Check for updates

Original Research Article

Assessment
2025, Vol. 32(4) 501–520
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10731911241256434
journals.sagepub.com/home/asm
Sage

Assessing Social, Emotional, and Behavioral Skills in Just a Few Minutes: 96-, 45-, and 20-Item Short Forms of the BESSI

Madison N. Sewell[1], Hee J. Yoon[1], Clemens M. Lechner[2], Christopher M. Napolitano[1], Beatrice Rammstedt[2], Brent W. Roberts[1,3], and Christopher J. Soto[4]

Abstract

Social, emotional, and behavioral (SEB) skills matter for individuals' well-being and success. The behavioral, emotional, and social skills inventory (BESSI) uses 192 items to assess 32 specific SEB skills across five broad skill domains. This research developed three short forms of the BESSI-192 and explored their measurement properties, predictive validity, and cross-cultural comparability. We found that BESSI-96, BESSI-45, and BESSI-20 largely captured the psychological content of the BESSI-192 measure, retained a robust multidimensional structure, and demonstrated adequate reliability. At the domain and facet level, the BESSI short forms showed patterns of associations with external criteria that were similar to the BESSI-192 and preserved most of the BESSI-192's predictive power. The BESSI short forms also demonstrated full or partial measurement invariance between the primarily U.S.-based and German adult samples. We conclude by discussing contexts in which the short forms may be useful for researchers and practitioners.

Keywords

socioemotional skills, short measures, noncognitive skills, psychometrics, social, emotional, and behavioral skills

Scholars and practitioners across disciplines are paying increasing attention to the noncognitive qualities that promote success and thriving (Allen et al., 2020; Casillas et al., 2015; Collaborative for Academic, Social, and Emotional Learning [CASEL], 2020; Kautz et al., 2014; Organisation for Economic Co-operation and Development [OECD], 2015). These qualities, which are largely independent of intelligence (Lechner et al., 2022), are sometimes termed noncognitive skills, soft skills, or *social, emotional, and behavioral (SEB) skills*—a person's capacities to manage goal—and learning-directed behaviors, maintain social relationships, and regulate emotions (Soto et al., 2021). Recently, a 192-item measure of SEB skills, the Behavioral, Emotional, and Social Skills Inventory (BESSI-192), was developed (Soto et al., 2022a). However, given the interest among researchers and practitioners in measuring and intervening on SEB skills (Domitrovich et al., 2017; Napolitano et al., 2021) and large-scale surveys with strict constraints on assessment time (OECD, 2021), a reliable, valid, and brief measure of SEB skills is needed to reduce assessment burden.

Therefore, the purpose of this research was to develop three short forms of the BESSI-192—a 96-item version, a 45-item version, and a 20-item version—and to investigate the extent to which these short forms retain the structure, reliability, and validity of the BESSI-192. In addition, this research explored whether the BESSI short forms were comparable across two languages and cultural contexts: Germany and the United States. In the following sections, we further define SEB skills, provide an overview of the BESSI-192, and discuss the need for short forms and measurement equivalence.

[1]University of Illinois at Urbana-Champaign, USA
[2]GESIS—Leibniz Institute for the Social Sciences, Mannheim, Germany
[3]Tübingen University, Germany
[4]Colby College, Waterville, ME, USA

Corresponding Author:
Madison N. Sewell, Department of Educational Psychology, University of Illinois at Urbana-Champaign, 1310 South Sixth Street, Champaign, IL 61820, USA.
Email: msewell2@illinois.edu

Defining and Measuring SEB Skills

SEB skills have been linked to a variety of positive outcomes, including higher levels of academic engagement and achievement, better well-being and physical health, higher quality relationships with family, friends, and teachers, and higher levels of civic engagement (Guo et al., 2022; Kautz et al., 2014; Oberle et al., 2014; Sewell et al., 2023; Soto et al., 2022a, 2022b). In addition, meta-analyses of school-based social and emotional learning interventions have found that SEB skill development is linked with improved mental health, including reduction in emotional distress, conduct problems, and clinical diagnoses (Durlak et al., 2011; Taylor et al., 2017). These findings underscore the promise of measuring and intervening in SEB skills in educational, clinical, and applied contexts.

The enthusiasm for SEB skills across the literatures has resulted in a proliferation of terms, frameworks, and inventories to define and measure these skills (Berg et al., 2017). In recent years, strides have been made to build consensus on SEB skills. For example, many scholars have drawn parallels between SEB skills and the Big Five personality traits, arguing that SEB skills can be organized in terms of a five-domain framework (Abrahams et al., 2019; Casillas et al., 2015; Napolitano et al., 2021; OECD, 2015, 2021; Soto et al., 2021; Walton et al., 2021). While SEB skills and personality traits share similar SEB content, SEB skills are not identical to traits. Whereas traits capture someone's typical patterns of thoughts, feelings, and behaviors, SEB skills capture functional capacities: how someone is *capable* of thinking, feeling, and behaving when needed (Soto et al., 2021). For example, a person who tends to feel anxious may also be highly capable of calming themselves when needed.

To reconcile the diversity in psychological content across SEB skill frameworks, reflect the growing consensus among researchers that SEB skills can be organized across five domains, and explicitly differentiate between trait and skill assessment, Soto and colleagues (2022a) developed the BESSI-192. The BESSI-192 measures 32 lower-order skill facets organized across five higher-order skill domains: Social Engagement Skills, Cooperation Skills, Self-Management Skills, Emotional Resilience Skills, and Innovation Skills. The BESSI-192 also has notable psychometric strengths. First, it reliably captures a wide range of SEB skills (Lechner et al., 2022; Soto et al., 2022). Second, it has a robust multidimensional structure and exhibits conceptually coherent convergence and divergence from the Big Five and other socioemotional skill measures (Soto et al., 2022a, 2022b). Finally, the BESSI-192 provides incremental validity over personality traits when predicting self-reported outcomes and observed behavior (Breil et al., 2022; Soto et al., 2022b, 2023; Yoon et al., 2024).

Goals for Developing Short Forms

The full BESSI-192 has a completion time of about 20 minutes. Researchers and practitioners wanting to measure SEB skills in different contexts will likely have varying assessment goals and face different constraints. To address these diverse needs, we aimed to develop three short forms for the BESSI-192 that differ in their overall length, as well as their focus on assessing broad skill domains versus narrower facets. With each short form, we also had an overarching aim to retain as much information as possible from the BESSI-192 for better comparability across future studies.

The first short form addresses the use case where researchers and practitioners would like to comprehensively assess SEB skills at both the domain and facet levels but cannot devote the time needed to administer the BESSI-192. Thus, we halved the number of items per facet to three items each (or 96 items in total). The BESSI-96 would maintain the BESSI-192's breadth of content, allow users to compute observed scores at both the domain and facet levels, and cut assessment time to approximately 10 minutes. In addition, users would be able to specify latent models of skill facets because three items per latent variable provide just enough information to estimate loadings and residual variances.

The second short form addresses the use case where researchers and practitioners face more stringent constraints on assessment time or participant attention, but still wish to thoroughly assess all five major skill domains. For example, they may want to include the BESSI in a survey battery alongside several other measures. To address such contexts, we next aimed to develop a shorter form of the BESSI-192 that uses nine items to assess each skill domain and thus includes 45 items in total. This form would still be long enough to include at least one item from every facet comprising each skill domain, and it would therefore capture almost as much domain-level information as the full-length measure (see Soto & John, 2019) while further reducing the required assessment time to only 5 minutes. However, with only one or two items selected from each BESSI-192 facet, it would be too short to reliably assess individual facets.

The third short form addresses the use case where researchers and practitioners face even more severe assessment constraints. For example, they may wish to include the BESSI in a large-scale panel survey or in a design that involves having each participant rate multiple target individuals (e.g., teachers rating students). To allow for extremely efficient assessment in such

cases, we aimed to develop an even shorter form of the BESSI that includes four items per skill domain, and thus 20 items total. This form yields less content breadth, reliability, and validity than the previously described versions. However, it still includes item content from most core skill facets, allowing users to compute observed scores and estimate latent variable models at the domain level and also reduce assessment time to only 2 minutes.

Investigating Measurement Equivalence Across Cultures

Given the international interest in assessing socioemotional skills, a secondary goal of this research was to begin investigating whether the BESSI functions equivalently across different cultural settings, thereby allowing for meaningful comparisons of scores and associations with other variables (Cheung & Rensvold, 2002; Church, 2001). Cross-cultural considerations in measurement development are important because respondents from different cultural settings might have different response styles (e.g., neutral vs. extreme responding styles; Mõttus et al., 2012) or use different reference groups for judging skills (Heine et al., 2002; McCrae et al., 1998). Furthermore, the meaning of certain items could be interpreted differently depending on the sociocultural context of assessment. Such between-culture variance can undermine the validity of research examining the mean levels and correlations of psychological inventories across cultures.

To begin testing the cross-cultural generalizability of the BESSI short forms, we explore their measurement equivalence by using a multigroup confirmatory factor analysis-based (CFA) approach to measurement invariance (Meredith, 1993; Millsap, 2011). Investigating measurement invariance entails testing a series of CFA models with differing levels of equality constraints: (a) no constraints (configural invariance), (b) loading equality constraints (metric invariance), and (c) threshold or intercept equality constraints (scalar invariance). Establishing configural invariance indicates that factor structures are equivalent across groups, metric invariance indicates that each item contributes similarly to the latent factor across groups, and scalar invariance indicates that mean differences in the latent SEB skill facets and domains capture all mean differences in the shared variance of the items (Putnick & Bornstein, 2016). If the BESSI short forms demonstrate measurement invariance, then that would suggest BESSI items function equivalently and mean the same thing to individuals from Germany and the United States.

Overview of This Research

In sum, this research aimed to develop three reliable and valid short forms of the BESSI-192, to explore the extent to which the BESSI-96, BESSI-45, and BESSI-20 retain the validity of the BESSI-192, and to investigate whether the BESSI short forms are comparable between Germany and the United States. We first developed the 96-, 45-, and 20-item versions of the BESSI-192 using a combination of empirical and rational criteria. Next, we examined the short forms' measurement properties, including factor structure and reliability, using data from six samples. Then, we compared the predictive validity of the BESSI short forms to the full measure using data from three samples. Finally, we tested the short forms' measurement equivalence across two cultural contexts—the United States and Germany.

