

Development of an Inventory Assessing Social and Emotional Skills in Brazilian Youth

Ricardo Primi,^{1,5} Daniel Santos,^{2,5} Oliver P. John,^{4,5} and Filip De Fruyt^{3,5}

¹Graduate Program in Psychological Assessment, Universidade São Francisco, Itatiba, Brazil,

²Economics Department, Universidade de São Paulo, Ribeirão Preto, Brazil, ³Department of Developmental, Personality, and Social Psychology, Ghent University, Belgium, ⁴Department of Psychology, University of California, Berkeley, CA, USA, ⁵EduLab21, Ayrton Senna Institute, São Paulo, Brazil

Abstract. Whereas the structure of individual differences in personal attributes is well understood in adults, much less work has been done in children and adolescents. On the assessment side, numerous instruments are in use for children but they measure discordant attributes, ranging from one single factor (self-esteem; grit) to three factors (social, emotional, and academic self-efficacy) to five factors (strength and difficulties; Big Five traits). To construct a comprehensive measure for large-scale studies in Brazilian schools, we selected the eight most promising instruments and studied their structure at the item level (Study 1; $N = 3,023$). The resulting six-factor structure captures the major domains of child differences represented in these instruments and resembles the well-known Big Five personality dimensions plus a negative self-evaluation factor. In a large representative sample in Rio de Janeiro State (Study 2; $N = 24,605$), we tested a self-report inventory (SENNA1.0) assessing these six dimensions of socio-emotional skills with less than 100 items and found a robust and replicable structure and measurement invariance across grades, demonstrating feasibility for large-scale assessments across diverse student groups in Brazil. Discussion focuses on the contribution to socio-emotional research in education and its measurement as well as on limitations and suggestions for future research.

Keywords: 21st century skills, social-emotional skills, instrument development, Big Five, five-factor model, measurement invariance, exploratory structural equation modeling

Recent developments in the fields of education, economics, psychology, and neuroscience have shown that the so-called “soft skills” or social-emotional skills (SEMS¹) are as important as cognitive measures such as IQ or fluid intelligence in predicting various life outcomes (OECD, 2015). SEMS are also frequently referred to as 21st Century Skills, personal characteristics that are considered crucial for individual’s life and work adaptation (Partnership for 21st Century Skills, www.p21.org; Santos, 2011, Trilling & Fadel, 2009). SEMS can be defined as: “individual characteristics that (1) originate in the reciprocal interaction between biological predispositions and environmental factors, (2) are manifested in consistent patterns of thoughts, feelings and behaviors, (3) continue to develop through formal and informal learning experiences, and (4) influence important socioeconomic outcomes throughout the individual’s life” (De Fruyt, Wille, & John, 2015; John & De Fruyt, 2014).

In the absence of a comprehensive and explicitly specified taxonomic framework, SEMS have been operationalized in research by an amalgam of different constructs

including, among others, self-concepts, attitudes and control beliefs, motivational constructs, and personality traits (OECD, 2015). Various subsets of constructs, drawn from this great variety, have been related to many important life outcomes. For example, there is now convincing evidence that self-efficacy is related to school achievement, but also that personality traits such as Conscientiousness and Openness to Experience are associated with educational attainment (Poropat, 2009); similar effects have been shown for the positive-psychology concept of grit (Duckworth & Quinn, 2009; MacCann, Duckworth, & Roberts, 2009). Conscientiousness and Openness in adolescence have also been linked to wages and other labor market outcomes (e.g., Almlund, Duckworth, Heckman, & Kautz, 2011).

A key characteristic of SEMS is that they develop in interaction with the environment and that schooling and school-based intervention programs can affect and moderate their developmental course (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). In a similar vein,

¹ We abbreviate social-emotional skills as SEMS to distinguish it unambiguously from the frequently used acronym SES (social-economic status) in this field.

contemporary personality research acknowledges that personality traits develop and change substantially during childhood and adolescence and tend to remain malleable well into adulthood (De Fruyt & Van Leeuwen, 2014; Denissen, 2014; Soto & John, 2014; Soto, John, Gosling, & Potter, 2011; Srivastava, John, Gosling, & Potter, 2003). The development of SEMS has thus become a key target to achieve in education, next to knowledge transfer and cognitive achievements, but also because there is evidence that the development of SEMS contributes to cognitive achievements at school (Durlak et al., 2011). Investing in SEMS development is thus not only important for its own sake, but also via its enhancing effects on traditional educational achievement outcomes.

Most of the educational systems around the world use content-based indicators (such as math and language scores) to monitor school performance. The previously described paradigm shift, however, emphasizes that schools are not only important for knowledge transmission but also for fostering SEMS (OECD, 2015). Attention for these two intertwined outcome domains is therefore considered crucial in education to promote individuals' success and happiness in future society. The implication is that this paradigm shift will have to be reflected in school monitoring and accountability systems, which will have to assess the traditional content-based indicators of achievement and SEMS. This evolution necessitates measurement tools to assess and evaluate SEMS comprehensively and reliably, so comparative school effectiveness evaluations can be conducted or the impact of a specific intervention on SEMS development can be examined.

To achieve these goals, assessment tools are needed. Whereas the structure of individual differences in personal, social, and emotional attributes is well understood in adults (e.g., John, Naumann, & Soto, 2008), much less work has been done in children and adolescents. On the assessment side, researchers from different disciplinary and theoretical backgrounds have developed instruments for use with children and adolescents. Many of these instruments, however, tend to focus on particular theoretical interests and practical concerns (e.g., children's beliefs about external control; conduct problems), rather than offer a comprehensive assessment of social and emotional attributes important in academic settings and for long-term life outcomes.

Therefore, as a first step to construct a comprehensive measure for large-scale studies in Brazilian schools, we reviewed the available self-report instruments and examined the eight most promising ones in Study 1, analyzing their factor structure at the item level to identify the most important dimensions of socio-emotional functioning in Brazilian youth.

