res.	Rel.	Rel./Family	Rel./Family
16.	I enjoy my caregiver's cultural and family traditions	Res.	Rel.	Rel./Family	Rel./Family
17.	Getting an education is important to me	Res.	Con.	Social/Cul.	Comp. Context
18.	I feel I belong at my school	Res.	Con.	Social/Cul.	Comp. Context
19.	I have role models (people I look up to)	Res.	Con.	Social/Cul.	Comp. Context
20.	I know how to behave in different social situations	Res.	Ind.	Social/Cul.	Comp. Context
21.	I am given opportunities to show others that I am becoming an adult	Res.	Ind.	Social/Cul.	Comp. Context
22.	I know where to go in my community to get help	Res.	Ind.	Social/Cul.	Comp. Context
23.	I have opportunities to develop skills that will be useful later in life	Res.	Ind.	Social/Cul.	Comp. Context
24.	I am proud of my cultural background	Res.	Con.	Social/Cul.	Comp. Context
25.	I am treated fairly in my community	Res.	Con.	Social/Cul.	Comp. Context
26.	I participate in organized activities (e.g., church, mosque, bible study)	Res.	Con.	Spirit/Com.	Comp. Context
27.	I enjoy my community's traditions	Res.	Con.	Spirit/Com.	Comp. Context
28.	I am proud of my nationality	Res.	Con.	Social/Cul.	Comp. Context

Note. Ind = individual; Rel = relational; Con = contextual; Cul = cultural; Com = community; Comp = composite.
The bifactor model was specified with the same item structure as the Alternative New Zealand with a general resilience factor also loaded to each item.
[a]Canada Alternative Model 1 from van Rensburg et al. (2017).
[b]Alternative New Zealand model was given the same name as given from van Rensburg et al. (2017) for consistency between studies.

is placed on a significant *p*-value given the known limitations of this estimate. Methodological research into the performance of the other indices in multilevel settings suggests that the model CFI, TLI, RMSEA, and SRMRW may detect misfit of the level 1 model, while only SRMRB appears to be sensitive to misfit at level 2 (Hsu et al., 2015). A cautionary note is that these authors used MLR to make these conclusions and these recommendations may not hold for WLSMV.

For steps three and four of modeling fitting, the resulting fit of models was determined by resulting fit statistics and plausibility of parameter estimates. For step three, the level specific models were evaluated for fit based on the resulting common fit statistics with criteria of fit outlined above. For step four, the level specific fit statistics were calculated based on the recommendations of Ryu (2014). No published guidelines existed for these level specific fit indices, so the aforementioned criteria were used at benchmarks.

In the estimation of these multilevel models, the factor loadings were constrained to be equal across levels when an equivalent factor model was theorized across levels. This specification aligns with a theoretical expectation that the same mechanism causes variation across communities. The level 2 model defines what mechanism is theorized to cause the group means to vary, and an equivalent structure is more easily interpretable as random effects (Jak, 2018). When the factor structure and loadings are invariant across levels, the model allows for the estimation of ICCs for the latent variables (Heck & Thomas, 2015). However, it should be noted that there is a possibility of incorrectly specifying the measurement model in the level 2 (between) part and the level 1 (within) part of the model. Given one of the goals of our research is to assess variability in measurement across communities, we used the simplifying assumption of equivalence of factor loadings across levels (when applicable). Further research into the measurement properties across levels and communities may be warranted.

Invariance Testing

Invariance of the level 1 measurement model was also investigated across sex and age categories. Invariance testing was conducted in R utilizing the lavaan package (Rosseel, 2012). For invariance testing, we only needed to analyze a level 1 model meaning that the level 2 structure was not subjected to testing. As the groups of interest were at level 1, we could not test for invariance at level 2. To account for the multilevel nature of these data, we created sampling weights for each case so that the clustering does not inadvertently influence parameter estimates.

We investigated six levels of measurement invariance across groups. We investigated invariance for sex (females versus males) and age categories (adolescents and early emerging adults). The levels of invariance tested are configural (baseline model), threshold invariance, loading invariance, strict, partial strict, and partial strict with equal means. Each level is progressively more restrictive than the previous level except for the partial strict invariance. For partial strict invariance, modification indices (MI) were used to guide the freeing of residual variances of items across groups. We used fit indices based on estimating each model with MLR and WLSMV to determine the level of invariance achieved. The rules we followed for assessing the invariance of sequential models were as follows: (1) $\Delta CFI \geq -0.01$ (Kim et al., 2017, p. 539), (2) $\Delta RMSEA \leq 0.03$ (Rutkowski & Svetina, 2014, p. 53; Rutkowski & Svetina, 2017, p. 48), (3) lowest information criteria when possible, and (4) plausibility of parameter estimates.

RESULTS

The results for the literature review will be discussed first, followed by the results of the secondary data analysis. The full text of 50 articles was screened to identify those eligible for inclusion. However, only 12 articles met inclusion criteria and contained enough information for aggregating the reliability evidence; 14 articles met inclusion criteria and contained enough information for the validity evidence.

Reliability Evidence

The studies that included reliability evidence included samples of a total of 7,746 adolescents from 19 unique groups within eight countries. The majority of the studies included samples from vulnerable populations, such as orphans (Makhnach, 2016), those deemed "at-risk" (Liebenberg et al., 2012), with multiple sclerosis (Rainone et al., 2017), or from low socioeconomic status settings (Daigneault et al., 2013). Three studies did not include an average age, but the average age of the participants from the other studies was 15.67 years.