We generally expected that the three short forms would converge strongly with the full BESSI-192 domain and facet scales, retain a robust multidimensional structure, and demonstrate adequate reliability. We also expected that the short forms would show patterns of skill-outcome associations similar to the full BESSI-192 (with column-vector correlations of .50 or higher) and would also retain most of the full measure's capacity to predict self-reported and school-reported outcomes (with explained variance ratios of .70 or higher). We based these criteria on findings from Soto and John (2017), who used a similar empirical–rational approach to developing the Big Five Inventory-2 (BFI-2) short forms. We did not formally hypothesize about cross-cultural comparability because our available datasets differed in terms of sampling methods, and, therefore, we treat these analyses as more exploratory. Despite this methodological confound, initial findings on cross-cultural comparability may inform future research on the utility of the BESSI short forms across various cultural contexts. We report all data exclusions and all measures in the study. No manipulations were used in this research.

Method

Samples and Procedure

We used data from four samples to select items for the BESSI-96, BESSI-45, and BESSI-20 short forms. The data included a U.S. High School Student sample, a U.S. College Student sample, a predominantly U.S.-based adult internet volunteer sample (Internet 1), and a U.S. adult observer sample; all of which were previously analyzed to validate the BESSI-192 (Soto et al., 2022a, Studies 2–5; see Sewell et al., 2022, for a detailed description of the samples). We then used these four selection samples and two independent validation samples, to

Table 1. Sample Demographics and Sampling Information.

Sample name	*N*	Age *M* (*SD*)	Sampling	% M.	% F.	%W.	% B.	% A.	%H.L.	%Mu.	% O.
High school[a]	499	15.6 (1.2)	Urban and suburban high schools	45.1	52.7	42.1	34.1	3.0	18.4	4.4	0.6
College[a]	322	19.4 (1.1)	Psychology classes from one college	29.8	69.6	74.5	6.5	20.2	8.7	11.8	2.2
Internet 1	400	25.2 (9.9)	Noncommercial website	50.0	50.0	67.8	5.3	13.8	8.3	8.0	10.8
Internet 2	600	26.4 (9.1)	Noncommercial website	50.0	50.0	75.5	2.3	13.5	8.8	7.0	6.3
Observer	488	49.3 (16.3)	Qualtrics Sample Service	50.0	50.0	84.2	6.1	4.7	3.1	0.0	1.8
Panel[b]	483	43.5 (14.1)	Respondi AG	50.5	49.5	—	—	—	—	—	—

Note. M. = male; F. = female; W. = White/Caucasian; B. = Black/African American; A. = Asian/Asian American; H.L. = Hispanic/Latino; Mu. = multiracial; O. = other race or ethnicity.

[a]Approximately 1.2% in the High School sample and 0.6% in the College sample indicated another gender identity. [b]Racial and ethnic information was not collected in the Panel sample.

investigate the short forms' measurement properties, predictive validity, and measurement invariance. The two validation samples included a Panel sample of German adults (see Lechner et al., 2022, Study 2) and a second sample of predominantly U.S.-based adult internet volunteers (Internet 2; see Soto et al., 2022a, Studies 2–3). In Internet 1, approximately 57% of participants reported residing in the United States, 10% reported residing in the United Kingdom, Ireland, Canada, New Zealand, or Australia, 18% reported another country, and 16% did not report any geographic information. In Internet 2, approximately, 55% of participants reported residing in the United States, 10% reported residing in the United Kingdom, Ireland, Canada, New Zealand, or Australia, 18% reported another country, and 17% did not report any geographic information. All participants in both internet samples reported that they spoke English fluently.

We did not exclude any participants that were included in the publicly available data. Full descriptions of exclusion criteria used for each dataset were reported in the prior studies (see Sewell et al., 2022; Lechner et al., 2022). Some exclusion criteria included (a) careless responding (e.g., straight-lining or failing attention checks), (b) not reporting demographics or reporting demographics inconsistently, (c) completing the survey faster than 1/3 of the median time, and (d) submitting more than one survey. Study procedures were approved by the local institutional review board or were in line with the Helsinki Declaration and the European Union's General Data Protection Regulation if an institutional review was not required by the country in which the research was conducted. Sampling and demographic information for each sample are provided in Table 1.

Measures

All participants completed the BESSI-192 in either English or German (Lechner et al., 2022; Soto et al., 2022a). External criteria were assessed in the College Student sample, High School Student sample, and Panel sample. Participants in the College Student sample completed five different measures of personal qualities. Participants in the High School sample completed a battery of developmentally relevant measures covering academic, occupational, social, and well-being outcomes. Finally, participants in the Panel sample completed a series of items assessing several occupational and well-being outcomes. The measures are described in detail below, and descriptive statistics for all measures and alpha reliability coefficients for external criteria can be found in the online Supplemental Tables (OSTs) S1 to S3.

The Behavioral, Emotional, and Social Skills Inventory. The BESSI-192 items are short phrases describing a behavior (e.g., "Lead a group of people"). All participants, except those from the Observer sample, self-reported how well they could perform each behavior using a 5-point scale ranging from "Not at all well" to "Extremely well." Participants in the Observer sample were asked to provide detailed and accurate ratings of another person's SEB skills. Items from the BESSI-192 were used to calculate observed domain and facet scale scores for the BESSI-96, BESSI-45, and BESSI-20 as well as to estimate the short forms' domain and facet latent factors. A subsample of participants from the College sample ($N = 219$) completed the BESSI-192 for a second time with an average retest interval of 17.6 weeks between assessments.

College Student Personal Qualities. All participants in the College Student sample completed the BFI-2 (Soto & John, 2017), and the positive youth development short form (PYD-SF; Geldhof et al., 2014). The BFI-2 measures the Big Five personality domains of extraversion, agreeableness, conscientiousness, negative emotionality, and open-mindedness with 60 items. Participants

indicated whether the BFI-2 statements were descriptive of them on a 5-point agreement scale (1 = disagree strongly, 5 = agree strongly). The PYD-SF measures developmental strengths across five domains—competence, confidence, character, caring, and connection—and consists of 34 items. Twenty-six of these items are pairs of contrasting statements that participants rated on a 5-point scale (e.g., 1 = left is really true, 5 = right is really true). The other eight items are individual statements or activities that participants rated on a 5-point scale reflecting their degree of agreement, importance, or descriptiveness.

A subsample of participants (N = 249) also completed the tripartite taxonomy of character (TTC, Park et al., 2017), the social and emotional competency assessment (SECA, Davidson et al., 2018), and social-emotional learning surveys (SELS, West et al., 2018). The TTC consists of 24 items and assesses interpersonal, intrapersonal, and intellectual strengths on a 7-point frequency scale (e.g., 1 = almost never, 7 = almost always). The SECA assesses self-management, social awareness, relationship skills, and responsible decision-making across 40 items on a 4-point scale indicating how difficult or easy it is for the respondent to enact that skill. The SELS uses 18 items measuring self-management, social awareness, self-efficacy, and growth mindset. All items are on a 5-point scale gauging either frequency, degree of confidence, agreement, or an item-specific outcome. For further details, see Soto et al. (2022a, Study 4).

High School Student Developmental Outcomes. All participants in the High School Student sample completed a battery of items assessing academic engagement, occupational interests, peer acceptance, friendship quality, mother and father relationship quality, volunteering, exercise, and satisfaction with life. Academic engagement was assessed using a 12-item version of the engagement versus disaffection with learning measure (EDL, Skinner et al., 2008). Occupational interests were assessed using the O*Net Mini Interest Profiler (Mini-IP, Rounds et al., 2016). Peer acceptance was assessed using two items: One item asked participants to rate their level of popularity among their peers on a 9-point scale (Paunonen, 2003), and the other item asked participants to rate their level of status within their social group on a 9-point scale (Anderson et al., 2001). Friendship quality with the participants' best friend was assessed using a 15-item version of the Friendship Qualities Scale (FQS, Bukowski et al., 1994). The quality of participants' relationships with their parents was assessed using six items adapted from the Dunedin Study of lifespan development (Belsky et al., 2001). Participants responded to four items assessing their current, past, and prospective volunteering activities (Carlo et al., 2005). Participants answered three items from the Godin Leisure-Time Exercise Questionnaire (GLTEQ) indicating the number of days in a typical week that they engage in strenuous, moderate, and mild physical exercise (Godin & Shephard, 1985). Finally, participants completed the 5-item Satisfaction With Life Scale (Diener et al., 1985). In addition, school-reported overall grade point averages for winter and spring courses were available for a subsample of participants (N = 474). For further details, see Soto et al. (2022a, Study 5).

Panel Occupational and Well-Being Outcomes. All participants in the Panel sample completed a series of items that assessed self-rated health, life satisfaction, feelings of burnout, work satisfaction, and income. Self-rated health was assessed by a single item: "All in all, how would you rate your state of health?" (1 = very good, 5 = very bad). Life satisfaction was also assessed by a single item: "All things considered, how satisfied are you at present with your life?" (1 = very dissatisfied, 11 = completely satisfied). Feelings of personal burnout were assessed by an adapted version of the six-item Copenhagen Burnout Inventory subscale (Kristensen et al., 2005). An example item for this subscale is "How often do you feel tired?" (1 = never, 5 = almost every day). Work satisfaction was measured by a single item: "How satisfied are you with your current professional activity/training/studies?" (1 = completely satisfied, 7 = completely dissatisfied). Income was assessed by asking participants, "What is your own monthly net income? (Employee: total after deduction of taxes and social security contributions; Self-employed: average net income less business expenses)." Responses ranged from 1 = under 300 euros to 17 = 10,000 euros or more. Participants could also indicate no "own" income.

Transparency and Openness

The High School Student, College Student, Internet 1, Internet 2, and Observer datasets are available at https://osf.io/4zgyr/. The Panel dataset is available at https://osf.io/9pvmj/. The item selection criteria, predictive validity quality control criteria, and analysis plan were preregistered at https://osf.io/2y7ua/?view_only=9cd00f088c7149abac50a7d5ceb5de18. All OSTs are available at https://osf.io/2y7ua/?view_only=9cd00f088c7149abac50a7d5ceb5de18. The BESSI short form items and scoring, the syntax for CFA and exploratory structural equation models (ESEM) analyses, and other materials are available at https://osf.io/2y7ua/?view_only=9cd00f088c7149abac50a7d5ceb5de18.

Results

Developing BESSI-96, BESSI-45, and BESSI-20 Short Forms

We selected items for the BESSI short forms using a diverse set of conceptual and empirical criteria in each of the four selection samples: the High School Student, College Student, Internet 1, and Observer samples. To promote coherence within each skill domain and facet, we computed each item's (a) corrected item-total correlation with its domain or facet scale, (b) standardized loading from a single-factor domain or facet CFA model, and (c) discrimination parameter from a two-parameter logistic item response theory (2PL IRT) graded response model of its domain or facet. To promote differentiation between domains and facets, we computed each item's (d) ratio of its corrected item-total correlation to the maximum correlation with another domain or facet scale. To avoid restricted range and ceiling effects, we computed each item's (e) closeness of its mean response to the scale midpoint, (f) standard deviation of responses, (g) percentage of responses that fell below the maximum response option, and (h) item difficulty parameter from the 2PL IRT model. Finally, we made expert judgments to (i) avoid overly redundant item content and wording, (j) select for easy-to-understand items, (k) select for item content developmentally relevant across childhood, adolescence, and adulthood (l) select for item content appropriate for observer-reports, and (m) avoid idiomatic expressions (thereby facilitating adaptation across languages and cultures).