In Study 2, we examined whether a short self-report instrument (called SENNA1.0) designed to measure the dimensions uncovered in Study 1 can be constructed and administered in large-scale school-based assessments in Brazil, including youth with heterogeneous and low social-economic backgrounds. Specifically, we examined whether the administration of a self-report instrument is feasible in a single classroom period in kids as young as 5th grade, whether the factor structure is stable and measurement-invariant across different grades, and whether the scales are suffi-

ciently reliable, so they can be used for large-scale monitoring of SEMS development in Brazilian schools.

Study 1: Analysis of the Item-Level Factor Structure of Available Child Instruments to Identify Dimensions to Assess with SENNA 1.0

Background and Rationale

The SEMS domain has been hampered by the lack of an integrative framework directing the domain's coverage and helping to identify its key defining dimensions (De Fruyt et al., 2015). To fill this gap and identify core qualities to be included in a more comprehensive measure, Santos and Primi (2014) carefully reviewed measures that assess social and emotional characteristics and were available in Brazil or could be translated into Portuguese. To guide their systematic search for SEMS-relevant instruments, they used four key criteria:

- (1) *Predictive power*: The measured constructs had to be related to at least one measure of well-being, achievement, or other future outcomes (the stronger and more reliable this association the better, as well as the larger the set of related outcomes);
- (2) *Feasibility*: Given the constraints in Brazil, the instrument should be low cost, take relatively little time to administer, be formulated in age-appropriate language, and be usable for self-descriptions without the assistance of a professional psychologist;
- (3) *Malleability*: The measured constructs should be malleable during school age, preferably with evidence that formal education can impact their course (again, the stronger and more reliable this evidence, the better); and finally
- (4) *Robustness*: The instrument should have good psychometric properties.

Adopting these criteria, the literature review identified eight instruments that are listed, by year of publication, in Table 1, along with their abbreviations, number and names of their scales, and their total number of items. This overview shows a wide range of scales measuring diverse constructs that are popular in educational, psychological, and economic research on SEMS, such as self-concepts, self-esteem and self-efficacy, motivation, attitudes and control beliefs (attributing control internally or externally), social and emotional behavior that may be adaptive or maladaptive, and personality traits. The overview in Table 1 also illustrates the state of the field: there is no consensus on the number and nature of the constructs needed to represent socio-emotional characteristics. For example, the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) suggests five constructs are needed, the Self-Efficacy

Table 1. Selected instruments: Names, abbreviations, number of items and scales, and scale names

Instrument	Items	Scales	Names of scales
Locus of Control Scale, Nowicki and Strickland (1973)	21	1	Locus of control
Rosenberg Self-Esteem Scale, Rosenberg (1979)	10	1	Self-esteem
Strengths and Difficulties Questionnaire (SDQ), Goodman (1997)	25	5	Emotional symptoms, Conduct problems, Hyperactivity, Peer problems, Prosocial skills
Big Five Inventory (BFI), John, Donahue, and Kentle (1991)	44	5	Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness
Self-Efficacy Questionnaire for Children (SEQ-C), Muris (2001)	24	3	Academic, Social, and Emotional self-efficacy
Big Five for Children (BF-C), Barbaranelli, Caprara, Rabasca, and Pastorelli (2003)	65	5	Neuroticism, Extraversion, Openness, Agreeableness, Conscientiousness
Core Self-Evaluations (CORE), Judge et al. (2003)	12	1	Positive vs. negative self-evaluation (combined self-esteem, generalized self-efficacy, locus of control, low neuroticism)
Grit Scale, Duckworth and Quinn (2009)	8	1	Grit
Total	209	21	16

Questionnaire for Children (SEQ-C) suggests three, and both Core Self-Evaluations (CORE; Judge, Erez, Bono, & Thoresen, 2003) and the Grit Questionnaire (Duckworth & Quinn, 2009), pointing to a single (but different) constructs.

In all, Table 1 shows that the eight instruments generate 21 distinct scales, which are measured with more than 200 unique items. Does the field need each of the individual concepts that are implied by all these scales? Do we need to measure each of them separately? As in the work on the structure of intelligence or personality traits, it seems likely that these scales do not all capture unique variance. Instead, they are likely to show considerable overlap, if not outright redundancy. In general, we expected some core themes to be represented across many of these instruments and their scales. One such core theme should be the experience of strong negative emotions, as contrasted with the appropriate regulation of these emotions. For example, we expected that the Emotional Symptoms scale on the Strengths and Difficulties Questionnaire (SDQ) may measure a concept and include items that are quite similar to what is called Negative Affect or Neuroticism on the Big Five Inventory (BFI), and that both may represent the conceptual opposite of the “Emotional” subscale on the Self-Efficacy Questionnaire.

Similarly, because completing tasks at school is so important in the educational life of children, we hypothesized that the degree to which kids can concentrate, persist, and be “on time” when working on school assignments and homework is likely addressed by many of these socio-emotional measures, whether the scales are called Academic Self-Efficacy, Conscientiousness, Grit, or even (low) Hyperactivity.

In short, we expected that the items of these instruments could be represented by a set of common latent factors, and that less than 21 factors would be needed. Thus, a first key contribution of this study is to examine the commonality and overlap that is likely to exist across the items in this set. Ideally, this common variance can be represented by a more manageable set of broader SEMS qualities (e.g., no more than 10), while at the same time ensuring comprehensiveness by providing better coverage of the entire domain than can be achieved by any one of the eight

instruments being considered. Once key underlying dimensions are identified, they can be used for item selection and the construction of a new, more comprehensive instrument.

The objectives of Study 1 were to (a) examine the latent structure underlying the items of these “21st century scales,” checking how many latent dimensions were needed to account for item covariances, (b) examine how the resulting dimensions can be interpreted and related to the dimensions of a generally agreed upon taxonomic framework for personality traits, namely the Big Five dimensions (e.g., John et al., 2008), and (c) select items that best represent these latent dimensions to construct a first version of a new comprehensive self-report inventory (SENNA).