Cronbach's alpha was reported in almost of all the articles as the evidence for reliability. Five studies reported Cronbach's alpha for the total CYRM-28, and six reported values for the three factors of individual, relational/familial, and contextual resilience. Two studies (Govender et al., 2017; Liebenberg et al., 2012) reported ICC values as estimates of internal consistency. Only one study (Makhnach, 2016) did not report any measure of reliability. One study (Daigneault et al., 2013) reported test–retest reliability values for each of the three factors (individual = .76, relational = .84, and contextual = .73) and the total CYRM-28 (.82). Based on the reliability estimates extracted from articles, we computed an aggregated estimate of reliability for general resilience based on the CYRM-28 (0.852, 95% CI = 0.830, 0.874).

Validity Evidence

There were correlations between 37 published measures and the CYRM-28. Thirty correlations were given to demonstrate the relationship with the measure and the total score on the CYRM-28, while 16 studies provided the correlation between the subscales of the CYRM-28 and subscales or total scores of other measures. The types of measures generally focused on issues related to the demonstration of resilience, such as trauma or stress levels, but also included scales related to constructs such as intelligence or personality. Only one study examined the relationship between this resilience measure with another resilience measure. With the exception of the relationship between intelligence and resilience from Avci et al. (2013), all relationships were in the expected direction. Convergent analysis results by means of H-S meta-analysis are reported in Table 3. Results indicate that resilience as measured with the CYRM-28 is positively related to positive relationships and negatively related to negative relationships, mental health issues, and risk.

In terms of factor analyses, four studies conducted a principal component analysis or exploratory factor analysis, four conducted a confirmatory analysis, and two studies conducted both exploratory and confirmatory analyses. For the articles that conducted a principal component analysis or an exploratory factor analysis, four studies identified a three-factor solution and two studies identified a four-factor solution. The studies that identified a three-factor solution identified the factors as those from the original study (i.e., individual, relational, and contextual factors of resilience), although the number of items within the relational and contextual factors differed. The two studies that identified four factors did not align; one study added a personality factor (Makhnach, 2016), and the other had two factors relating to context (Sanders et al., 2017) and both had different numbers of items within each factor. The percent explained from these three factors ranged from 34.47% (Zand et al., 2017) to 50.59% (Sanders et al., 2017). Despite the low number of studies in this sample, these results indicate that more work should be done to identify the structure and other items applicable to measuring resilience.

TABLE 3
Summary of Hunter–Schmidt Meta-Analysis for Convergent Validity Evidence

Method	τ^2	I^2	*Est*	SE	*LL*	*UL*
Total CYRM and Mental Health						
Bare-Bones	0.0084	85.7	−0.321	0.056	−0.434	−0.207
H-S	0.0159	89.7	−0.376	0.077	−0.527	−0.224
Total CYRM and Risk						
Bare-Bones	0.001	31.9	−0.264	0.022	−0.307	−0.222
H-S	0.003	50.8	−0.306	0.030	−0.366	−0.247
Total CYRM and Positive Relationships						
Bare-Bones	0.006	81.1	0.378	0.039	0.302	0.455
H-S	0.012	87.4	0.441	0.054	0.336	0.546
Total CYRM and Negative Relationships						
Bare-Bones	0	0	−0.320	0.040	−0.399	−0.241
H-S	0	0	−0.367	0.046	−0.458	−0.277

Note. τ^2 = estimated amount of heterogeneity; I^2 (total heterogeneity/ total variability) × 100.
Est = estimated correlation between CYRM and outcome; SE = standard error of estimate; LL = 95% CI lower limit; UL = 95% CI upper limit; Bare-Bones = meta-analysis correcting for sample size; H-S = Hunter–Schmidt meta-analysis correcting for sample size and measurement error.

The articles that utilized CFA reported different results in terms of model fit. CFI values revealed a marginal level of fit in two of the studies (Govender et al., 2017; Panter-Brick et al., 2017), with higher CFI values reported in the other four studies. TLI values were consistent with the CFI values reported in the studies, with the exception of one study (Panter-Brick et al., 2017) that did not report a TLI value. RMSEA values were good or acceptable in all but one of the studies (Panter-Brick et al., 2017). This evidence suggests that the CYRM-28 has either three- or four-factor structure across samples, but the small number of studies available limits generalizability.

ML-CFA Model Fit

A summary of the fit of all ML-CFA models estimated is presented in Table 4. The results for all five steps of the fitting process are available upon request. Based on the fit for the single-factor model on the level 2 covariance matrix, we included an additional ML-CFA model where the level 1 model is the best-fitting bifactor model and the level 2 is the single-factor model. In total, this combined model with a unique structure at level 2 provides the best-fitting based on all available data. This final model fit best among tested models (CFI = 0.93, TLI = 0.93, RMSEA = 0.01, SRMRW = 0.05, SRMRB = 0.17). The path diagram of the final model is shown in Figure 1. The general resilience factor at level 1 and level 2 reflects a construct that varies in magnitude across communities. The measurement of resilience is therefore reflecting individual components along with a general influence of one's environment above what can be explained by items pertaining to contextual aspects of resilience.

Examination of ICCs

The estimated ICCs for all observed indicators are reported in Table 5. The estimates range from .038 (Item 12) to .442 (Item 3) with median ICC of .11. This means that, for example, 44.2% of the variability in the response to Item 3 can be attributed to variability among communities. The vast range of ICCs for the observed variables gives evidence that at least some items on then CYRM-28 are not interpreted the same across communities.