Reflecting the different goals of the BESSI short forms, we weighted domain and facet-level selection criteria somewhat differently across forms. Because the BESSI-96 is intended to efficiently assess all 32 skill facets and five domains, we selected three items per facet with equal attention to domain-level and facet-level characteristics. Because the BESSI-45 is intended to efficiently assess the five skill domains, we selected nine items per domain with greater attention to domain-level than facet-level characteristics. For the Self-Management domain, we selected one item from each of its nine core facets; for each of the other four skill domains, we selected two items from each of its four core facets plus one item from its interstitial facet. Because the BESSI-20 is intended to assess the five skill domains very efficiently, we selected four items per domain, again with greater attention to domain-level characteristics. To preserve as much content breadth as possible, for each domain, we selected one item from four different core facets.

To maximize comparability across forms, we first selected items for the BESSI-96 from the full-length BESSI-192. We then selected items for the BESSI-45 from the BESSI-96 and finally selected items for the BESSI-20 from the BESSI-45. The complete set of items selected for the BESSI-96, BESSI-45, and BESSI-20 are provided in the Appendix.

Measurement Properties of the BESSI Short Forms

Do the Short Forms Adequately Capture BESSI-192 Content? Next, we investigated the short forms' measurement properties in the four item-selection samples, as well as the two independent validation samples: the Internet 2 sample and the Panel sample. We began by computing part–whole correlations in which the BESSI domain scales from each short form were correlated with the full measure. We also computed part–whole correlations between the BESSI-96 and BESSI-192 facets. These correlations allowed us to test the degree to which the short forms captured the content of the full BESSI-192 measure (see OSTs S5 and S6 for complete matrices).

Across all six samples, part–whole correlations between BESSI-96 and BESSI-192 domain pairs averaged .98 (range across samples = .96–.99). Part–whole correlations between BESSI-96 and BESSI-192 facet pairs averaged .96 (range = .94–.97). For BESSI-45, part–whole correlations between domain pairs averaged .96 (range = .94–.98). Finally, for BESSI-20, part–whole correlations between domain pairs averaged .91 (range = .87–.95).

Do the Short Forms Retain the Facet and Domain Structure of BESSI-192? Next, we assessed whether the short forms retained the BESSI-192's multidimensional structure using a series of CFAs for strict tests of the short forms' facet and domain structures and ESEMs to further explore domain structure while allowing for greater complexity. Fit statistics for these models are provided in Table 2, and complete parameter estimates, including loadings and intercorrelations, for all samples and models are provided in OSTs S7 to S14. All items were defined as ordinal indicators, and models were fit using the means and variance-adjusted weighted least squares (WLSMV) estimator in Mplus 8 or the lavaan package in *R* (Muthén & Muthén, 2017; Rosseel, 2012). For all models, we used conventional standards to determine whether model fit was adequate (comparative fit index [CFI] and Tucker–Lewis index [TLI] $\geqslant$ 0.90, root-mean-square error of approximation [RMSEA], and standardized root mean square residual [SRMR] $\leqslant$ 0.08) or good (CFI and TLI $\geqslant$ 0.95, RMSEA and SRMR $\leqslant$ 0.05; Barrett, 2007). We began by investigating the loadings of the short form items onto their

Table 2. Fit Statistics for CFAs and ESEMs of the BESSI Short Forms' Facet and Domain Structure.

Form	Model	Sample	χ^2	*df*	CFI	TLI	RMSEA	SRMR
BESSI-96	32 Facet CFA	Internet 1	6,171	3,968	0.92	0.90	0.04	0.06
BESSI-96	Higher-order CFA	Internet 1	9,301	4,406	0.81	0.80	0.05	0.10
BESSI-96	Bifactor ESEM	Internet 1	5,176	3,994	0.93	0.92	0.03	0.05
BESSI-96	32 Facet CFA	HS.	6,193	3,968	0.95	0.94	0.03	0.05
BESSI-96	Higher-order CFA	HS.	8,311	4,406	0.91	0.91	0.04	0.07
BESSI-96	Bifactor ESEM	HS.	5,806	3,994	0.96	0.95	0.03	0.04
BESSI-96	32 Facet CFA	College	5,707	3,968	0.94	0.93	0.04	0.06
BESSI-96	Higher-order CFA	College	7,415	4,406	0.89	0.89	0.05	0.08
BESSI-96	Bifactor ESEM	College	5,242	3,997	0.95	0.95	0.03	0.04
BESSI-96	32 Facet CFA	Observer	6,840	3,968	0.96	0.96	0.04	0.04
BESSI-96	Higher-order CFA	Observer	9,746	4,406	0.93	0.93	0.05	0.06
BESSI-96	Bifactor ESEM	Observer	6,148	3,994	0.97	0.97	0.03	0.03
BESSI-96	32 Facet CFA	Internet 2	7,778	3,968	0.91	0.90	0.04	0.06
BESSI-96	Higher-order CFA	Internet 2	12,582	4,406	0.81	0.80	0.06	0.09
BESSI-96	Bifactor ESEM	Internet 2	6,713	3,998	0.94	0.93	0.03	0.04
BESSI-96	32 Facet CFA	Panel	6,036	3,968	0.96	0.96	0.03	0.04
BESSI-96	Higher-order CFA	Panel	8,129	4,406	0.93	0.93	0.04	0.06
BESSI-96	Bifactor ESEM[a]	Panel	5,950	3,994	0.96	0.96	0.03	0.04
BESSI-45	5 Domain CFA	Internet 1	3,094	915	0.82	0.81	0.08	0.10
BESSI-45	Targeted ESEM	Internet 1	1,405	759	0.95	0.93	0.05	0.04
BESSI-45	5 Domain CFA	HS.	2,628	915	0.91	0.90	0.06	0.07
BESSI-45	Targeted ESEM	HS.	1,494	759	0.96	0.95	0.04	0.03
BESSI-45	5 Domain CFA	College	2,129	915	0.90	0.90	0.06	0.08
BESSI-45	Targeted ESEM	College	1,262	759	0.97	0.96	0.04	0.03
BESSI-45	5 Domain CFA	Observer	3,341	915	0.93	0.92	0.07	0.06
BESSI-45	Targeted ESEM	Observer	1,611	759	0.97	0.97	0.05	0.03
BESSI-45	5 Domain CFA	Internet 2	4,299	915	0.83	0.81	0.08	0.09
BESSI-45	Targeted ESEM	Internet 2	1,742	759	0.95	0.93	0.05	0.04
BESSI-45	5 Domain CFA	Panel	2,343	915	0.95	0.94	0.06	0.06
BESSI-45	Targeted ESEM	Panel	1,477	759	0.97	0.97	0.04	0.03
BESSI-20	5 Domain CFA	Internet 1	758	160	0.82	0.79	0.10	0.09
BESSI-20	Targeted ESEM	Internet 1	204	100	0.97	0.94	0.05	0.02
BESSI-20	5 Domain CFA	HS.	601	160	0.93	0.92	0.07	0.06
BESSI-20	Targeted ESEM	HS.	241	100	0.98	0.96	0.05	0.02
BESSI-20	5 Domain CFA	College	537	160	0.91	0.89	0.09	0.07
BESSI-20	Targeted ESEM	College	174	100	0.99	0.97	0.04	0.02
BESSI-20	5 Domain CFA	Observer	853	160	0.94	0.93	0.09	0.06
BESSI-20	Targeted ESEM	Observer	317	100	0.98	0.97	0.07	0.02
BESSI-20	5 Domain CFA	Internet 2	1,067	160	0.83	0.79	0.10	0.08
BESSI-20	Targeted ESEM	Internet 2	323	100	0.96	0.92	0.06	0.03
BESSI-20	5 Domain CFA	Panel	546	160	0.95	0.94	0.07	0.05
BESSI-20	Targeted ESEM	Panel	258	100	0.98	0.96	0.06	0.02

Note. HS = high school; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root-mean-square error of approximation; SRMR = standardized root mean square residual. Degrees of freedom differ across bifactor ESEM models because some facet item loadings were constrained to equality to address item negative residual variance.
[a]The German Panel bifactor ESEM includes 31 facet factors and five exploratory domain ESEM.

respective domains and facets as well as domain factor intercorrelations. Then, we inspected the model fit.

BESSI-96. To investigate whether the BESSI-96 retains the facet structure of the BESSI-192, we specified a CFA model in which each BESSI-96 item was allowed to load on its respective facet factor, and facet factors were allowed to intercorrelate. Across all six samples, items loaded very strongly on their assigned facet (M = 0.82, range = 0.58–0.94). This model demonstrated adequate to good fit across all samples (CFI = 0.91–0.96, TLI = 0.90–0.96, RMSEA = 0.03–0.04, SRMR = 0.04–0.06).

To investigate whether the BESSI-96 retained the domain structure of the BESSI-192, we first tested a higher-order five-domain CFA which nested the 32 BESSI facets within five correlated domain factors. Each BESSI-96 item was allowed to load onto its respective facet factor, and each facet factor was allowed to load onto one or more domain factors (see Soto et al.,

2022a). Specifically, only facets defined as interstitial (energy regulation, ethical competence, impulse regulation, and information processing skill) or compound (self-reflection skill, capacity for independence, and adaptability) were allowed to load onto more than one domain. Across all six samples, single-domain facets loaded strongly on their assigned domain factors ($M_{\text{self-management}}$ = 0.66, $M_{\text{social engagement}}$ = 0.65, $M_{\text{cooperation}}$ = 0.65, $M_{\text{emotional resilience}}$ = 0.67, $M_{\text{innovation}}$ = 0.57), and all single-domain facets demonstrated meaningful pooled loadings (range = 0.34–0.80). Interstitial facets demonstrated moderate loadings across their assigned domains (M = 0.35, range = 0.27–0.46). Two compound facets—adaptability and capacity for independence—demonstrated pooled loadings above 0.20 on three domains, while self-reflection had pooled loadings above 0.20 on two domains. Domain intercorrelations ranged from moderate in Internet 1 (M = .43, range within sample = .21–.61) to strong in the Observer sample (M = .77, range within sample = .77–.89), indicating greater differentiation between skill domains in self-ratings than in observer-ratings. Similar to the BESSI-192 (Soto et al., 2022a), fit statistics across samples ranged from adequate (e.g., $\text{CFI}_{\text{Panel}}$ = .93, $\text{TLI}_{\text{Panel}}$ = .93, $\text{RMSEA}_{\text{Panel}}$ = .04, $\text{SRMR}_{\text{Panel}}$ = .06) to inadequate (e.g., $\text{CFI}_{\text{Internet1}}$ = .81, $\text{TLI}_{\text{Internet1}}$ = .80, $\text{RMSEA}_{\text{Internet1}}$ = .05, $\text{SRMR}_{\text{Internet1}}$ = .10).