Method

Participants

Participants were 3,023 students who attended one of 86 classes within 16 schools; these schools were selected to represent typical levels of performance on standardized achievement in Rio de Janeiro State. Students were in 5th ($N = 697$), 6th ($N = 710$), 9th ($N = 674$), 10th ($N = 488$), and 12th (454) grades. They were from two educational systems, one funded by the local municipality (student $N = 2,081$) and the other funded by the state ($N = 942$). The number of students who completed each of the instruments listed in Table 1 was as follows. Locus: $N = 1,026$, BFI: $N = 927$, SDQ: $N = 1,055$, Self-efficacy: $N = 1,011$, BFC: $N = 920$, Grit: $N = 985$, Rosenberg: $N = 602$ (only 4–6th) and Core: $N = 409$ (only 10–12th).

Measures

The eight instruments included in this study can be found in Table 1. In order to conduct an item-level factor analysis, it is necessary to estimate the full inter-item correlation matrix. Ideally this is accomplished by having all student participants answer all items of all instruments. However,

in practice, with more than 200 items, this is not feasible because diminished concentration and fatigue are likely to occur when children have to answer more than even 100 items.

In order to obtain a full correlation matrix, a Balanced Incomplete Block design was employed (BIB; Sailer, 2005). We created booklets of two instruments for 5th and 6th grade students and three instruments for the 9th, 10th, and 12th grades (see Appendix). We systematically produced combinations of instruments in booklets in a way to balance all pairwise combinations of the eight scales, such that each pair was included in at least one booklet, and instruments were equally distributed among students. These combinations generated 20 booklets for younger children and seven for older students. Each booklet had an average of 67 items.

At the time the data were collected in the classrooms, each student was administered one booklet. Booklets were systematically distributed within classes, always returning to the first booklet after the 20th student (younger children) or the 7th student (older children) had received a booklet. This spiraling strategy produced equivalent random samples of respondents for each booklet. Therefore, for each pair of items, we had a random sample of children allowing the calculation of the full inter-item correlation matrix. Considering a statistical power of 0.80 for the detection of correlations with a magnitude of 0.30 (i.e., an 80% chance of detecting a correlation with a magnitude of 0.30 or more as being statistically significant), it was established that samples of 85 students would be necessary (Cohen, 1992) for each booklet of items. In the final sample the number of students answering each booklet ranged from 93 to 252, so final numbers were well beyond those indicated by the power analysis.

Procedure

Translation, Adaptation, and Content Review of Instruments

Translation of items and adaptation of instruments aimed to adapt item content to the Brazilian target culture in a manner that preserved the intended relationships between indicators and construct (Borsa, Damásio, & Bandeira, 2012). To do so five steps were followed:

- (1) *Translation*: Two to four researchers prepared independent translations of the instruments. When a Portuguese version of an instrument already existed, this version was included as an additional translation.
- (2) *Sensitivity analysis*: Two groups of school administrators from the Rio de Janeiro State and the municipal education systems examined the items and the available translations to check their suitability for the target population and evaluate possible negative reactions to particular items by teachers, school directors, and parents.
- (3) *Consensus translation*: A final translation was produced summarizing and implementing all suggestions obtained in the previous phases.

- (4) *Pilot study*: A qualitative pilot study was conducted with 48 children from 5th to 9th grades, divided into groups of four or five children each. In some sessions, students read the items with the help of researchers; in others, they just responded to the items and explained afterwards to the researchers how they interpreted the items. Researchers checked whether children understood the items and properly used the Likert 5-point rating scale. They also evaluated whether the item content was suitable for the respective age groups and understandable within the sociocultural experience of these Brazilian children. The information obtained from this careful piloting served to further review the items.
- (5) *Back translation*: After the revisions suggested during previous phases had been implemented, a final version of each instrument was prepared, translated back into English, and sent to the original authors for approval.

Length of the Booklets

Little is known about how many items children as young as 10 years can answer during a 40-min period, within the usual 50-min classroom teaching period (10 min was reserved for demographic and socioeconomic background questions). To find out, a sample of 228 children (5th grade: $N = 60$, 6th grade: $N = 42$, and 10th grade: $N = 126$) responded to the eight instruments (with their order randomly distributed) until the class period ended. They were told to answer items in the order in which they were listed in the booklet. Start and end times of each instrument were recorded for each individual student. While the students answered items, researchers timed the progress of the sessions and had students switch to a different-colored pen every ten min, thus enabling us to measure the number of items completed in 10-min intervals. On average, students answered 4.1 items per minute but, as expected, we found substantial age differences: The youngest students (5th grade) answered only 2.4 items per minute and a subgroup of slower students answered even fewer (1.5 items per minute). In terms of the total number of items answered in 40 min, this pilot study suggested three conclusions: (a) the slowest younger children would answer about 60 items, (b) the average-speed younger children up to 94 items, and (c) the typical 7th and 8th graders up to about 160 items. Taking into account the slower students, the booklets were limited to 60 items for 5th graders, and to 120 items for 9th–12th graders. These thresholds formed the constraints for the BIB design.

Results and Discussion

Factor Structure Underlying this Broadly Defined SEMS Item Set

Exploratory Factor Analyses (EFAs) were conducted at the item level, examining solutions with a different number of

factors, using the Unweighted Least Squares (ULS) method with GEOMIN oblique rotation. It was further taken into account that, strictly speaking, the item responses were ordered categorical variables, circumventing the problem of non-normality of the distribution of items (equaling an analysis of a polychoric inter-item correlation matrix). Analyses were performed using MPLUS software (Muthén & Muthén, 2012).

Parallel analysis indicated a maximum of 12 factors with eigenvalues higher than random generated eigenvalues, whereas scree plot analysis suggested no more than eight factors. Given that the two criteria did not agree and the fact that parallel analyses often suggest too many factors, between 5 and 10 factors were extracted, and the resulting six matrices of factor loadings were inspected. Indeed, in the 10-factor solution, four of the rotated factors were clearly under-defined and had few substantial loadings, suggesting a substantial overextraction of factors. However, the six well-defined factors in that solution were already present in the six-factor solution and replicated in all of the analyses retaining larger numbers of factors. Moreover, these six stable factors also proved the most interpretable in all solutions examined. Sample items defining each of the six factors, from the different instruments, are shown in Table 2. The full matrix of factor loadings can be obtained from the first author.