For the latent variables, only the ICC for the global resilience can be calculated for the final model. The final ML-CFA model contained the bifactor model at level 1 and a single-factor model at level

TABLE 4
Summary of Full ML-CFA Model Fit

Model	*Estimation*	χ^2	df	*CFI*	*TLI*	*RMSEA*	*SRMRW*	*SRMRB*	*AIC*	*BIC*	*aBIC*
Single Factor[a]	WLSMV	1202.12	727	0.92	0.91	0.01	0.05	0.47			
Canada[a]	WLSMV	1120.02	719	0.93	0.93	0.01	0.05	0.44			
Alt. New Zealand[a]	WLSMV	1173.11	715	0.92	0.92	0.01	0.05	0.47			
New Zealand[a]	WLSMV	1121.31	712	0.93	0.93	0.01	0.05	0.44			
Bifactor[a]	WLSMV	1086.08	696	0.93	0.93	0.01	0.05	0.45			
Bifactor and Single Factor[b,a]	WLSMV	1178.38	780	0.93	0.93	0.01	0.05	0.17			
Single Factor[a]	MLR	7068.89	727	0.77	0.76	0.04	0.04	0.31	504574.3	505335.6	504976.6
Canada[a]	MLR	6614.84	719	0.78	0.77	0.04	0.04	0.30	504155.6	504970.8	504586.3
Alt. New Zealand[a]	MLR	6609.98	715	0.78	0.77	0.04	0.04	0.31	504027.0	504869.2	504472.0
New Zealand[a]	MLR	6599.73	712	0.78	0.77	0.04	0.04	0.31	504013.2	504875.6	504468.9
Bifactor[a]	MLR	5777.08	696	0.81	0.80	0.03	0.03	0.32	503107.5	504077.7	503620.1
Bifactor and Single Factor[b,a]	MLR	5749.14	671	0.81	0.79	0.04	0.04	0.16	503010.8	504149.4	503612.4

Note. [a]Model structure is invariant across levels.
[b]Model structure is different across levels with the first reported structure describing level 1 and the second describing level 2 (e.g., bifactor at level 1 and single factor at level 2). Methodological research into the performance of these indices in a multilevel setting suggests that the model χ^2, CFI, TLI, RMSEA, and SRMRW may detect misfit of the level-1 model while only SRMRB appears to be sensitive to misfit at level 2 (Hsu et al., 2015). The information criterion (AIC, BIC, aBIC) is not available under estimation with WLSMV. Summary and fit information for steps 1–4 are omitted here due to space limitations but are available upon request.

2, with equal factor loadings across levels for general resilience. The level 1 variance was fixed to one for identification. The level 2 variance was .153. Therefore, 15.3% of the variance in resilience can be attributed to community membership.

Invariance Testing

The results of testing the invariance of the level 1 bifactor model are presented in Table 6. Based on these results, the measurement of resilience with

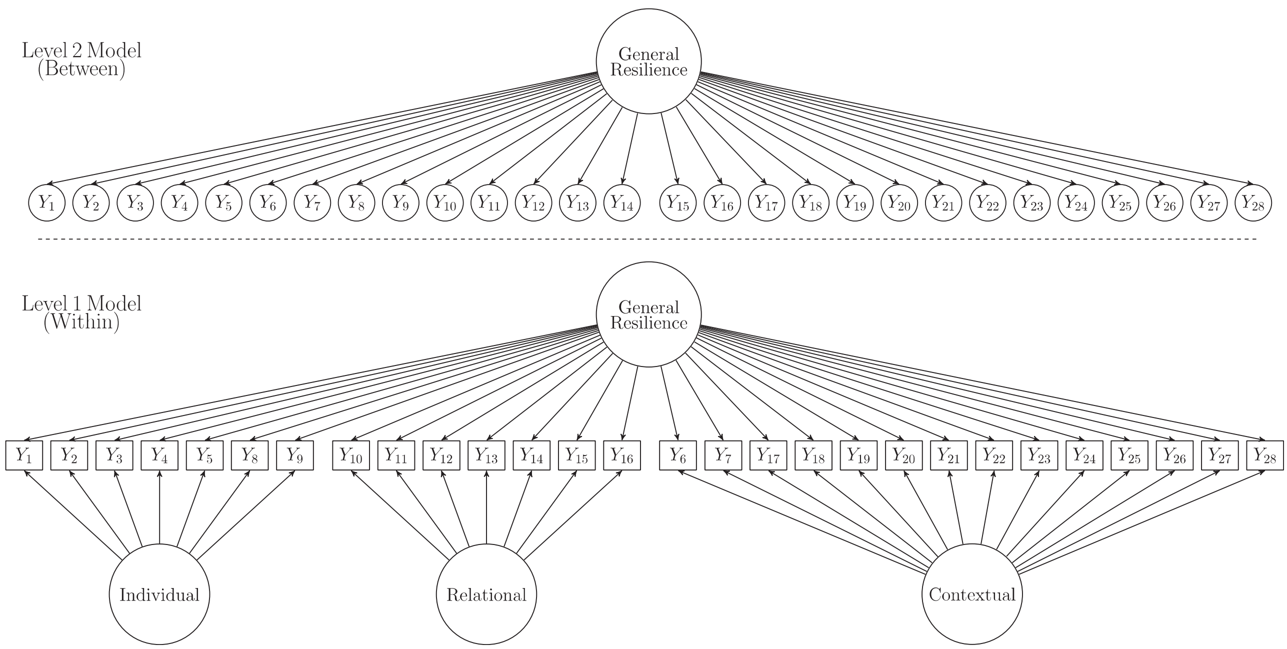

FIGURE 1 Multilevel path diagram of CYRM-28.*Note:* Error variances of observed indicators and level-1 and level-2 latent factors are omitted from this diagram for simplicity. Y_1 is the observed scores on item 1 at level 1 and is therefore shown as a box (manifest variables), while Y_1 at level 2 is a "latent" variable that represents the group aggregate score on item 1. Level 2 represents the mechanism by which groups vary in average response to each item. The factor loadings for level 1 and level 2 of the general resilience factor are equal.