To explore unmodeled complexity, we next tested a bifactor ESEM that (a) allowed indicators to load onto multiple factors at different levels of abstraction (e.g., both skill facets and skill domains), and (b) combined a confirmatory approach at the facet level with an exploratory approach at the domain level. In this bifactor ESEM, each BESSI-96 item is loaded directly onto its facet factor and the five exploratory domain factors. To approximate the domain-level structure of the confirmatory model, item loadings on domain factors were targeted at zero except for loadings on each item's assigned domains. The five exploratory domain factors were allowed to intercorrelate, but the 32 facet factors[1] were constrained to be independent of other facets and the domains.

On average, most BESSI-96 items demonstrated meaningful pooled loadings on their assigned facets (M = 0.42, range = 0.07–0.83) and their targeted domain factor or factors (M = 0.41).[2] Single-domain items demonstrated the strongest pooled loadings (M = 0.58, range = 0.33–0.78) while, as expected, interstitial items (M = 0.36, range = 0.17–0.53) and compound items (M = 0.17, range = −0.19 to 0.61) demonstrated weaker loadings. Most items demonstrated insignificant pooled loadings on the other domains (M = 0.05), but some items demonstrated meaningful pooled cross-loadings (range = −0.24 to 0.44). Domain factor intercorrelations ranged from weak in the internet samples ($M_{\text{Internet1 \& 2}}$ = .17, range = .05–.30) to moderate in the Observer sample (M = .46, range within sample = .28–.62), indicating better differentiation between skill domains than in the strict CFA model. As shown in Table 2, the bifactor ESEM models provided consistently better fit than the higher-order CFA models (ΔCFI = .04–.12, ΔTLI = .04–.12, ΔRMSEA = .01–.03, ΔSRMR = .03–.05) and demonstrated good overall fit across all samples.

BESSI-45. We first tested a five-domain CFA model with cross-loadings and correlated residuals to represent the BESSI facets. In this model, each single-domain BESSI-45 item is loaded directly onto its respective domain factor, and interstitial BESSI-45 items loaded onto two domain factors. This model also included correlated residuals between items belonging to the same skill facet. Across all six samples, single-domain items loaded strongly on their assigned factors ($M_{\text{self-management}}$ = 0.68, $M_{\text{social engagement}}$ = 0.64, $M_{\text{cooperation}}$ = 0.68, $M_{\text{emotional resilience}}$ = 0.72, $M_{\text{innovation}}$ = 0.59), and almost all single-domain items demonstrated meaningful pooled loadings (range = 0.28–0.84). Interstitial items demonstrated moderate loadings across their assigned domains (M = 0.37, range = 0.19–0.61). As with BESSI-96, domain intercorrelations ranged from moderate in Internet 1 (M = .41, range within sample = .15–.62) to strong in the Observer sample (M = .76, range within sample .67–.90). This model provided adequate fit in four samples (see Table 2), but inadequate fit in the two internet samples (e.g., CFI = .82–.83, TLI = .81, RMSEA = .08, SRMR = .09–.10).

Therefore, we also tested an ESEM model with five exploratory domain factors. In this model, we targeted BESSI-45 item loadings at zero on all domains except for the items' assigned domains. Most items loaded meaningfully on their targeted domain factor or factors (M = 0.54) with single-domain items demonstrating stronger pooled loadings (range = 0.37–0.77) than interstitial items (range = 0.16–0.56). Most items demonstrated insignificant pooled loadings on the other domains (M = 0.05), but some showed more substantial cross-loadings (range = −0.24–0.47). Domain factor intercorrelations ranged from weak in the internet samples ($M_{\text{Internet1 \& 2}}$ = .19, range = −.01 to .32) to moderate in the Observer sample (M = .48, range within sample = .33–.65). As demonstrated in Table 2, the targeted ESEM model significantly improved model fit over the strict confirmatory model (ΔCFI = 0.02–0.13, ΔTLI = 0.03–0.12, ΔRMSEA = 0.02–0.03, ΔSRMR = 0.03–0.06) and provided good fit across all datasets.[3]

Table 3. Internal Consistency of BESSI Short Form Skill Domains.

	Alpha reliability					Omega reliability				
Form/sample	SM	SE	CO	ER	IN	SM	SE	CO	ER	IN
BESSI-96										
Internet 1	.93	.89	.89	.90	.85	.92	.79	.79	.81	.87
High School	.94	.89	.88	.92	.86	.96	.88	.74	.88	.86
College	.94	.91	.89	.91	.87	.95	.84	.92	.84	.88
Observer	.97	.92	.95	.95	.93	.98	.89	.93	.92	.91
Internet 2	.93	.87	.87	.90	.84	.92	.79	.73	.82	.85
Panel	.96	.94	.92	.95	.92	.94	.88	.79	.91	.89
M	.95	.90	.90	.92	.88	.95	.85	.82	.87	.88
BESSI-45										
Internet 1	.82	.84	.85	.86	.81	.84	.85	.87	.87	.84
High School	.88	.84	.83	.88	.79	.88	.85	.84	.89	.79
College	.86	.86	.85	.87	.81	.87	.87	.87	.88	.83
Observer	.93	.86	.93	.92	.89	.94	.87	.94	.93	.90
Internet 2	.81	.81	.80	.86	.75	.84	.81	.82	.86	.77
Panel	.86	.91	.88	.92	.89	.88	.92	.89	.93	.90
M	.86	.85	.86	.89	.82	.88	.86	.87	.89	.84
BESSI-20										
Internet 1	.75	.71	.76	.72	.66	.79	.76	.81	.77	.76
High School	.82	.69	.75	.75	.59	.85	.73	.80	.80	.64
College	.80	.74	.75	.75	.68	.84	.77	.80	.80	.75
Observer	.88	.72	.87	.84	.78	.91	.74	.90	.88	.82
Internet 2	.78	.62	.75	.72	.59	.82	.67	.79	.76	.67
Panel	.78	.82	.77	.84	.78	.82	.85	.82	.87	.82
M	.80	.72	.78	.77	.68	.84	.75	.82	.81	.74

Note. SM = self-management skills; SE = social engagement skills; CO = cooperation skills; ER = emotional resilience skills; IN = innovation skills.

BESSI-20. We first tested a five-domain CFA model in which each BESSI-20 item was loaded directly onto its respective domain. Across all six samples, BESSI-20 items loaded strongly on their assigned domain factors ($M_{\text{self-management}} = 0.75$, $M_{\text{social engagement}} = 0.66$, $M_{\text{cooperation}} = 0.73$, $M_{\text{emotional resilience}} = 0.72$, $M_{\text{innovation}} = 0.65$), and all items had meaningful loadings (range = 0.46–0.88). Domain intercorrelations ranged from small in Internet 1 ($M_{\text{Internet1}} = .39$, range within sample = .21–.59) to strong in the Observer sample ($M = .75$, range within sample = .69–.86). As shown in Table 2, this model provided adequate fit for four samples and inadequate fit in the two internet samples (e.g., CFI = .82–.83, TLI = .79, RMSEA = .10, SRMR = .08–.09).

We therefore tested a targeted ESEM similar to the ESEM for the BESSI-45. In this model, each item's loadings were targeted at zero for four of the five domain factors and freely estimated for its assigned domain. Items loaded substantially on their targeted domain factor ($M = 0.63$), and all items' pooled loadings were meaningful (range = 0.42–0.76). Most items demonstrated insignificant loadings on the other domains, though a few items had small cross-loadings ($M = 0.04$, range = −0.22 to 0.27). Domain intercorrelations ranged from weak in Internet 1 ($M = .21$, range within sample = .10–.32) to moderate in the Panel sample ($M = .50$, range within sample = .35–.62). The targeted ESEM model significantly improved model fit over the strict confirmatory model (ΔCFI = .03–.15, ΔTLI = .02–.15, ΔRMSEA = .01–.05, ΔSRMR = .03–.07) and had good fit across all datasets (see Table 2).

Are the Short Forms' Skill Domains and Facets Reliable? Finally, we assessed the internal consistency and retest reliability of the short forms' skill domains and facets. Table 3 presents both alpha (calculated from observed scores) and omega (calculated from the best-fitting CFA models) reliability coefficients for the five skill domains across samples, and OST S15 presents alpha and omega reliabilities for the BESSI-96 facets. For the BESSI-96 skill domains, the pooled Cronbach's alpha across the six samples was .91, and the pooled Omega was .87. For BESSI-45, the pooled alpha was .85, and Omega was .87. For BESSI-20, the pooled alpha was .75, and Omega was .79. The pooled Cronbach's alpha of the BESSI-96 facets was .82, and the pooled Omega was .86.

Because a subset of the College Student sample completed the BESSI-192 a second time, we also calculated the test–retest reliability for the short forms' domain and facet scales by correlating each pair of domain or facet scale scores. For the domain scales, the retest reliability averaged .80 for BESSI-96 (range = .79–.81), .79 for BESSI-45 (range = .75–.81), and .75 for BESSI-20 (range = .74–.78). For the BESSI-96 facet scales, retest reliability averaged .71 (range = .56–.83). Retest reliability reflects measurement error (unreliability), true change in the construct, and state deviations. Thus, there are no universally accepted cut-offs for interpreting retest reliability. However, these domain and facet retest reliabilities are similar to those reported for the BFI-2 short forms' trait domains across a 3-month interval (Soto & John, 2017) and the BESSI-192 facet stabilities across a 6-week and 8-month interval (Lechner et al., 2022). Taken together, the present findings indicate that the BESSI short forms' domains and facets demonstrate substantial but not perfect stability across time.

Summary. In summary, our rational–empirical item selection process yielded three short forms that largely captured the psychological content of the full BESSI-192 measure. In addition, across analyses, the BESSI short forms retained the full measure's facet and domain structure. Across all six samples, items loaded substantially on their respective facet and domain factors. In addition, domain intercorrelations did not vary much across short forms, indicating that reducing the number of items does not impact discriminability between domains. However, like the BESSI-192, a strictly confirmatory five-factor approach to this structure may be ill-suited for some samples and a more flexible structure may be necessary to concurrently model all five domains. This finding reflects the broad, inclusive nature of the BESSI. Rather than narrowly defining domains to produce well-fitting models, we prioritized domain construct breadth during the short form item selection process, and we also intentionally included interstitial and compound facets and items. This inclusive approach means that fairly complex models are needed to adequately capture the BESSI's domain-level structure (Soto et al., 2022a). Finally, the short forms demonstrated adequate internal consistency at the domain and facet levels, as well as moderate levels of stability across a 4-month time span. Our findings also suggest that the BESSI-96 and BESSI-45 provide greater construct coverage and reliability than the BESSI-20.