Conscientiousness: Working Hard and Persevering at Tasks at School

This was by far the largest factor (explaining 13% of the total variance) and represented across almost all the instruments studied here. It captures the effort (“hard work”) and perseverance needed to successfully complete school work and other achievement-related tasks in an efficient and timely way. As shown in Table 2, six of the instruments included items that loaded highly on this factor. The highest loading items came from the Grit scale, emphasizing working hard and diligently, followed by Big Five Conscientiousness items about efficiency, dutifulness, and perseverance. Additional item content came from the CORE Self-evaluation and from the Academic Self-efficacy scales, both emphasizing self-appraisals as being successful at completing tasks. The SDQ Hyperactivity scale further added strong attentional focus, as illustrated by the ability to finish “the work I’m doing.” These are all well-established elements of the Conscientiousness (C) domain (e.g., MacCann et al., 2009), suggesting that this Brazilian factor is a close child analog of an important dimension in the Big Five taxonomy.

Emotional Stability: Managing Negative Emotions Versus Experiencing Negative Affect

This factor accounted for 8% of the total variance and captures how the child experiences and regulates core features of negative emotion, such as anger/temper, sadness/hopelessness, and fear/nervousness. The theme of strong

negative emotional experiences (getting nervous, sad, or angry easily) on the high pole is contrasted with the ability to control or prevent such feelings (handling stress well; able to prevent becoming nervous) on the low pole. Note that items related to this emotional stability or regulation factor were represented in items from five of the instruments studied, including CORE Self-evaluations (things looking bleak and hopeless), SDQ Conduct Problems (get angry and lose my temper), and Emotional Self-efficacy (prevent becoming nervous). Items from the single-dimension Grit questionnaire are noticeably absent here; Grit does not involve emotional reactions but is focused solely on those aspects of task performance captured by Factor 1.

Extraversion: Energetic Approach to the Social World

The third largest factor (5% of the variance) captures themes related to Extraversion in the Big Five, with a focus on positive energy (lively, full of energy, not quiet, satisfied with self) and social skills that facilitate making contact with peers and initiating friendships (can tell a funny event to other children, good at making friends, talkative). Again, as shown in Table 2, five instruments contributed high loading items, including not only the BFI and BF-C but also items from the Self-efficacy, CORE Self-evaluation, and SDQ scales.

Agreeableness: Prosocial Skills in Peer Relationships

This factor (4% of the variance) captured themes related to cooperation, compassion, and forgiveness; these interpersonal strengths are described as Prosocial skills on the SDQ (caring about other people’s feelings) and as Agreeableness in the Big Five taxonomy. Clearly, the ability to feel compassion and empathy for one’s classmates, treat them with kindness and respect, help them when they have difficulties, and forgive their transgressions are both social and emotional in nature and represent a core agenda for socio-emotional teaching in schools. It is thus disconcerting that only three of the instruments included relevant item content. In other words, despite their names, the Social Self-efficacy and CORE Self-evaluation scales do not include these prosocial considerations.

Open-Mindedness: Curiosity, Imagination, and Invention

Although accounts of 21st century skills highlight the importance of a curious and imaginative mind-set and the need for greater creativity and invention, only the BFI and BF-C included items that loaded substantially on the relatively small factor (3% of the variance) that captured these themes. Best labeled as Openness or Open-mindedness, this factor also included particular likes and preferences associated with intrinsic motivation, such as liking to know and learn things, liking scientific shows on television, and creating new games or entertainments.

Table 2. Six common factors underlying the socio-emotional skills instruments for youth in Brazil: Eight example items for each factor, the original measure, and the loading of the item on its primary factor

Factor and item	Original measure	Item number	Item text in English original	Primary loading
C 1	GRIT	T8_04GR1	I am a hard worker	.725
C 2	GRIT	T8_08GR1	I am diligent (hardworking and careful)	.711
C 3	BFI	T2_33C1	Does things efficiently	.706
C 4	BF-C	T1_3801	I do my own duty	.679
C 5	BFI	T2_28C1	Perseveres until the task is finished	.671
C 6	CORE	T7_05SE1	I complete tasks successfully	.609
C 7	SEQ-C	T5_04AC1	How well do you succeed in finishing all your homework every day?	.605
C 8	SDQ	T3_25HA0	I finish the work I'm doing. My attention is good	.567
ES 1	BF-C	T1_22N1	I easily lose my calm	-.759
ES 2	BFI	T2_39N0	Gets nervous easily	-.720
ES 3	BFI	T2_29N0	Can be moody	-.697
ES 4	BF-C	T1_20N1	I am sad	-.635
ES 5	CORE	T7_12SE0	There are times when things look pretty bleak and hopeless to me	-.602
ES 6	SDQ	T3_05CP1	I get very angry and often lose my temper	-.543
ES 7	BFI	T2_09N1	Is relaxed, handles stress well	.513
ES 8	SEQ-C	T5_19EM1	How well can you prevent becoming nervous	.420
E 1	BFI	T2_21E0	Tends to be quiet	-.642
E 2	BF-C	T1_52E1	I am happy and lively	.567
E 3	SDQ	T302HA1	I am restless, I cannot stay still for long	.559
E 4	BFI	T201E1	Is talkative	.537
E 5	BFI	T2_11E1	Is full of energy	.530
E 6	CORE	T7_07SE1	Overall, I am satisfied with myself	.523
E 7	SEQ-C	T5_10SC1	How well can you become friends with other children?	.486
E 8	SEQ-C	T5_14SC1	How well can you tell a funny event to a group of children?	.474
A 1	BFI	T2_32A1	Is considerate and kind to almost everyone	.711
A 2	SDQ	T3_01PS1	I try to be nice to other people. I care about their feelings.	.627
A 3	BFI	T2_42A1	Likes to cooperate with others	.615
A 4	BFC	T1_06A1	I treat my peers with love and warmth	.600
A 5	BFI	T2_07A1	Is helpful and unselfish with others	.573
A 6	BFC	T1_07A1	I behave with others with great kindness	.570
A 7	BFI	T2_17A1	Has a forgiving nature	.525
A 8	BFC	T1_09A1	If a classmate has some difficulty, I help him/her	.520
O 1	BFI	T2_25O1	Is inventive	.587
O 2	BFI	T2_05O1	Is original, comes up with new ideas	.571
O 3	BFI	T2_20O1	Has an active imagination	.533
O 4	BFI	T2_10O1	Is curious about many different things	.494
O 5	BF-C	T1_57C1	I like scientific TV shows	.409
O 6	BF-C	T1_29O1	I have a great deal of fantasy	.402
O 7	BF-C	T1_48E1	I am able to create new games and entertainments.	.391
O 8	BF-C	T1_64C1	I like to know and learn new things	.341
LC 1	SDQ	T3_18CP1	I am often accused of lying or cheating	.612
LC 2	LOCUS	T4_19LC1	Do you usually feel that it's almost useless to try in school because most other children are just plain smarter than you are?	.598
LC 3	SDQ	T3_22CP1	I take things that are not mine from home, school, or elsewhere	.532
LC 4	LOCUS	T4_11LC1	Do you feel that when a kid your age decides to hit you, there's little you can do to stop him or her?	.452
LC 5	LOCUS	T4_02LC1	Are you often blamed for things that just aren't your fault?	.442
LC 6	LOCUS	T4_10LC1	Do you feel that one of the best ways to handle most problems is just not to think about them?	.410
LC 7	LOCUS	T4_08LC1	Do you feel that when you do something wrong there's very little you can do to make it right?	.366
LC 8	LOCUS	T4_13LC0	Most of the time, do you feel that you can change what might happen tomorrow by what you do today?	-.335