TABLE 5
ICC Estimates for CYRM-28 Items

Item	*ICC*	*Item*	*ICC*	*Item*	*ICC*	*Item*	*ICC*
1	.136	8	.068	15	.053	22	.171
2	.059	9	.054	16	.065	23	.060
3	.442	10	.180	17	.141	24	.152
4	.069	11	.088	18	.117	25	.149
5	.106	12	.038	19	.142	26	.096
6	.080	13	.120	20	.143	27	.040
7	.259	14	.041	21	.182	28	.307

Note. ICC = intraclass correlation coefficient.

the CYRM-28 is invariant across males and females. The invariance testing offered, at a minimum, evidence in support of metric invariance, meaning that the factor loadings are invariant across groups. Equality of factor loadings across levels indicates that items are reflecting similar aspects of resilience across groups. However, the amount of measurement error may depend on group membership because evidence for strict invariance was not found. Across two broad age categories (adolescents vs. early emerging adults), we found evidence for at least scalar invariance across groups. Scalar invariance means that each group has approximately the same response probabilities to each category for each item along with equal factor loadings (see Table 6 for more details).

DISCUSSION

Researchers and practitioners across the world use measures of resilience to determine promotive factors within the lives of youth, yet few studies have tried to understand whether collective evidence supports the cross-cultural applicability of such measures. In this study, we reviewed existing evidence, aggregated the provided reliability and validity evidence, and reanalyzed multiple samples to examine the psychometric properties of the CYRM, a resilience measure that was constructed for utility and comparisons across communities. Findings from previous literature indicate that researchers tend to promote the use of the CYRM

TABLE 6
Results from Invariance Testing Across Sex and Age Categories

								RMSEA					
Level	*Estimation*	*k*	χ^2	df	p	*CFI*	*TLI*	*Est*	*CI LL*	*CI UL*	*SRMR*	*AIC*	*BIC*
Sex: Females vs. Males													
Configural	MLR	224	7007.79	644	0	0.823	0.792	0.056	0.055	0.058	0.043	507275	508784
Metric	MLR	168	7238.29	700	0	0.818	0.803	0.055	0.054	0.056	0.047	507393	508525
Scalar	MLR	144	7484.3	724	0	0.811	0.803	0.055	0.053	0.056	0.047	507591	508561
Strict	MLR	116	7767.28	752	0	0.804	0.803	0.055	0.054	0.056	0.049	507818	508600
Partial Strict	MLR	122	7720.25	746	0	0.806	0.803	0.055	0.054	0.056	0.049	507783	508605
Eq. Means	MLR	118	7771.92	750	0	0.804	0.803	0.055	0.054	0.056	0.049	507827	508622
Configural	WLSMV	336	7085.96	644	0	0.944	0.934	0.056	0.055	0.058	0.051		
Metric	WLSMV	280	7592.88	700	0	0.940	0.935	0.056	0.055	0.057	0.052		
Scalar[a]	WLSMV	200	7844.06	780	0	0.938	0.940	0.053	0.053	0.055	0.052		
Age Categories: Adolescents vs. Early Emerging Adults													
Configural	MLR	224	7000.01	644	0	0.818	0.786	0.056	0.055	0.058	0.043	507818	509327
Metric	MLR	168	7312.45	700	0	0.810	0.795	0.055	0.054	0.056	0.047	508018	509150
Scalar	MLR	144	7472.74	724	0	0.806	0.798	0.055	0.054	0.056	0.048	508130	509101
Strict	MLR	116	7694.25	752	0	0.801	0.800	0.054	0.053	0.056	0.047	508296	509078
Partial Strict	MLR	124	7632.86	744	0	0.802	0.799	0.055	0.053	0.056	0.047	508251	509086
Eq. Means	MLR	120	7689.48	748	0	0.801	0.800	0.054	0.053	0.056	0.048	508300	509108
Configural	WLSMV	336	7085.53	644	0	0.944	0.934	0.057	0.055	0.058	0.050		
Metric	WLSMV	280	7985.11	700	0	0.936	0.931	0.058	0.057	0.059	0.051		
Scalar[b,a]	WLSMV	200	7859.05	780	0	0.938	0.940	0.054	0.053	0.055	0.050		

Note. AIC = Akaike's information criteria; BIC = Bayesian information criterion; CFI = comparative fit index; *df* = degrees of freedom; K = number of parameters; *p* = *p*-value from chi-square test of model fit; RMSEA = root-mean-square error of approximation with 90% confidence interval; SRMR = standard root-mean-square residual; TLI = Tucker–Lewis Index.

[a]When assessing invariance based on the WLSMV estimator, where data were assumed to be ordered categorical, scalar is the highest level of measurement invariance testing possible. This is because the WLSMV assumes that the ordered categorical data have an underlying continuous response that is being approximated and estimated based on the item response frequencies, and the residual variance is fixed to 1 (theta parameterization).

but do not include all relevant information to support these claims. Follow-up analyses indicate the CYRM works the same for different age-groups and genders but does not measure resilience in the same way across communities.

Reliability estimates from the 12 studies reporting on such evidence suggest the overall scale and individual subscales of the CYRM have moderate levels (Cronbach's alpha estimates >.73) of internal consistency. Based on evidence of validity reported in 14 studies on the CYRM, correlations suggest the measure has evidence of validity. Associations were found between the CYRM and 37 published measures on constructs relating to trauma, mental health, and cognition. With the exception of one finding, correlations were in the expected direction.

Exploratory and confirmatory factor analysis results from approximately six previous studies provide mixed validity evidence for the CYRM-28. Exploratory results indicate that resilience may contain three or four factors, and confirmatory results indicate the three-factor model does not fit across all samples. This could be due to the nature of the third and fourth factors. Generally, the fourth factor is a division of the third factor into two specific categories relating to social and spiritual components. For some samples, such as those with a large focus on religious practices, it makes sense for these items to be distinct from general social resources. Church supports are different than extended family supports which are also different from peer resources despite all falling under the umbrella of social support. Researchers should examine the factor structures of similar societies and explore whether the three- or four-factor structure might work better when using the CYRM to measure resilience factors. This is not especially surprising given that the experts who were involved in the instrument development initially argued for four general factors of resilience related to those seen here.