Predictive Validity of the BESSI Short Forms

Are Skill-Outcome Associations Similar Between the BESSI-192 and the Short Forms? To assess whether the BESSI short forms retained the validity of the full measure when predicting consequential outcomes, we first computed partial correlations between the BESSI-192, BESSI-96, BESSI-45, and BESSI-20 skill domains and the 43 total self-reported or school-reported criteria in the U.S. College, U.S. High School, and German Panel samples, controlling for participant gender and grade level or age. Next, we conducted multiple regression analyses in which each criterion was regressed on each set of domain scales, as well as gender and age or grade level. To parallel domain-level analyses, we also computed partial correlations of BESSI-192 and BESSI-96 facets with the 43 outcomes and conducted multiple regression analyses in which each criterion was regressed on each set of 32 facet scales.

We expected that patterns of criterion associations for the BESSI short forms would be similar to the full measure, with column-vector correlations of .50 or higher (see OST S16 for a complete list of coefficients). Supporting this hypothesis, column-vector correlations comparing the domain-criterion partial correlations for BESSI-192 and BESSI-96 were at least .86 (M = .99) for each of the 43 criteria. When comparing patterns of standardized regression coefficients, all column-vector correlation coefficients were at least .88 (M = .99). For comparisons between BESSI-192 and BESSI-45, column-vector correlations of the partial correlations were at least .50 (M = .96) for each criterion, and at least .68 (M = .97) for the regression coefficients. Column-vector correlations comparing the partial correlations for BESSI-192 against BESSI-20 were at least .76 (M = .95) for each criterion, and at least .79 (M = .95) for the regression coefficients. Finally, for comparisons between the BESSI-192 facets and BESSI-96 facets, column-vector correlations of the partial correlations were at least .92 (M = .97), and at least .60 (M = .88) for regression coefficients. Thus, column-vector correlations exceeded our .50 threshold in every case, with most correlations at .90 or higher. These results indicate that when a particular criterion variable has stronger associations with some BESSI-192 skill domains or facets than others, these differential patterns also tend to hold for the short forms. This finding suggests that the BESSI short forms' skill domains and facets exhibit discriminant validity similar to the BESSI-192.

How Well Do the BESSI Short Forms Retain the Predictive Power of the BESSI-192? To assess whether the short forms retain the predictive power of the BESSI-192, we examined adjusted R^2 values from the regression models. Table 4 presents these adjusted R^2 values for domain- and facet-level analyses. On average, the BESSI-192 domains explained 30.1%, BESSI-96 domains explained 28.6%, BESSI-45 domains explained 27.6%, and

Table 4. Adjusted R^2 Values From BESSI Domain and Facet Scales Predicting Outcomes.

Criterion	BESSI-20 domains	BESSI-45 domains	BESSI-96 domains	BESSI-192 domains	BESSI-96 facets	BESSI-192 facets
College student						
PYD competence	.16	.15	.15	.16	.22	.25
PYD confidence	.35	.34	.35	.35	.47	.46
PYD character	.28	.28	.30	.33	.38	.43
PYD caring	.30	.34	.31	.34	.33	.34
PYD connection	.22	.20	.20	.21	.31	.27
BFI-2 conscientiousness	.51	.59	.63	.63	.72	.71
BFI-2 extraversion	.57	.65	.65	.67	.74	.76
BFI-2 agreeableness	.49	.54	.55	.58	.64	.67
BFI-2 negative emotionality	.62	.62	.65	.69	.70	.73
BFI-2 open-mindedness	.54	.50	.59	.64	.61	.69
TTC interpersonal strengths	.46	.49	.49	.52	.53	.55
TTC intrapersonal strengths	.54	.54	.57	.60	.63	.66
TTC intellectual strengths	.51	.50	.50	.51	.59	.59
SECA self-awareness	.36	.38	.38	.40	.55	.55
SECA self-management	.54	.51	.52	.54	.58	.58
SECA social awareness	.34	.34	.37	.39	.38	.41
SECA relationship skills	.45	.43	.44	.45	.48	.51
SECA responsible decision-making	.25	.26	.28	.30	.35	.38
SELS self-management	.42	.41	.40	.43	.53	.56
SELS social awareness	.39	.41	.41	.42	.46	.45
SELS self-efficacy	.25	.26	.26	.27	.32	.34
SELS growth mindset	.24	.25	.24	.26	.29	.31
High school student						
Winter GPA	.07	.08	.09	.11	.21	.23
Spring GPA	.02	.03	.04	.06	.13	.14
Academic engagement	.32	.32	.35	.38	.44	.46
Realistic interests	.07	.07	.08	.08	.10	.11
Investigative interests	.04	.05	.04	.04	.05	.05
Artistic interests	.20	.20	.23	.25	.23	.30
Social interests	.19	.21	.21	.21	.25	.23
Enterprising interests	.12	.12	.13	.13	.15	.17
Conventional interests	.04	.04	.04	.04	.04	.06
Peer acceptance	.19	.23	.24	.25	.28	.30
Friendship quality	.20	.20	.21	.22	.26	.29
Mother quality	.08	.08	.08	.09	.10	.13
Father quality	.08	.10	.09	.11	.15	.17
Volunteerism	.09	.11	.12	.12	.18	.21
Exercise	.10	.09	.11	.12	.11	.14
Life satisfaction[a]	.22	.21	.22	.23	.26	.28
Panel						
Life satisfaction[a]	.16	.15	.16	.16	.26	.24
Job satisfaction	.16	.17	.17	.17	.20	.21
Health	.20	.19	.20	.20	.25	.25
Income	.05	.04	.05	.05	.08	.08
Burnout	.20	.20	.22	.22	.27	.28

Note. PYD = positive youth development; BFI-2 = Big Five Inventory-2; TTC = tripartite taxonomy of character; SECA = social and emotional competency assessment; SELS = social-emotional learning surveys. All adjusted R^2 values have a p-value of $< .001$. All models included gender and age or grade level as covariates.

[a]The High School sample used a five-item measure of life satisfaction (Diener et al., 1985), and the Panel sample used a single-item measure of life satisfaction.

BESSI-20 domains explained 26.9% of the variance in outcomes. At the facet level, the BESSI-192 explained 36.1%, and the BESSI-96 explained 34.4% of the variance in outcomes, on average.

Next, we calculated the proportion of the BESSI-192's validity retained by each short form's domain and facet scales when predicting outcomes. These proportions from domain- and facet-level analyses are

Table 5. Fit Statistics Comparing Configural, Metric, and Scalar Models Between Internet 1 and Panel Samples.

	BESSI-96 Semi-higher-order CFA				BESSI-45 Five-domain CFA				BESSI-20 Five-domain CFA			
Fit statistics	Con.	Met.	Sca. F.O.	Sca. S.O.	Con.	Met.	F. Sca.	P. Sca.	Con.	Met.	F. Sca.	P. Sca.
χ^2	17,466	17,288	17,961	18,328	5,511	5,433	6,000	5,947	1,325	1293	1,497	1,442
df	8,806	8,903	9,154	9,181	1,830	1,874	2,004	2,003	320	335	390	387
CFI	.898	.901	.896	.892	.912	.915	.905	.906	.917	.921	.908	.913
RMSEA	.047	.046	.047	.048	.068	.066	.067	.067	.084	.081	.080	.079
SRMR	.075	.082	.076	.076	.074	.079	.076	.076	.065	.068	.066	.066
$\Delta\chi^2$		172*	348*	292*		222*	155	63		59*	179*	112
Δdf		97	251	27		44	130	129		15	55	52
ΔCFI		.003	.005	.004		.003	.010	.009		.004	.013	.008
ΔRMSEA		.001	.006	.001		.002	.001	.001		.003	.001	.002

Note. CFI = comparative fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root mean square residual; Con. = Configural; Met. = Metric; Sca. = Scalar; F.O. = scalar model for first order (facets) latent factors; S. O. = scalar model for second order (domains) latent factors; F. = full; P. = partial.
*$p < .05$.

presented in OST S17. We expected that the short forms would retain at least 70% of the BESSI-192's predictive power. Supporting this hypothesis, on average, the BESSI-96 domains retained approximately 94%, BESSI-45 domains retained approximately 91%, and BESSI-20 domains retained approximately 88% of the predictive power of the BESSI-192 domains. Similarly, the BESSI-96 facets retained approximately 93% of the predictive power of the BESSI-192 facets. Moreover, the explained variance ratios exceeded our preregistered 70% benchmark for the vast majority of individual cases (167 of 172 possible combinations of a short form scale with an outcome), with most ratios exceeding 90%.

Summary. Our predictive validity analyses indicated that the BESSI short forms demonstrated patterns of associations similar to the BESSI-192 and retained most of BESSI-192's predictive power. These results also suggest a gradual decrease in predictive power as the BESSI short forms lose items and, thus, construct coverage and information. However, with only a 3% loss in predictive power, the BESSI-20 still offers a strong alternative to the BESSI-192 in terms of validity. In summary, these findings suggest that in many cases the short forms can serve as an adequate proxy for BESSI-192.

Cross-Cultural Comparability of the BESSI Short Forms

Finally, to investigate whether the short forms are comparable between two cultural contexts, we conducted multigroup measurement invariance analyses using the German Panel sample and the Internet 1 sample because these samples were the most directly comparable in that they were both self-report, adult samples. We started by utilizing the best-fitting CFA model[4] for each short form. This baseline, configural measurement model had no equality constraints between groups. The configural models provide a basis for comparison to the metric measurement models. The metric measurement model constrained all item loadings to equality between the two groups and served as the comparison for the scalar measurement models. The scalar measurement model constrained all item loadings and additionally constrained item thresholds to equality between the two groups. Following recommendations by Cheung and Rensvold (2002), a ΔCFI smaller than .01 between the models indicates measurement invariance.