Notes. Factor labels are C = Conscientiousness, ES = Emotional Stability (vs. negative affect), E = Extraversion, A = Agreeableness, O = Open-mindedness; LC = External Locus of Control.

The underrepresentation of this important skill domain in the child instruments is noteworthy.

External Locus of Control/Negative Valence: Ineffective Coping and Hopeless Beliefs

Whereas five of the Brazilian youth factors could be aligned quite well with the Big Five personality domains, we found an additional dimension (4% of the variance) that was defined primarily by a number of highly undesirable, external Locus of Control items, complemented by some SDQ Conduct problem items. These items all indicate feelings of ineffectiveness, distress, and hopelessness and had some secondary loadings on the previous factors (especially low Conscientiousness and Emotional Stability). Youth with high scores indicated that there was little they could do to make things right or stop others treating them badly; that it was useless to even try in school and best to not think about problems; and that they get accused or blamed for things they did not actually do, including lying, cheating, and even stealing things. These negative, even hopeless self-views suggest broadly inadequate coping skills. This dimension can be best labeled External Locus of Control/Negative Valence.

We decided to retain these six factors because they were the most stable and interpretable of all the solutions considered. In fact, even in the solution retaining up to 10 factors, these six factors were still apparent. The additional factors consisted of (1) three introversion items split off from the broader Extraversion factor found in the six-factor solution, (2) a few intellectual skills' items split off from the broader Open-mindedness factor; and (3) two narrow factors split off from the Locus of Control factor. In short, none of these additional factors was large, well defined, or conceptually promising.

Parallels With the Big Five Framework

These results show that the majority of the 209 child and adolescent SEMS items, and five of their underlying factors, could be conceptually aligned and interpreted with the Big Five framework. As can be concluded from the example items provided in Table 2, all Big Five dimensions were clearly represented within the comprehensive SEMS item set. Inspection of the full factor loading matrix further shows that the majority of items included in the SEMS instruments for children had their primary loadings on dimensions that paralleled the Big Five. In other words, Conscientiousness, Emotional Stability, Extraversion, Agreeableness, and Open-mindedness (in that order) turned out to be defining qualities and key dimensions to frame a spectrum of SEMS constructs.

In addition to the Big Five, we found a sixth factor, tapping into the more specific beliefs about self and the future defined by the personality trait of external Locus of Control. This factor was complemented with items referring to conduct problems (being accused of lying, cheating, stealing) and low self-esteem bordering on hopelessness,

representing extremely negative valence in self-perception, an additional dimension described first by Tellegen and Waller (1987).

The analyses of more complex factor solutions at the item level clearly showed that this factor split in smaller factors and was less stable across analyses, whereas EFAs conducted using the 21 individual scales (not reported in this paper) showed that this sixth factor also explained the least amount of variance. Given these unstable characteristics and the possibility that this factor partly represents method variance (i.e., infrequently endorsed items), it is premature to conclude that this sixth factor is a robust factor beyond the Big Five. Nevertheless, it was retained as a key dimension for item representation in the new SENNA 1.0 item pool, to be as comprehensive as possible while safeguarding parsimony.

Development of the SENNA Item Pool

These six factors served as the starting point to develop a new instrument with groups of internally consistent items that marked these six dimensions well, providing a manageable (parsimony) though comprehensive account (bandwidth) of the variance enclosed in the frequently used instruments to assess SEMS. The item selection process was conducted by four independent reviewers taking into account that the final set of items should: (1) represent a broad range of constructs enclosed in the primary eight instruments, (2) have high loadings and variance, (3) represent the positive and negative poles of the descriptive dimensions, (4) come from different instruments whenever possible, (5) have passed the qualitative pilot analysis and the focus group study (without presenting any problems), and (6) add up to a maximum of approximately 90 items. Only items that received approval to be used in the SENNA instrument from the original instrument authors were retained. Closely parallel items were written as replacements for those items for which no permission was granted. The selections were contrasted and reviewed by a four-member team and a final consensus version of 92 items was decided upon for grades 10–12, and a smaller subset of 62 items (from the original 92) was selected for the 5th grade. The distribution of the 92 items across the six factors was 17 items for Conscientiousness, 14 items for Emotional Stability, 14 items for Extraversion, 15 items for Agreeableness, 17 items for Open-mindedness, and 15 items for Locus of Control/Negative Valence. The response format was a Likert rating scale with five points.