In examining the structure of the CYRM using the raw data that were obtained, additional issues with the factorial validity of the measure were identified. Results of the ML-CFA indicated that resilience in adolescents and young people can be represented by three interrelated domains: individual, relational, and contextual. This is consistent with the recent assertions of resilience as a complex system comprising a combination of personal and environmental resources (i.e., Ungar, 2011). In a recent review of factors promoting resilience for at-risk youth, Meng et al. (2018) found there were also three types of resources—individual, familial, and community—that appear to be akin to the domains of the CYRM identified in this study.

However, we found evidence that a general construct of resilience is also measured by the CYRM items. The validity of describing differences in measured components of resilience across communities is due, in part, to this general resilience factor. The cultural component to general resilience is partially highlighting that communities are known to differ in resilience. This could indicate differences in resilience at the community level or that resilience factors are being measured differentially across communities. It appears as though attempts at using a single measure to assess resilience in different settings may insufficiently capture nuanced variations in the cultural influences that might affect resilience. The CYRM was found to be invariant across sex and age-groups, which offers support for these group comparisons thanks to this large and diverse sample. While the extant literature generally supports the CYRM as a reliable and valid measure of resilience among adolescents and young people, the findings of the analyses performed in this study caution against using the CYRM in its current form to directly compare the resilience of adolescents and young people from different communities.

Implications for Research and Practice

The findings of this study indicate that continued research is necessary to improve our understanding and measurement of resilience. On the whole, publications that reported validity and reliability evidence for the CYRM tended to support the instrument. However, when examined collectively, the results suggest that context-specific research on resilience is needed to understand distinctions in resilience between communities. The developers of the CYRM-28 ought to be commended for creating and validating the original instrument, but researchers should not assume a clear, three-factor structure across all contexts. This is likely due to the limited number of cultures examined. While the creators developed the instrument using 11 cultures, there will be differences within cultures, such as with high- and low-risk populations. It is impossible to create a measure with focus groups from every culture, but the countries from the initial development and the follow-up studies are not representative of the globe; as such, researchers should focus on targeting a narrower sample that will benefit the most from the measure (e.g., refugees and sexual abuse survivors). It would be a pipe dream to hope for one instrument to fully address the complexities of experiences for adolescents, and resilience research suggests that there will be heterogeneity across culture. For researchers who want to compare the resilience of two vastly different cultures, the CYRM would only be a starting place.

Users of the assessment should adhere to the guidelines of the instrument. Although scholars (e.g., Ungar & Liebenberg, 2011) have advocated preliminary testing of the CYRM in communities of interest to determine whether the items may be suitability appropriated, relatively few studies included in this review provided evidence of such procedures. Contextual nuances of resilience, coupled with variations in language and differences in posttranslation interpretability of the items, are likely to affect the cross-cultural validity of the CYRM. With consideration to these issues, further research is required to identify and integrate culturally specific aspects of resilience into measures designed to assess the concept.

The findings also point to large inconsistencies and gaps in scientific reporting of the psychometric evidence on the CYRM. Out of the 50 articles initially identified for inclusion in the review, approximately a quarter (12–14%) of the articles provided reliability *or* validity evidence of the measure. Even fewer articles reported reliability information for the sample under investigation. Not only do sample-specific estimates of reliability provide important precursory information that can affect interpretations of more complex analyses (Kashy, Donnellan, Ackerman, & Russell, 2009), evidence of reliability is an essential part of establishing the cross-cultural applicability of measures and needs to be routinely integrated into reporting practices of scholars involved in research on resilience (and the CYRM).

An important contribution of this study is that it provides evidence of the cross-cultural utility of a widely used measure of resilience, one which is publicly available, can be administered relatively quickly, and assesses multiple components of resilience in adolescents and young people. Invariance testing indicated that adolescents and young people of different sex and age-groups appear to understand and respond to the items in similar ways, suggesting that the CYRM is suitable for measuring and drawing comparisons on resilience across these groups. This is an important step in assessing needed supports within specific communities.

However, the findings suggest that caution should be applied when attempting to use the CYRM to measure and compare resilience levels *across* communities. Certain items, such as those that ask about serving the community (item 7) and being fun to be around (item 3), appear more problematic and should be determined as relevant based on input from the community. Researchers and practitioners are encouraged to evaluate the CYRM-28 in subpopulations and communities of interest before employing it as a measure of resilience (as recommended by the measure developers), especially when evidence of its validity in contexts and subpopulations is mixed or has yet to be established.

Limitations and Future Research Directions

Alongside the strengths of this study, there are several limitations to consider. First, only a portion of published data on the CYRM was made available by other authors. Additional data from more communities may provide different results. Furthermore, more research should examine fit statistics for multilevel models to help support or refute the structural validity conclusions from the ML-CFA. We investigated the invariance of the level 1 model across communities to help find more evidence of the latent structure where we generally found supporting evidence of the level 1 model (see Appendix 1). Given few resilience measures attempt to capture promotive factors within and between communities, future research should examine the factorial validity of the CYRM before using the instrument for unstudied groups. Second, we did not investigate community-level covariates on the measurement of resilience, which could influence model fit and conclusions of which factor structure offers the best fit. During the examination of measurement invariance across age and sex, the organizational structure of these data was accounted for by computing sampling weights. However, we did not have access to the full range of information necessary to compute the most precise sampling weights (i.e., number of possible individuals within each community that could have been sampled), which may have resulted in some communities having larger contributions to the results generated. This would lead to issues of interpretation if, for example, resilience was more homogenous among adolescents and young people from larger communities as compared to their counterparts from smaller communities.