Table 5 presents the goodness-of-fit indices for the configural, metric, and scalar models across BESSI short forms. These results show that ΔCFI was smaller than .01 between these models for the BESSI-96, thus indicating similar structure and scaling at the item, facet,

and domain levels across the U.S. and German cultural contexts. However, for the BESSI-45 and BESSI-20, full scalar invariance could not be established. Modification indices indicated that the three thresholds most contributing to misfit for BESSI-45 all belonged to the Innovation item "invent things," and freeing one threshold allowed us to establish partial scalar measurement invariance. This finding indicates that while the structure of SEB skills and item loadings onto domains may be comparable across U.S. and German contexts, participants' interpretations and ratings of "invent things" differ, and thus could inflate or suppress mean level differences in Innovation skills. For BESSI-20, a threshold for the Innovation item "come up with new ideas" and two thresholds for the Self-Management item "work towards my goals" contributed to misfit. Freeing these three thresholds allowed us to establish partial scalar measurement invariance. Similar to the BESSI-45, the BESSI-20 domain scores are largely comparable between these two cultures, but participants' interpretations and ratings of these two specific items differed and could bias mean-level comparisons of Self-Management and Innovation skills.

Discussion

The burgeoning interest in understanding, measuring, and intervening on SEB skills in research and applied contexts necessitates inventories that address the breadth of possible assessment goals for researchers and practitioners as well as their assessment constraints. This research sought to develop reliable and valid short forms of the BESSI-192, a measure of 32 skill facets and five skill domains, and to investigate the short forms' cross-cultural comparability. Using a combination of rational and empirical criteria, we developed 96-item, 45-item, and 20-item BESSI short forms that preserved the full-length BESSI-192's strong psychometric characteristics. These short forms largely captured the psychological content of the full measure and demonstrated adequate internal consistency and retest stability at both the domain and facet levels.

The short forms also retained the multidimensional structure of BESSI-192. At the facet level, the BESSI-96 demonstrated a clear structure with the 32-facet, fully confirmatory model demonstrating at least adequate fit across all samples. An important finding from our analyses is that, like the Big Five (Wetzel & Roberts, 2020), the complex, multidimensional domain structure of SEB skills may not be adequately captured by strict confirmatory models. From both a conceptual and measurement standpoint, SEB skill domains are not best represented as independent constructs but instead overlap to some degree (Soto et al., 2022a), and we prioritized domain construct coverage over defining independent but narrow domains. Thus, relatively complex structural models are needed to adequately account for the relations between SEB skill facets and domains. Despite this complexity, discriminability across domains stayed consistent across forms, even with substantial reductions in the number of items.

In addition to exploring their measurement properties, we tested the short forms' criterion validity and found that they demonstrated patterns of skill-outcome associations similar to the BESSI-192 and preserved a large proportion of the BESSI-192's predictive validity. Previous studies have shown that each BESSI-192 skill domain has distinctive relations with consequential outcomes and criterion variables (Lechner et al., 2022; Soto et al., 2022a), and the present findings show that the BESSI short forms retain these distinctive skill-outcome relations.

Finally, measurement invariance analyses indicated that the BESSI-96 showed equivalent structure and scaling between a predominantly U.S.-based sample and a German sample. For the BESSI-45 and BESSI-20, partial scalar invariance could be established by freeing only one to three specific item thresholds. This latter finding may suggest that a few BESSI items are interpreted differently across cultures. However, we are hesitant to overinterpret these findings because the samples analyzed here differed in other ways besides culture (e.g., demographic characteristics and recruitment strategy). Taken as a whole, the present findings indicate that the BESSI short forms are reliable and valid measures of SEB skills and provide promising initial evidence for their cross-cultural comparability.

In What Contexts Should Short Forms Be Used?

Before selecting which version of the BESSI to use, researchers and practitioners should consider a few key points. The first consideration is whether the assessment goals involve both SEB skill facets and domains. Both the BESSI-192 and BESSI-96 provide breadth and depth in SEB skill content and would therefore be useful in contexts where users wish to assess skills at both the domain and facet levels. For example, research on personality development suggests that facet-level traits follow distinct developmental trajectories during adolescence that would be masked if only trait domains were assessed (Ringwald et al., 2023; Soto et al., 2011). Utilizing the BESSI-192 or BESSI-96 in longitudinal research would capture facet-level specificity that the BESSI-45 and BESSI-20 could not. In addition, facets and nuances that are better captured in the BESSI-96 and BESSI-192 may be more predictive of particular outcomes than broad domains (Soto & John, 2017;

Stewart et al., 2022). Finally, in some assessment contexts, researchers and practitioners may be interested in assessing specific facets. In these contexts, it may be better to choose full item sets for particular facets from the BESSI-192 than to use the shorter measures.

By contrast, if the assessment goal is to measure only the broad SEB skill domains, then the BESSI-45 or BESSI-20 could serve as an adequate proxy for the BESSI-192. Because the BESSI-45 includes at least one item per facet within the five major skill domains, it may be particularly useful in work that needs to balance measurement efficiency with a breadth of content at the domain level. On the contrary, the BESSI-20 necessarily sacrifices some breadth of content due to its very small number of items.

Another consideration is assessment time. The BESSI-192 takes about 20 minutes to administer. This makes it well suited to assessment contexts that can devote considerable time to SEB skill assessment but may be too long for other settings. In contrast, the BESSI-96 requires about 50% (10 minutes), BESSI-45 about 25% (5 minutes), and BESSI-20 about 10% (2 minutes) of the BESSI-192's administration time. These time savings may be consequential in studies that need to assess many variables within a single session and in contexts where respondent fatigue is of concern, such as when children and adolescents are supplying self-ratings or when observers are rating multiple individuals. However, researchers and practitioners should keep in mind that these time savings do come at a cost to reliability and validity. We recommend using the BESSI-20 only when there are severe constraints on assessment time.

For researchers who want to preserve the full breadth of the BESSI-192 or the BESSI-96 but shorten the assessment time, a planned missingness design is also a viable option (PMD; Graham et al., 2006; see Lechner et al., 2022, for an example of a PMD with BESSI-192). In a PMD, all items in the measures are used, but each respondent is administered a subset of the items. Thus, the resulting data are missing at random, and missing data methods, such as multiple imputation or FIML, can be used without introducing bias (Zhang & Yu, 2022). However, because of the statistical expertise required to utilize a PMD, the BESSI short forms present an easy-to-use option for assessing SEB skills in applied settings.

Limitations and Future Directions

Though this work has important strengths including its use of samples and criterion measures that span developmental periods and rater perspectives, there are also some important limitations that highlight promising directions for future research. This research used a scale development approach designed to select items for the BESSI short forms that would retain as much domain- and facet-level information as possible from the full-length BESSI (for a similar, previous example, see Soto & John, 2017). However, another approach to short form development would be to select BESSI items that are relatively pure indicators of only one domain (Donnellan et al., 2006). Further research is needed to compare such approaches in terms of measurement properties, validity, and cross-cultural measurement invariance.

Because there is no standard in the literature for selecting thresholds for some scale validation criteria (e.g., evaluating the magnitude of column-vector correlations or retention of predictive power), our selection of specific thresholds was by necessity somewhat arbitrary. Despite this limitation, we believe that preregistering these thresholds is valuable, rather than only interpreting results post hoc. Moreover, many of the present results would remain encouraging even when evaluated against stricter thresholds (e.g., most column-vector correlations and explained variance ratios were greater than .90).

We administered the full BESSI to all participants in this research rather than administering each short form separately, which could bias results. However, supplemental analyses using a sample of high school students who completed the BESSI-45 indicated that there were no major discrepancies in terms of measurement properties compared to the High School Student sample that completed BESSI-192 (see Supplemental Tables S4, S11, and S12). In addition, a recent study using BESSI-45 demonstrated SEB skill domain associations similar to the full BESSI across a variety of criteria (Soto et al., 2022b). Future research could similarly explore the structure, reliability, and validity of BESSI-96 and BESSI-20 when they are assessed as standalone measures rather than when they are embedded in the full BESSI.

Given the calls for assessment and intervention on SEB skills during childhood (Domitrovich et al., 2017), future research should explore the reliability and validity of the BESSI short forms with younger samples. Future work can also assess whether the BESSI short forms are developmentally invariant across other factors (e.g., age, gender, and race/ethnicity) by collecting data from diverse samples at developmentally meaningful measurement points. Finally, an important future direction of this work would be to investigate SEB skills in clinical populations and mental health contexts, given meta-analytic findings indicating that K-12 students' SEB skill development is linked with better mental health (Durlak et al., 2011; Taylor et al., 2017).

Constraints on Generality

One limitation on the generalizability of this research is that the cross-cultural comparisons involved samples from only the United States and Germany. Thus, a critical future direction is to explore the cross-cultural comparability of the BESSI measure beyond U.S. and European contexts (Laajaj et al., 2019). In addition, future cross-cultural work on the BESSI should also consider the historical, political, and cultural contexts in which SEB skills are measured (Mahfouz & Anthony-Stevens, 2020; Tiwari & Ruedas-Gracia, 2023). For one, current conceptualizations and definitions of SEB skills, created and validated in the United States and Europe, may miss relevant SEB strengths such as spirituality and cultural connectedness (Hodge et al., 2015; Snowshoe et al., 2015). In addition, manifestations of SEB skills that are admired and encouraged in some contexts may be inappropriate in others depending on cultural norms and environmental constraints (Tiwari & Ruedas-Gracia, 2023). The relevance and validity of SEB skill measures, including the BESSI, among more diverse samples warrants careful exploration.

Conclusion

The 192-item BESSI assesses 32 specific SEB skill facets and five broad skill domains. This research developed three short forms consisting of 96, 45, and 20 items. These short forms preserved the BESSI-192's multidimensional structure, retained most of the full measure's reliability and validity, and demonstrated partial or full measurement invariance between U.S. and German cultural contexts. The BESSI short forms save assessment time and can help reduce participant fatigue in studies and interventions in which administering the full BESSI would not be feasible. Researchers and practitioners should carefully consider assessment goals when deciding which form of the BESSI best fits their context, needs, and constraints.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Christopher J. Soto, Christopher M. Napolitano, and Brent W. Roberts hold the copyright for the Behavioral, Emotional, and Social Skills Inventory (BESSI), which was used in the present research.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Data collection was supported by faculty research grants from Colby College and the University of Illinois at Urbana-Champaign.

ORCID iDs

Madison N. Sewell https://orcid.org/0000-0002-6531-3038
Clemens M. Lechner https://orcid.org/0000-0003-3053-8701
Christopher M. Napolitano https://orcid.org/0000-0001-8626-2914

Data Availability Statement

Data collection for the High School Student sample was supported by Character Lab and facilitated through the Character Lab Research Network, a consortium of schools across the country working collaboratively with scientists to advance scientific insights that help kids thrive. The item selection criteria, predictive validity quality control criteria, and analysis plan were preregistered, see https://osf.io/yxder and https://osf.io/nyt74. Data for this article are available, see https://osf.io/4zgyr/ and https://osf.io/9pvmj/. Online Supplemental Tables for this article are available, see https://osf.io/nyt74. Other study materials can be found at https://osf.io/2y7ua/.