Study 2: SENNA Confirmatory and Measurement Invariance Study

The purpose of Study 2 was to evaluate the psychometric properties of this SENNA item set in a large-scale sample in order to test the proposed internal structure using confirmatory factor analysis and to test measurement invariance

Table 3. Summary results of item-factor analysis and factor reliability

Reliability	F1 Cons	F2 Extr	F3 Neur	F4 Agre	F5 Opns	F6 Loc
Alpha	.91	.83	.89	.81	.78	.75
Fit indices	Chi-square	df	n.FreeP	CFI	TLI	RMSEA
CFA	418.777	4,079	475	0.700	0.692	0.064
ESEM	120.588	3,649	905	0.915	0.903	0.036
ESEM factors	F1 Cons	F2 Extr	F3 Neur	F4 Agre	F5 Opns	F6 LoC/NV
Expected factors						
F1	17					
F2		14				
F3			14			
F4				14	1	
F5	1				13	3
F6		2			1	12

Notes. Cons = Conscientiousness; Extr = Extraversion; Neur = Neuroticism; Agre = Agreeableness; Opns = Openness; LoC/NV = Locus of control/Negative Valence; df = Degrees of freedom; n.FreeP = Number of free parameters; CFI = Comparative fit index; TLI = Tucker-Lewis index; RMSEA = Root mean square of error of approximation; CFA = Confirmatory factor analysis; ESEM = Exploratory structural equation modeling; F = Factor.

Table 4. Fit indices of the multiple-group exploratory structural equations modeling assessing invariance across grades

Model	χ^2	$\Delta\chi^2$	df	Δdf	n.FreeP	CFI	TLI	RMSEA	SRMR
Configural	41.413*	–	4,602	–	1,443	0.901	0.878	0.031	0.022
Metric	43.884*	2,471*	5,274	672	771	0.896	0.889	0.030	0.026
Scalar	46.219*	2,335*	5,386	112	659	0.890	0.885	0.030	0.026

Notes. χ^2 = Chi-square; $\Delta\chi^2$ = Chi-square difference test; df = Degrees of freedom; Δdf = Degrees of freedom difference; n.FreeP = Number of free parameters; CFI = Comparative fit index; TLI = Tucker-Lewis index; RMSEA = Root mean square of error of approximation; SRMR = Standardized root mean square residual. * $p < .001$.

of the instrument across grades. Our main hypothesis was that a six-factor structure would explain the covariance among all 92 items and that the 62 items (which were available for kids in all grades) would show measurement invariance across grades.

Method

Participants

Participants were a representative sample of students from the Rio de Janeiro State. The total number of students was $N = 24,605$ (58.2% boys and 41.8% girls). They came from 14 regions, 79 cities, 431 schools, and 1,062 classes (mean N for each class $M = 23$, $SD = 7.6$). With respect to grades, the sample was subdivided as follows: 5th grade: $N = 1,472$ (6%), 10th grade: $N = 14,504$ (58.9%), and 12th grade: $N = 8,629$ (35.1%). The ethnic origin of the students was white: 28.9%, pardo (mulatto): 40.4%, black: 21.1%, and Asian: 5.9%. The average ages in each grade were: 5th: $M = 11.9$ years, $SD = 1.46$; 10th grade: $M = 16.45$ years, $SD = 1.05$ and 12th grade: $M = 18.21$ years, $SD = 1.02$.

Data Analysis

First, we performed an item-level confirmatory factor analysis (CFA), verifying whether it was possible to recover the structure of the six factors as defined in Study 1. We ran an Exploratory Structural Equation Modeling (ESEM) analysis with the Weighted Least Square parameter estimator (WLSMV) using MPLUS (Muthén and Muthén, 2012) to test whether the structure of six latent factors with the same indicators could be recovered. An ESEM approach is an intermediate approach halfway between exploratory approaches and the fully independent confirmatory approach, imposing only a restriction on the number of factors, but leaving their loadings free to be estimated on all extracted factors. This method proved to be more suitable for multidimensional spaces where secondary loadings can be expected and particularly for item-level factor analysis (Marsh et al., 2010). A parallel analysis was performed to investigate the number of factors using simulation of matrices of random correlations with the MPLUS package. Item indicators were assumed to be categorical variables that accommodate non-normality in their distributions. The model was specified to retain six correlated factors using the GEOMIN oblique rotation.

Second, multiple-group confirmatory factor analyses were conducted in the samples from the 5th, 10th, and 12th grades, testing the invariance of the measurement model parameters across these three groups. The basic hypothesis tested in this analysis was whether SENNA 1.0 measured the same factors in the same way across grades. The analysis of measurement invariance was performed with MPLUS using the Exploratory Structural Equation Modeling (ESEM) for Multiple Groups. This analysis tested three increasingly restrictive models: (1) The configural invariance model tested whether the basic six-factor structure model was suitable to represent the underlying structure of the 62-item covariance matrix when calculated separately, based upon samples from different school years. In this model, all the parameters (item-factor loadings and item intercepts – item indicator averages) are freely estimated. Only the general structure (number of factors) is held constant between the groups; (2) The metric invariance model that tested whether the factor loadings of the items, on the factors, varied across grades or not when calculated separately. The additional restrictions imposed here are that factorial loadings of the items on factors are constant across grades. This is equivalent to assuming that an item does not show Differential Item Functioning (DIF) in the discrimination parameter “*a*” according to Item Response Theory (IRT); (3) The scalar invariance model that tested whether the intercepts (averages in the items) were equal when calculated separately across grades (controlling for possible differences in the latent factor among groups). The additional restriction imposed on this model is that, in addition to assuming six factors and equal item loadings, the averages of the items are kept constant across the three groups, but the differences that could possibly exist due to the different averages in the latent constructs are controlled. This model tests whether students with the same average latent factor, in different grades, showed different averages in the items. As a result, this is equivalent to the analysis of DIF under the difficulty parameter “*b*” (in this case, the ease of agreeing with the assertions) in IRT.