Third, a confirmatory analytic approach was used to examine a select set of models, which may be seen as restrictive. An alternative approach would have involved splitting the data and sequentially performing exploratory and confirmatory factor analyses, but multilevel CFA results can be severely influenced by sample size, particularly the results of the level 2 model (Wu & Kwok, 2012). Such an approach would have led to a lower level 1 sample size for the multilevel CFA, which would have influenced the precision of the estimation of the level 2 model. Additionally, some of the community groups had very small sample sizes and group sample sizes ranged from 7 to 499. Not only would splitting smaller groups severely limit their contributions to model estimation, but it would also be lower average within group totals that could bias results when the number of units within a group is unequal across groups (Hox & Maas, 2001).

CONCLUSION

In summary, the CYRM has been studied multiple times in countries across the globe. Previous research has supported the internal consistency and convergent as well as discriminant validity of the measure, although evidence of its factorial validity has been mixed. From a systematic review of the literature, researchers who have used the CYRM often did not have rigorous reporting; only a few of the studies included crucial information regarding reliability or validity evidence. The findings of this study suggest there is considerable variability in the contextual suitability of the CYRM, indicating that resilience is difficult to compare across settings using this measure yet may be possible for researchers to examine resilience between age-groups, sexes, and from ages nine to 22.

REFERENCES

Ahern, N. R., Kiehl, E. M., Sole, M. L., & Byers, J. (2006). A review of instruments measuring resilience. *Issues in Comprehensive Pediatric Nursing*, *29*, 103–125. https://doi.org/10.1080/01460860600677643

Avci, G., Hanten, G., Schmidt, A., Li, X., Orsten, K., Faber, J., ⋯ Newsome, M. R. (2013). Cognitive contributors to resilience in youth from underserved populations: A brief report. *Journal of Public Mental Health*, *12*, 165–170. https://doi.org/10.1108/JPMH-02-2013-0005

Blakemore, S., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Bonanno, G. A., Westphal, M., & Mancini, A. D. (2011). Resilience to loss and potential trauma. *Annual Review of Clinical Psychology*, *7*, 511–535. https://doi.org/10.1146/annurev-clinpsy-032210-104526

Clauss-Ehlers, C. S. (2008). Sociocultural factors, resilience, and coping: Support for a culturally sensitive measure of resilience. *Journal of Applied Developmental Psychology*, *29*(3), 197–212. https://doi.org/10.1016/j.appdev.2008.02.004

Daigneault, I., Dion, J., Hebert, M., McDuff, P., & Collin-Vezina, D. (2013). Psychometric properties of the Child and Youth Resilience Measure (CYRM-28) among samples of French Canadian Youth. *Child Abuse & Neglect*, *37*, 160–171. https://doi.org/10.1016/j.chiabu.2012.06.004

Doyle, C., & Cicchetti, D. (2017). From the cradle to the grave: The effect of adverse caregiving environments on attachment and relationships throughout the lifespan. *Clinical Psychology*, *24*, 203–217. https://doi.org/10.1111/cpsp.12192

Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *Leadership Quarterly*, *16*, 149–167. https://doi.org/10.1016/j.leaqua.2004.09.009

Finney, S., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 439–492). Charlotte, NC: Information Age Publishing.

Furman, W., & Winkles, J. K. (2012). Transformations in heterosexual romantic relationships across the transition into adulthood. *Relationship pathways: From adolescence to young adulthood*, 191–213.

Garmezy, N. (1981). Children under stress: Perspectives on antecedents and correlates of vulnerability and resistance to pathology. In A. Rabin, J. Arnoff, A. Barclay, & R. Zuckers (Eds.), *Further explorations in personality* (pp. 196–269). New York, NY: Wiley Interscience.

Govender, K., Cowden, R. G., Asanta, K. O., George, G., & Reardon, C. (2017). Validation of the child and youth resilience measure among South African adolescents. *PLoS One*, *12*, 1–13. https://doi.org/10.1371/journal.pone.0185815

Hartman, J. L., Turner, M. G., Daigle, L. E., Exum, L., & Cullen, F. T. (2008). Gender differences in protective factors: Implications for understanding resiliency. *International Journal of Offender Therapy and Comparative Criminology*, *53*, 250–278.

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus*, 3rd edn. New York, NY: Routledge.

Hox, J. J., & Maas, C. J. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling*, *8*(2), 157–174. https://doi.org/10.1207/S15328007SEM0802_1

Hsu, H., Kwok, O., Lin, H., & Acosta, S. (2015). Detecting misspecified multilevel structural equation models with common fit indices: A Monte Carlo study. *Multivariate Behavioral Research*, *50*, 197–215. https://doi.org/10.1080/00273171.2014.977429

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Hunter, J., & Schmidt, F. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*, 2nd edn. Thousand Oaks, CA: SAGE Publications Inc.

Infurna, F. J., & Luthar, S. S. (2016). Resilience to major life stressors is not as common as thought. *Perspectives on Psychological Science*, *11*, 175–194. https://doi.org/10.1177/1745691615621271

Jak, S. (2018). Cross-level invariance in multilevel factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(4), 607–622. https://doi.org/10.1080/10705511.2018.1534205

Jorgensen, T. D., Kite, B., Chen, P.-Y., & Short, S. D. (2017). Finally! A valid test of configural invariance using permutation in multigroup CFA. In L. A. van der Ark, M. Wiberg, S. A. Culpepper, J. A. Douglas, & W. C. Wang (Eds.), *Quantitative psychology: IMPS 2016* (pp. 93–103). Asheville, NC: Psychometric Society.