Notes

1. The initial bifactor ESEM with 32 facets and five exploratory domains did not converge in the Panel sample. Because abstract thinking skill and creative skill were highly correlated (r = .84) in a CFA model with only the innovation facets, we allowed abstract thinking skill and creative skill items to load directly onto one facet factor in the ESEM. These 31 facet factors and five exploratory domain ESEM converged and demonstrated a good fit.
2. In some samples, the bifactor ESEM produced negative residual variances for some items, indicating that scores on these items were fully explained by the domain and facet factors. This issue can be addressed by constraining each problematic loading to have a facet loading equal to its two other same-facet items or by using the marker-variable approach to specify latent factors.
3. To further investigate BESSI-45's measurement properties, we also analyzed data from a newly available sample of U.S. high school students (N = 838) who completed BESSI-45 independently. Results from structural and reliability analyses were nearly identical to the High School Student sample that completed the BESSI-192. Full results from these analyses can be found in the online Supplemental Tables S4, S11, and S12.
4. We slightly changed the model used for the BESSI-96 higher-order test because the fully higher-order model did not converge. Instead, we used a semi-higher-order model in which all single domain and interstitial facets were loaded onto one or more domains. Compound facets were allowed to freely correlate with the five domain factors (instead of hierarchically loading onto the domains). This semi-higher-order model has an identical fit to the fully higher-order model in terms of CFI, TLI, RMSEA, and SRMR.

References

Abrahams, L., Pancorbo, G., Primi, R., Santos, D., Kyllonen, P., John, O. P., & De Fruyt, F. (2019). Social-emotional skill assessment in children and adolescents: Advances and challenges in personality, clinical, and educational contexts. *Psychological Assessment*, *31*(4), 460–473. https://doi.org/10.1037/pas0000591

Allen, J., Belfi, B., & Borghans, L. (2020). Is there a rise in the importance of socioemotional skills in the labor market? Evidence from a trend study among college graduates. *Frontiers in Psychology*, *11*, 1710. https://doi.org/10.3389/fpsyg.2020.01710

Anderson, C., John, O. P., Keltner, D., & Kring, A. M. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology*, *81*(1), 116–132. https://doi.org/10.1037/0022-3514.81.1.116

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*(5), 815–824.

Belsky, J., Jaffee, S., Hsieh, K.-H., & Silva, P. A. (2001). Child-rearing antecedents of intergenerational relations in young adulthood: A prospective study. *Developmental Psychology*, *37*(6), 801–813. https://doi.org/10.1037/0012-1649.37.6.801

Berg, J., Osher, D., Same, M. R., Nolan, E., Benson, D., & Jacobs, N. (2017). *Identifying, defining, and measuring social and emotional competencies*. American Institutes for Research.

Breil, S. M., Mielke, I., Ahrens, H., Geldmacher, T., Sensmeier, J., Marschall, B., & Back, M. D. (2022). Predicting actual social skill expression from personality and skill self-concepts. *Journal of Intelligence*, *10*(3), 1–16.

Bukowski, W. M., Hoza, B., & Boivin, M. (1994). Measuring friendship quality during pre- and early adolescence: The development and psycho-metric properties of the Friendship Qualities Scale. *Journal of Social and Personal Relationships*, *11*(3), 471–484. https://doi.org/10.1177/02654075 94113011

Carlo, G., Okun, M. A., Knight, G. P., & de Guzman, M. R. T. (2005). The interplay of traits and motives on volunteering: Agreeableness, extraversion and prosocial value motivation. *Personality and Individual Differences*, *38*(6), 1293–1305. https://doi.org/10.1016/j.paid.2004.08.012

Casillas, A., Way, J., & Burrus, J. (2015). Behavioral skills. In W. Camara, R. O'Connor, K. Mattern, & M. A. Hanson (Eds.), *Beyond academics: A holistic framework for enhancing education and workplace success* (pp. 25–38). ACT.

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

Church, A. T. (2001). Personality measurement in cross-cultural perspective. *Journal of Personality*, *69*(6), 979–1006. https://doi.org/10.1111/1467-6494.696172

Collaborative for Academic, Social, and Emotional Learning. (2020). *SEL: What are the core competence areas and where are they promoted?* https://casel.org/sel-framework/

Davidson, L. A., Crowder, M. K., Gordon, R. A., Domitrovich, C. E., Brown, R. D., & Hayes, B. I. (2018). A continuous improvement approach to social and emotional competency measurement. *Journal of Applied Developmental Psychology*, *55*, 93–106.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development*, *88*(2), 408–416. https://doi.org/10.1111/cdev.12739

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five Factors of personality. *Psychological Assessment*, *18*(2), 192–203. https://doi.org/10.1037/1040-3590.18.2.192

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

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.

Godin, G., & Shephard, R. J. (1985). A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Sciences*, *10*(3), 141–146.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, *11*(4), 323–343.

Guo, J., Tang, X., Marsh, H. W., Parker, P., Basarkod, G., Sahdra, B., Ranta, M., & Salmela-Aro, K. (2022). The roles of social–emotional skills in students' academic and life success: A multi-informant and multicohort perspective. *Journal of Personality and Social Psychology*, *124*(5), 1079–1110.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group effect. *Journal of Personality and Social Psychology*, *82*(6), 903–918. https://doi.org/10.1037/0022-3514.82.6.903

Hodge, D. R., Zidan, T., & Husain, A. (2015). Validation of the Intrinsic Spirituality Scale (ISS) with Muslims. *Psychological Assessment*, *27*(4), 1264–1272. https://doi.org/10.1037/pas0000130

Kautz, T., Heckman, J. J., Diris, R., Weel, B. T., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success* (NBER Working Paper No. w20749). National Bureau of Economic Research. https://doi.org/10.3386/w20749

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress*, *19*(3), 192–207. https://doi.org/10.1080/02678370500297720

Laajaj, R., Macours, K., Hernandez, D. A. P., Arias, O., Gosling, S. D., Potter, J., Rubio-Codina, M., & Vakis, R. (2019). Challenges to capture the big five personality traits

in non-WEIRD populations. *Science Advances*, *5*(7), 1–13. https://doi.org/10.1126/sciadv.aaw5226

Lechner, C. M., Knopf, T., Napolitano, C. M., Rammstedt, B., Roberts, B. W., Soto, C. J., & Spengler, M. (2022). The Behavioral, Emotional, and Social Skills Inventory (BESSI): Psychometric properties of a German-language adaptation, temporal stabilities of the skills, and associations with personality and intelligence. *Journal of Intelligence*, *10*(3), 63.

Mahfouz, J., & Anthony-Stevens, V. (2020). Why trouble SEL? The need for cultural relevance in SEL. *Occasional Paper Series*, *2020*(43), 6. https://doi.org/10.58295/2375-3668.1354

McCrae, R. R., Yik, M. S. M., Trapnell, P. D., Bond, M. H., & Paulhus, D. L. (1998). Interpreting personality profiles across cultures: Bilingual, acculturation, and peer rating studies of Chinese undergraduates. *Journal of Personality and Social Psychology*, *74*(4), 1041–1055. https://doi.org/10.1037/0022-3514.74.4.1041

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543. https://doi.org/10.1007/BF02294825

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. Routledge.

Mõttus, R., Allik, J., Realo, A., Rossier, J., Zecca, G., Ah-Kion, J., Amoussou-Yéyé, D., Bäckström, M., Barkauskiene, R., Barry, O., Bhowon, U., Björklund, F., Bochaver, A., Bochaver, K., De Bruin, G., Cabrera, H. F., Chen, S. X., Church, A. T., Cissé, D. D., . . . Johnson, W. (2012). The effect of response style on self-reported conscientiousness across 20 countries. *Personality and Social Psychology Bulletin*, *38*(11), 1423–1436. https://doi.org/10.1177/0146167212451275

Muthén, B., & Muthén, L. (2017). *Mplus user's guide* (8th ed.).

Napolitano, C. M., Sewell, M. N., Yoon, H. J., Soto, C. J., & Roberts Brent, W. (2021). Social, emotional, and behavioral skills: An integrative model of the skills associated with success during adolescence and across the life span. *Frontiers in Education*, *6*, 679561. https://doi.org/10.3389/feduc.2021.679561

Oberle, E., Schonert-Reichl, K. A., Hertzman, C., & Zumbo, B. D. (2014). Social–emotional competencies make the grade: Predicting academic success in early adolescence. *Journal of Applied Developmental Psychology*, *35*(3), 138–147.

Organisation for Economic Co-Operation and Development. (2015). *Skills for social progress: The power of social and emotional skills*. http://www.oecd.org/education/skills-for-social-progress-9789264226159-en.htm

Organisation for Economic Co-Operation and Development. (2021). *Beyond academic learning: First results from the survey of social and emotional skills*. https://doi.org/10.1787/92a11084-en

Park, D., Tsukayama, E., Goodwin, G. P., Patrick, S., & Duckworth, A. L. (2017). A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual competencies in children. *Contemporary Educational Psychology*, *48*, 16–27.

Paunonen, S. V. (2003). Big Five factors of personality and replicated predictions of behavior. *Journal of Personality and Social Psychology*, *84*(2), 411–424. https://doi.org/10.1037/0022-3514.84.2.411

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90.

Ringwald, W. R., Kaurin, A., Lawson, K. M., Wright, A. G., & Robins, R. (2023). The development of personality—From metatraits to facets—Across adolescence and emerging adulthood in a sample of Mexican-origin youth. *PsyArXiv*. https://doi.org/10.31234/osf.io/ft3cq

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36.

Rounds, J., Ming, C. W. J., Cao, M., Song, C., & Lewis, P. (2016). *Development of an O*NET Mini Interest Profiler (Mini-IP) for mobile devices: Psychometric characteristics*. Department of Labor O* NET Resource Center.

Sewell, M. N., Napolitano, C. M., Roberts, B. W., Soto, C. J., & Yoon, H. J. (2023). The social, emotional, and behavioral skill antecedents to college students' volunteering during the COVID-19 pandemic. *Journal of Research on Adolescence*, *33*(2), 618–631.