Results and Discussion

Table 3 presents the main findings from the first analysis of the internal structure of the new instrument. In the upper part, Table 3 shows internal consistency coefficients of the new Brazilian scales, which were all above .75 and satisfactory. Table 3 also shows model fit indices from the ESEM analysis along with a traditional independent confirmatory factor analysis (CFA) for comparative purposes. The six-factor ESEM model showed good fit using the commonly used benchmarks (CFI and TLI > 0.90, RMSEA < .05). The lower part of Table 3 shows a cross-tabulation of expected item-factor classifications based on Study 1 (rows) with those obtained from ESEM results in Study 2 (columns). In the ESEM analysis, an item was classified in a factor if it had a factor loading greater than or equal to 0.29 and content analysis suggested that it was

theoretically coherent. The numbers presented in the body of the table show the number of items assigned to each of the 6 expected factors found in Study 1 (rows) by their final assignments based on the ESEM analysis. As can be seen, the item-factor organization is particularly well recovered in the ESEM analysis. Only a few items crossed loadings from Study 1 to Study 2.

The second analysis investigated whether we could assume measurement invariance of the instrument for students of different grades verifying if the instrument measures the same factors in the same way in different grades. For this analysis we used the subset of 62 items which had been administered across all grades.

Results in Table 4 show that the restrictions imposed by the more demanding models containing the invariance assumption did not greatly reduce the fit achieved by the base model. The RMSEA and SRMR indexes remained essentially unchanged. The CFI and TLI indexes dropped a little in the metric model and stayed the same with the additional restrictions of the scalar model. This suggests that there might be some changes in item loadings for some items across grades that need to be further explored. But the most important aspect is that there were no large differences in item intercepts (scalar model). Therefore a partially invariant model could be assumed for SENNA 1.0. In summary, these results provide evidence for measurement invariance, suggesting that the new measure is psychometrically equivalent across different grades.

General Discussion

The central goals of the present studies were to examine the factor structure underlying a diverse set of instruments commonly used to assess SEMS, and to construct a new, more comprehensive instrument that is useful, reliable, stable, and measurement-invariant as a self-report measure for education monitoring and evaluation purposes in Brazil. This objective was examined in a separate development sample (Study 1) and a very large replication sample (Study 2).

A first contribution of this work is that it identifies six key dimensions underlying the amalgam of constructs and operationalizations that were used in the past to refer to SEMS. The availability of a robust taxonomic framework defining the core characteristics of a content domain (in this case SEMS) is a sine qua non to comprehensively represent this domain in a new inventory. The present work suggests that the Big Five framework is not necessarily isomorphic with this underlying structure, but provides a compelling approach to structure SEMS constructs and considerably facilitates the interpretation and labeling of the factors. The final retained model included six factors, the familiar five along with an “additional” Locus of Control/Negative Valence factor, not unlike the additional self-evaluative factor Tellegen and Waller (1987) had discovered. In other words, the existence of a sixth factor does not invalidate the Big Five personality taxonomy but supplements it with item content that may be of particular importance to the

SEMS domain (beliefs about personal control versus hopelessness and feeling defeated) that stems from a specific instrument (LoC) and links, in part, to low levels of Emotional Stability.

Considered, from the opposite direction, however, the current study also showed that most single SEMS instruments tapped into a more limited number of dimensions than our new taxonomic framework suggested. GRIT, for example, almost exclusively included items referring to conscientiousness, whereas the BFI and the BF-C, for obvious reasons, included item content for five of the six factors. The joint factor analysis at the item level across different instruments provides a good test whether individual items, even from measures assumed to assess the Big Five, are placed correctly in their instrument scoring system. Here, our analyses of the Brazilian adaptations show that some Openness/Intellect items from the BF-C had content more closely aligned with the Conscientiousness factor, findings that have been previously observed by Mervielde, Buyst, and De Fruyt (1995); conversely, some of the best BF-C Open-mindedness items in our analyses were designed initially for other factors, such as creating games for Extraversion.

From a practical point of view, this work has produced an initial version of an instrument that will be easy-to-administer and robust and that can be used for educational monitoring, social-policy research, and evaluation programs in Brazil. The availability of a 92-item version for older students and a 62-item version for younger age groups is consistent with the attention span typical for each of these targeted age groups. The way SENNA 1.0 was developed, including the extensive pilot testing, further ensures that the items are well understood and align with the target groups' cultural background. Study 2 provided compelling evidence that the structure of SENNA 1.0 was replicable and robust and shows the characteristics required to make comparisons across grades.

The present work is not without its limitations and raised a number of questions to be addressed in future research. A first constraint could be the starting point of the project, that is, the sampling of SEMS instruments relying on the different inclusion criteria. For example, the current review did not include specifically focused measures related to emotion knowledge or awareness, recognition, understanding, and management. The inclusion of other SEMS measures may affect the dimensions to be found, so future research should expand this initial list.

Second, a common methodological weakness of many SES evaluation studies is their reliance on self-reports. Therefore, we encourage studies with informant ratings to examine whether the psychometric properties reported for self-reports also replicate in parent/teacher or peer reports.

A final limitation involves the broad nature of the dimensions specified by SENNA 1.0. The bandwidth-fidelity discussion, or whether it is better to use smaller or broader constructs to compare groups or predict outcomes, is applicable here. Therefore, future research efforts may concentrate on creating facets with one or more of these SENNA 1.0 dimensions to sharpen the specificity to predict when necessary or enhance sensitivity to pick up developmental or targeted change in SEMS. SENNA

1.0 points to the broad underlying qualities of more specific SEM skill sets. The dimension we have here labeled, for historical consistency, as *Extraversion*, for example, refers to a multitude of skills related to active and energetic engagement with others and the world. Teachers and practitioners will be most interested to target their interventions at the more specific skill level; that is, they will be interested in developing their students' public speaking and presentation skills, or their leadership skills, rather than trying to affect the underlying broad dimension of Extraversion. Whereas SENNA 1.0 grasps the common core necessary for structuring the SEMS field, the next version of SENNA will have to be more fine-tuned to the needs of practitioners and teachers in the field.

To conclude, the present work was designed to contribute to building the much-needed SEMS framework and taxonomy and delivered an initial tool to assess 21st century skills in Brazil. The tool will be available for researchers and practitioners, and we encourage to share your experiences with us.

Acknowledgments

This article is part of a research project supported by the Ayrton Senna Foundation. The first author also received a scholarship from the National Council for Scientific and Technological Development (CNPq).