Kashy, D. A., Donnellan, M. B., Ackerman, R. A., & Russell, D. W. (2009). Reporting and interpreting research in PSPB: Practices, principles, and pragmatics. *Personality and Social Psychology Bulletin*, *35*, 1131–1142. https://doi.org/10.1177/0146167208331253

Kim, E. S., Cao, C., Wang, Y., & Nguyen, D. T. (2017). Measurement invariance testing with many groups: A comparison of five approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*, 524–544. https://doi.org/10.1080/10705511.2017.1304822

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York, NY: Guilford Publications.

Liebenberg, L., Ungar, M., & Vijver, F. V. D. (2012). Validation of the child and youth resilience measure-28 (CYRM-28) among Canadian youth. *Research on Social Work Practice*, *22*, 219–226. https://doi.org/10.1177/1049731511428619

Linders, A. (2017). Deconstructing adolescence. In A. L. Cherry, V. Baltag, & M. E. Dillon (Eds.), *International handbook on adolescent health and development: The public health response* (pp. 15–28). Cham, Switzerland: Springer International Publishing.

Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development*, *71*(3), 543–562. https://doi.org/10.1111/1467-8624.00164

Madsen, S. D., & Collins, W. A. (2011). The salience of adolescent romantic experiences for romantic relationship qualities in young adulthood. *Journal of Research on Adolescence*, *21*, 789–801. https://doi.org/10.1111/j.1532-7795.2011.00737.x

Makhnach, A. V. (2016). Resilience in Russian youth. *International Journal of Adolescence and Youth*, *21*, 195–214. https://doi.org/10.1080/02673843.2013.815116

Masten, A. S. (2014a). *Ordinary magic: Resilience in development*. New York, NY: Guilford Press.

Masten, A. S. (2014b). Global perspectives on resilience in children and youth. *Child Development*, *85*, 6–20. https://doi.org/10.1111/cdev.12205

McDougall, P., & Vaillancourt, T. (2015). Long-term adult outcomes of peer victimization in childhood and adolescence: Pathways to adjustment and maladjustment. *American Psychologist*, *70*, 300–310. https://doi.org/10.1037/a0039174

Meng, X., Fleury, M. J., Xiang, Y. T., Li, M., & D'Arcy, C. (2018). Resilience and protective factors among people with a history of child maltreatment: a systematic review. *Social Psychiatry and Psychiatric Epidemiology*, *53* (5), 453–475. https://doi.org/10.1007/s00127-018-1485-2

Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research.*, *22*, 376–398. https://doi.org/10.1177/0049124194022003006

Muthén, L. K., & Muthén, B. O. (1998–2018). *Mplus user's guide*, 8th edn. Los Angeles, CA: Muthén & Muthén.

Pangallo, A., Zibarras, L., Lewis, R., & Flaxman, P. (2015). Resilience through the lens of interactionism: A systematic review. *Psychological Assessment*, *27*, 1–20. https://doi.org/10.1037/pas0000024

Panter-Brick, C., Dajani, R., Ager, A., Hadfield, K., Eggerman, M., & Ungar, M. (2017). Resilience in context: A brief and culturally grounded measure for Syrian and Jordanian host-community adolescents. *Child Development*, *89*(5), 1–18, https://doi.org/10.1111/cdev.12868

Park, S., & Schepp, K. G. (2015). A systematic review of research on children of alcoholics: Their inherent resilience and vulnerability. *Journal of Child and Family Studies*, *24*, 1222–1231. https://doi.org/10.1007/s10826-014-9930-7

Rainone, N., Chiodi, A., Lanzillo, R., Magri, V., Napolitano, A., Morra, V. B., ⋯ Freda, M. F. (2017). Affective disorders and Health-Related Quality of Life (HRQoL) in adolescents and young adults with Multiple Sclerosis (MS): The moderating role of resilience. *Quality of Life Research*, *26*, 727–736. https://doi.org/10.1007/s11136-016-1466-4

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Ruhl, H., Dolan, E. A., & Buhrmester, D. (2014). Adolescent attachment trajectories with mothers and fathers: The importance of parent-child relationship experiences and gender. *Journal of Research on Adolescence*, *25*, 427–442. https://doi.org/10.1111/jora.12144

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, *74*(1), 31–57. https://doi.org/10.1177/0013164413498257

Rutkowski, L., & Svetina, D. (2017). Measurement invariance in international surveys: Categorical indicators & fit measure performance. *Applied Measurement in Education*, *30*(1), 39–51. 10.1080/08957347.2016.1243540

Ryu, E., & West, S. G. (2009). level specific evaluation of model fit in multilevel structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 583–601. https://doi.org/10.1080/10705510903203466

Ryu, E. (2014). Model fit evaluation in multilevel structural equation models. *Frontiers in Psychology*, *5*, 1–9. https://doi.org/10.3389/fpsyg.2014.00081

Sanders, J., Munford, R., Thimasarn-Anwar, T., & Liebenberg, L. (2017). Validation of the child and youth resilience measure (CYRM-28) on a sample of at-risk New Zealand youth. *Research on Social Work Practice*, *27*, 827–840. https://doi.org/10.1177/1049731515614102

Sanders, J., Munford, R., Thimasarn-Anwar, T., Liebenberg, L., & Ungar, M. (2015). The role of positive youth development practices in building resilience and enhancing wellbeing for at-risk youth. *Child Abuse & Neglect*, *42*, 40–53.

Smith-Osborne, A., & Whitehill Bolton, K. (2013). Assessing resilience: A review of measures across the life course. *Journal of Evidence-Based Social Work*, *10*, 111–126. https://doi.org/10.1080/15433714.2011.597305

Southwick, S. M., Bonanno, G. A., Masten, A. S., Panter-Brick, C., & Yehuda, R. (2014). Resilience definitions, theory, and challenges: interdisciplinary perspectives. *European Journal of Psychotraumatology*, *5*(1), 25338. https://doi.org/10.3402/ejpt.v5.25338

Stapleton, L. M. (2013). Using multilevel structural equation modeling techniques with complex sample data. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course*, 2nd edn (pp. 521–562). Charlotte, NC: Information Age Publishing.