Sewell, M. N., Soto, C. J., Napolitano, C. M., Yoon, H. J., & Roberts, B. W. (2022). Survey data of social, emotional, and behavioral skills among seven independent samples. *Data in Brief*, *40*, 107792. https://doi.org/10.1016/j.dib.2022.107792

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, *100*(4), 765–781. https://doi.org/10.1037/a0012840

Snowshoe, A., Crooks, C. V., Tremblay, P. F., Craig, W. M., & Hinson, R. E. (2015). Development of a Cultural Connectedness Scale for First Nations youth. *Psychological Assessment*, *27*(1), 249–259. https://doi.org/10.1037/a0037867

Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory-2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, *68*, 69–81. https://doi.org/10.1016/j.jrp.2017.02.004

Soto, C. J., & John, O. P. (2019). Optimizing the length, width, and balance of a personality scale: How do internal characteristics affect external validity? *Psychological Assessment*, *31*(4), 444–459. https://doi.org/10.1037/pas0000586

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*(2), 330–348. https://doi.org/10.1037/a0021717

Soto, C. J., Napolitano, C. M., & Roberts, B. W. (2021). Taking skills seriously: Toward an integrative model and agenda for social, emotional, and behavioral skills. *Current Directions in Psychological Science*, *30*(1), 26–33.

Soto, C. J., Napolitano, C. M., Sewell, M. N., Yoon, H. J., Murano, D., Casillas, A., & Roberts, B. W. (2023). What I do and what I can do: Testing the incremental validity of social, emotional, and behavioral skills vs. traits for predicting academic success. *Journal of Research in Personality*, *104*, 104382. https://doi.org/10.1016/j.jrp.2023.104382

Soto, C. J., Napolitano, C. M., Sewell, M. N., Yoon, H. J., & Roberts, B. W. (2022a). An integrative framework for conceptualizing and assessing social, emotional, and behavioral

skills: The BESSI. *Journal of Personality and Social Psychology*, *123*(1), 192–222.

Soto, C. J., Napolitano, C. M., Sewell, M. N., Yoon, H. J., & Roberts, B. W. (2022b). Going beyond traits: Social, emotional, and behavioral skills matter for adolescents' success. *Social Psychological and Personality Science*, *15*(1), 33–45.

Stewart, R. D., Mõttus, R., Seeboth, A., Soto, C. J., & Johnson, W. (2022). The finer details? The predictability of life outcomes from Big Five domains, facets, and nuances. *Journal of Personality*, *90*(2), 167–182. https://doi.org/10.1111/jopy.12660

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, *88*(4), 1156–1171. https://doi.org/10.1111/cdev.12864

Tiwari, A., & Ruedas-Gracia, N. (2023). *SEL is not universal: Addressing the cross-cultural complexities*. UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development. https://mgiep.unesco.org/article/sel-is-not-universal-cross-cultural-complexities-in-conceptualizing-and-applying-socioemotional-skills

Walton, K. E., Murano, D., Burrus, J., & Casillas, A. (2021). Multimethod support for using the big five framework to organize social and emotional skills. *Assessment*, *30*(1), 144–159. https://doi.org/10.1177/10731911211045744

West, M. R., Buckley, K., Krachman, S. B., & Bookman, N. (2018). Development and implementation of student social-emotional surveys in the CORE districts. *Journal of Applied Developmental Psychology*, *55*, 119–129.

Wetzel, E., & Roberts, B. W. (2020). Commentary on Hussey and Hughes (2020): Hidden invalidity among 15 commonly used measures in social and personality psychology. *Advances in Methods and Practices in Psychological Science*, *3*(4), 505–508. https://doi.org/10.1177/251524 59 20957618

Yoon, H. J., Roberts, B. W., Sewell, M. N., Napolitano, C. M., Soto, C. J., Murano, D., & Casillas, A. (2024). Examining SEB skills' incremental validity over personality traits in predicting academic achievement. *PLOS ONE*, *19*(1), e0296484. https://doi.org/10.1371/journal.pone.02 96484

Zhang, C., & Yu, M. C. (2022). Planned missingness: How to and how much? *Organizational Research Methods*, *25*(4), 623–641. https://doi.org/10.1177/10944281211016534s

Appendix

The BESSI Short Form Items

Instructions

Here is a list of activities or things you could do. For each one, please select a response to indicate *how well you can do that thing*. For example, how well can you *follow the instructions for an assignment*? Note that how well you can do something may be different from how often you do it, or how much you like to do it. For each activity, you should rate *how well* you can do that thing.

1	2	3	4	5
Not at all well	Not very well	Pretty well	Very well	Extremely well

1.	Lead a group of people.
2.	Sympathize with other people's feelings.
3.	Show up for things on time.
4.	Understand abstract ideas.
5.	Stop myself from worrying.
6.	Organize my personal spaces.
7.	Find the energy to get things done.
8.	See the good in people.
9.	Repeat a task consistently.
10.	Understand my emotions.
11.	Look on the bright side of things.
12.	Keep working until a task is finished.
13.	Win debates with other people.
14.	Make people feel comfortable.
15.	Take care of details.
16.	Use my imagination.
17.	Express my thoughts and feelings.
18.	Follow instructions.
19.	Try new things.
20.	Calm down when I'm feeling angry.
21.	Manage my responsibilities.
22.	Make sense of complex information.
23.	Work as part of a group.
24.	Set clear goals.
25.	Meet new people.
26.	Have confidence in myself.
27.	Make careful decisions.
28.	Draw or paint.
29.	Take responsibility when I've made a mistake.
30.	Resist temptations.
31.	Make decisions on my own.
32.	Learn about other cultures.
33.	Assert myself as a leader.
34.	Take another person's perspective.
35.	Follow a schedule.
36.	Discuss complicated topics and ideas.
37.	Calm down when I'm feeling anxious.
38.	Keep things neat and tidy.
39.	Maintain a high energy level.
40.	Forgive people quickly.
41.	Do the same task over and over again.
42.	Pay attention to my thoughts and feelings.
43.	Stay positive when something bad happens.
44.	Work efficiently, without wasting time.
45.	Speak up when I disagree with others.
46.	Get along with people.
47.	Find and correct mistakes.
48.	Invent things.
49.	Tell people how I am feeling.
50.	Do what I'm supposed to do.
51.	Try something that's unfamiliar.
52.	Control my temper.
53.	Keep track of my promises and commitments.

(continued)

54. Learn things quickly.
55. Work with people toward a shared goal.
56. Make plans to achieve a goal.
57. Talk to people.
58. Find reasons to feel good about myself.
59. Think before acting.
60. Create art.
61. Stop myself from lying or cheating.
62. Control my impulses.
63. Do things on my own.
64. Understand people from different backgrounds.
65. Take charge of a situation.
66. Understand how other people feel.
67. Plan out my time.
68. Think deeply about things.
69. Settle down when I'm feeling nervous.
70. Put things back in their proper place.
71. Keep myself motivated.
72. Trust people.
73. Do tasks that are routine or repetitive.
74. Examine myself and my life.
75. Keep a positive attitude.
76. Concentrate on a task.
77. Win arguments.
78. Make a positive impression on people.
79. Double-check my work.
80. Come up with new ideas.
81. Explain what's on my mind.
82. Follow the rules.
83. Adapt to change.
84. Stop myself from getting angry.
85. Follow through on promises.
86. Find logical solutions to problems.
87. Cooperate with other people.
88. Work toward my goals.
89. Start a conversation.
90. See my good qualities.
91. Think things through carefully.
92. Write stories or poems.
93. Be honest with people.
94. Stop myself from acting on impulse.
95. Get things done by myself.
96. Appreciate different cultures.

Note. BESSI items copyright ©2019 by Christopher J. Soto, Christopher M. Napolitano, and Brent W. Roberts. Reprinted with permission.

Scoring the BESSI-96 Skill Facets

Each of the 32 facet scales should be scored by averaging its three BESSI items.

Time management = 3, 35, 67
Organizational skill = 6, 38, 70Capacity for consistency = 9, 41, 73
Task management = 12, 44, 76
Decision-making skill = 27, 59, 91
Leadership skill = 1, 33, 65
Persuasive skill = 13, 45, 77
Expressive skill = 17, 49, 81
Conversational skill = 25, 57, 89
Energy regulation = 7, 39, 71
Perspective-taking skill = 2, 34, 66
Capacity for trust = 8, 40, 72
Capacity for social warmth = 14, 46, 78
Teamwork skill = 23, 55, 87
Ethical competence = 29, 61, 93
Stress regulation = 5, 37, 69
Detail management = 15, 47, 79
Rule-following skill = 18, 50, 82
Responsibility management = 21, 53, 85
Goal regulation = 24, 56, 88
Capacity for optimism = 11, 43, 75
Anger management = 20, 52, 84
Confidence regulation = 26, 58, 90
Impulse regulation = 30, 62, 94
Abstract thinking skill = 4, 36, 68
Creative skill = 16, 48, 80
Artistic skill = 28, 60, 92
Cultural competence = 32, 64, 96
Information processing skill = 22, 54, 86
Self-reflection skill = 10, 42, 74
Adaptability = 19, 51, 83
Capacity for independence = 31, 63, 95

Scoring the BESSI-96 Skill Domains

Each of the five domains should be scored by averaging its component facets. As noted below, interstitial facets should be given half weight on each of their two assigned domains.

Self-Management Skills = Time Management, Organizational Skill, Capacity for Consistency, Task Management, Detail Management, Rule-Following Skill, Responsibility Management, Goal Regulation, Decision-Making Skill, Energy Regulation (half weight), Ethical Competence (half weight), Impulse Regulation (half weight), Information Processing Skill (half weight)

Social Engagement Skills = Leadership Skill, Persuasive Skill, Expressive Skill, Conversational Skill, Energy Regulation (half weight)

Cooperation Skills = Perspective-Taking Skill, Capacity for Trust, Capacity for Social Warmth, Teamwork Skill, Ethical Competence (half weight)

Emotional Resilience Skills = Stress Regulation, Capacity for Optimism, Anger Management, Confidence Regulation, Impulse Regulation (half weight)
Innovation Skills = Abstract Thinking Skill, Creative Skill, Artistic Skill, Cultural Competence, Information Processing Skill (half weight)

Scoring BESSI-45 Skill Domains

Each BESSI-45 domain scale should be scored by averaging its nine BESSI items.
Self-Management Skills = 38, 53, 67, 73, 76, 79, 82, 88, 91
Social Engagement Skills = 1, 17, 33, 39, 45, 49, 57, 77, 89
Cooperation Skills = 2, 8, 23, 29, 40, 46, 66, 78, 87
Emotional Resilience Skills = 37, 43, 52, 58, 62, 69, 75, 84, 90
Innovation Skills = 4, 22, 28, 32, 36, 48, 60, 64, 80

Scoring BESSI-20 Skill Domains

Each BESSI-20 domain scale should be scored by averaging its four BESSI items.
Self-Management Skills = 53, 67, 76, 88
Social Engagement Skills = 1, 17, 45, 89
Cooperation Skills = 8, 46, 66, 87
Emotional Resilience Skills = 37, 52, 58, 75
Innovation Skills = 4, 32, 60, 80