References

- Almund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. D. (2011). Personality psychology and economics. In E. Hanushek, S. Machin, & L. Woessman (Eds.), *Handbook of the Economics of Education* (Vol. 4, pp. 1–181). Amsterdam, The Netherlands: Elsevier.
- Barbaranelli, C., Caprara, G., Rabasca, A., & Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences*, *34*, 645–664.
- Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Adaptação e validação de instrumentos psicológicos entre culturas: algumas considerações [Cross-Cultural Adaptation and Validation of Psychological Instruments: Some Considerations]. *Paidéia (Ribeirão Preto)*, *22*, 423–432. doi: 10.1590/S0103-863X2012000300014
- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, *112*, 155–159.
- De Fruyt, F., & Van Leeuwen, K. (2014). Advancements in the field of personality development. *Journal of Adolescence*, *37*, 763–769. doi: 10.1016/j.adolescence.2014.04.009
- De Fruyt, F., Wille, B., & John, O. P. (2015). Employability in the 21st Century: Complex (Interactive) Problem Solving and Other Essential Skills. *Industrial and Organizational Psychology*, *8*, 276–281. doi: 10.1017/iop.2015.33
- Denissen, J. J. A. (2014). A Roadmap for Further Progress in Research on Personality Development. *European Journal of Personality*, *28*, 213–215. doi: 10.1002/per.1967
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, *91*, 166–174.
- 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* (pp. 405–432). doi: 10.1111/j.1467-8624.2010.01564.x
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586.
- John, O. P., & De Fruyt, F. (2014, September). *An integrative framework for the assessment of social-emotional skills*. Lecture presented at the Institute Ayrton Senna, Sao Paulo.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory – Versions 4a and 54*. Berkeley, CA: Institute of Personality and Social Research, University of California.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and research* (pp. 114–158). New York, NY: Guilford Press.
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The core self-evaluations scale: Development of a measure. *Personnel Psychology*, 56, 303–331. doi: 10.1111/j.1744-6570.2003.tb00152.x
- MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences*, 19, 451–458. doi: 10.1016/j.lindif.2009.03.007
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, 22, 471.
- Mervielde, I., Buyst, V., & De Fruyt, F. (1995). The validity of the Big-5 as a model for teachers ratings of individual-differences among children aged 4–12 years. *Personality and Individual Differences*, 18, 525–534.
- Muris, P. (2001). A brief questionnaire for measuring self-efficacy in youths. *Journal of Psychopathology and Behavioral Assessment*, 23, 145–149.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Nowicki, S., & Strickland, B. R. (1973). A locus of control scale for children. *Journal of Consulting and Clinical Psychology*, 40, 148–154. doi: 10.1037/h0033978
- OECD. (2015). *Draft proposal: OECD Longitudinal Study of Skills Development in Cities*. Paris, France: OECD.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135, 322–338. doi: 10.1037/a0014996
- Rosenberg, M. (1979). *Conceiving the Self*. New York, NY: Basic Books.
- Sailer, O. (2005). crossdes: A package for design and randomization in crossover studies. *Rnews*, 5, 24–27.
- Santos, D. (2011). *The socio-economic importance of personality characteristics*. Retrieved from <http://educacaosec21.org.br/wp-content/uploads/2013/08/The-socio-economic-importance-of-personality-characteristics.pdf>
- Santos, D., & Primi, R. (2014). *Social and emotional development and school earning: a measurement proposal in support of public policy*. São Paulo, Brazil: Ayrton Senna Institute.
- Soto, C. J., & John, O. P. (2014). Traits in transition: The structure of parent-reported personality traits from early childhood to early adulthood. *Journal of Personality*, 82, 182–199. doi: 10.1111/jopy.12044
- 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, 330–348.
- Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, 84, 1041–1053.
- Tellegen, A., & Waller, N. G. (1987). *Reexamining basic dimensions of natural language trait descriptors*. Paper presented at the 95th Annual Meeting of the American Psychological Association, New York.
- Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. San Francisco, CA: Jossey-Bass.

Received: November 4, 2014

Revision received: May 1, 2015

Date of acceptance: October 15, 2015

Published online: April 22, 2016

Ricardo Primi

Universidade São Francisco
 Laboratory of Psychological and Educational Assessment
 Rua Alexandre Rodrigues Barbosa, 45
 CEP 13251-900
 Itatiba, São Paulo
 Brazil
 E-mail rprimi@mac.com

Appendix

The Balanced Incomplete Block Design

Booklets for 5th and 6th grades

Booklet	Instrument 1	Instrument 2	# of items
1*	BF-C	BFI	*109
2	BF-C	SDQ	90
3	BF-C	Locus	86
4	BF-C	Self-ef	89
5**	BF-C	Rosenberg/Core**	83
6	BF-C	Grit	75
7	BFI	SDQ	69
8	BFI	Locus	65
9	BFI	SEQ-C	68
10	BFI	Rosenberg/Core	62
11	BFI	Grit	54
12	SDQ	Locus	46
13	SDQ	SEQ-C	49
14	SDQ	Rosenberg/Core	43
15	SDQ	Grit	35
16	Locus	SEQ-C	45
17	Locus	Rosenberg/Core	39
18	Locus	Grit	31
19	SEQ-C	Rosenberg/Core	42
20	SEQ-C	Grit	34
21	Grit	Rosenberg/Core	28

Notes. *Since this booklet had too many items with two instruments measuring the same constructs we did not administer Booklet No. 1. **Rosen/Core: This means that if the sample was comprised of students of 4th and 6th grades Rosenberg's self-esteem scale was applied. If the sample was comprised of students of 10th and 12th grades Core was applied.

Booklets for 9th, 10th, and 12th grades

Booklet	Instrument 1	Instrument 2	Instrument 3	# of items
1	BFI	SEQ-C	Rosen/Core	86
2	BFC	SDQ	SEQ-C	114
3	BFI	SDQ	Locus	90
4	BFC	Locus	Rosen/Core	104
5	SDQ	Rosen/Core	Grit	53
6	BFC	BFI	Grit	119
7	Locus	SEQ-C	Grit	55