Stapleton, L. M., Yang, J. S., & Hancock, G. R. (2016). Construct meaning in multilevel settings. *Journal of Educational and Behavioral Statistics*, *41*(5), 481–520. https://doi.org/10.3102/1076998616646200

Steinberg, L. (2014). *Age of opportunity: Lessons from the new science of adolescence*. New York, NY: Houghton Mifflin Harcourt.

Sun, J., & Stewart, D. (2007). Age and gender effects on resilience in children and adolescents. *International Journal of Mental Health Promotion*, *9*(4), 16–25.

Tierney, N., Cook, D., McBain, M., & Fay, C. (2018).). *naniar: Data structures, summaries, and visualisations for missing data. R package version 0.2.0*. Retrieved from https://CRAN.R-project.org/package=naniar

Ungar, M. (2011). The social ecology of resilience: Addressing contextual and cultural ambiguity of a nascent construct. *American Journal of Orthopsychiatry*, *81*, 1–17. https://doi.org/10.1111/j.1939-0025.2010.01067.x

Ungar, M., & Liebenberg, L. (2009). Cross-cultural consultation leading to the development of a valid measure of youth resilience: The International Resilience Project. *Studia Psychologica*, *51*, 259–268.

Ungar, M., & Liebenberg, L. (2011). Assessing resilience across cultures using mixed methods: Construction of the child and youth resilience measure. *Journal of Mixed Methods Research*, *5*, 126–149. https://doi.org/10.1177/1558689811400607

Ungar, M., & Liebenberg, L. (2013). A measure of resilience with contextual sensitivity—The CYRM-28: Exploring the tension between homogeneity and heterogeneity in resilience theory and research. In S. Prince-Embury, & D. Saklofske (Eds.), *Resilience in children, adolescents, and adults* (pp. 245–255). New York, NY: Springer.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*, 1–67. Retrieved from https://www.jstatsoft.org/v45/i03/

van Harmelen, A.-L., Kievit, R. A., Ioannidis, K., Neufeld, S., Jones, P. B., Bullmore, E., ⋯ Goodyer, I. (2017). Adolescent friendships predict later resilient functioning across psychosocial domains in a healthy community cohort. *Psychological Medicine*, *47*, 2312–2322. https://doi.org/10.1017/S0033291717000836

van Hoorn, J., van Dijk, E., Meuwese, R., Rieffe, C., & Crone, E. A. (2014). Peer influence on prosocial behavior in adolescence. *Journal of Research on Adolescence*, *26*, 90–100. https://doi.org/10.1111/jora.12173

van Rensburg, A. C., Theron, L. C., & Ungar, M. (2017). Using the CYRM-28 with South African young people: A factor structure analysis. *Research on Social Work Practice*, *29*, 93–102. https://doi.org/10.1177/1049731517710326

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48. Retrieved from http://www.jstatsoft.org/v36/i03/

Wexler, L. M., DiFluvio, G., & Burke, T. K. (2009). Resilience and marginalized youth: Making a case for personal and collective meaning-making as part of resilience research in public health. *Social Science & Medicine*, *69*(4), 565–570. https://doi.org/10.1016/j.socscimed.2009.06.022

Windle, G., Bennett, K. M., & Noyes, J. (2011). A methodological review of resilience measurement scales. *Health and Quality of Life Outcomes*, *9*, 8. https://doi.org/10.1186/1477-7525-9-8

Wu, J., & Kwok, O. (2012). Using SEM to analyze complex survey data: A comparison between design-based single-level and model-based multilevel approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(1), 16–35. https://doi.org/10.1080/10705511.2012.634703

Zand, B. K., Liebenberg, L., & Shamloo, Z. S. (2017). Validation of the factorial structure of the child and youth resilience measure for use with Iranian youth. *Child Indicators Research*, *10*, 797–809. https://doi.org/10.1007/s12187-016-9412-0

APPENDIX

COMMUNITY-LEVEL INVARIANCE TESTING

Invariance testing of the level 1 model across communities was conducted using a permutation test (Jorgensen et al., 2017). The permutation test is conducted by first fitting the configural model and saving the fit indices and model results. Secondly, the group identifiers are randomly reassigned (this is most easily done by sampling without replacement on a vector of all the IDs). Third, the configural model is refit using the "new" group IDs that have been randomly shuffled across all the cases. Steps 2 and 3 are repeated many times (1,000 in this case; more could be done, but this process is computationally intensive), and all fit indices are stored. Lastly, the fit indices (e.g., model chi-square, RMSEA, and CFI) of the permutations are compared to the fit of the model with the true group IDs. For the fit indices, we computed the proportion of permutations that were more extreme than the original model. For example, for the chi-square statistic, we computed the proportion of permutations that results in a chi-square value less than the value from the model with the original group IDs. For CFI, the p-value is computed as the proportion of permutations that are higher than the original value. If this p-value estimate is > 0.05, then we have evidence of configural invariance.

RESULTS

The permutation test gave evidence of configural invariance (chi-square p-value = 0.119; CFI p-value = 0.092; and RMSEA p-value = 0.130). These results suggest that the level 1 factor structure is at least approximately equal, or at least similar enough, to suggest that the same factor structure is plausible. However, the MI testing (comparing the configural model to metric model) did not support equal factor loadings ($\Delta\chi^2 = 2{,}993$, $\Delta df = 1{,}716$, $p < .001$; $\Delta CFI = -0.05$; $\Delta RMSEA = 0.007$). This gives evidence to the variability in measurement across communities, which lines up with our finding that resilience varies across communities. The multilevel factor model aimed to help identify a mechanism that we can use to help explain these cross-community differences.