

Developing a new tool for international youth programs: the YouthPower Action Youth Soft Skills Assessment (YAYSSA)

Carina Omoeva¹

Nina Menezes Cunha¹

Patrick Kyllonen²

Sarah Gates³

Andres Martinez¹

Holly M. Burke¹

¹FHI360, Washington D.C.

²Educational Testing Service, Princeton, NJ

³Statistical Consultant, Washington D.C.

Abstract

We developed and evaluated the YouthPower Action Youth Soft Skills Assessment (YAYSSA), a self-report soft-skills measure. YAYSSA targets 15- to 19-year-old youth in lower resource environments. In Study 1, we identified 16 key constructs based on a review of those associated with positive youth outcomes in sexual and reproductive health, violence prevention, and workforce success. We adapted promising items measuring those constructs from existing and openly available tools. We conducted cognitive interviews with 50 youth from 6 schools in Uganda, for wording and response formats, leading to a first draft tool. In Study 2 we administered that tool to N = 1,098 youth in 59 schools in Uganda. Confirmatory factor analyses did not support the hypothesized 16-factor structure, but exploratory factor analyses suggested a four-factor solution (Positive self-concept, Higher order thinking skills, Social and Communication skills, and Negative self-concept). In Study 3, a revised tool was administered to Uganda youth (N = 1,010, 59 sites) and, after cognitive testing with 45 youth in Guatemala, administered to youth (N = 794; 59 sites) once, then 5 months later, with a mixture of retested and new participants (N = 784; 67 sites). Factor analytic results supported the four-factor structure with 48 retained items and indicated that the instrument was reliable by internal consistency and test-retest correlations. The instrument correlated with demographic variables and outcomes in expected directions. We found evidence for measurement invariance across country, gender, and socio-economic status. We discuss implications for scale validation and use in future research.

Keywords: youth soft skills, measurement, adolescent development, scale development, scale validation

Evidence across fields and disciplines highlights the importance of “soft” or “life” skills to long-term education, employment, health, and violence prevention outcomes (Deming 2017; Almlund et al., 2011; Heckman et al., 2006; Carneiro et al. 2007). These skills are also referred to as “21st century skills” or “socio-emotional skills” and include broad non-academic constructs such as positive self-concept or self-efficacy, problem solving and decision-making, communication and social skills, as well as emotional regulation and the ability to exercise self-control. The growing recognition of the centrality of these soft skills for a range of life outcomes for a young person has led to increased investments in school-based, out-of-school, and workplace-based programs and activities that promote developing soft skills among adolescents and young adults. Internationally, aid and development agencies such as USAID, the World Bank, and UNICEF are increasingly turning their attention to youth soft and life skills, calling for greater levels of programming for youth and greater focus on skill building across training programs and curricula. As investments grow, so does interest in measurement instruments and metrics that can help quantify levels of soft skill competencies and register differences in soft skills across groups or over time.

Research over the past decade has identified promises and pitfalls associated with a range of soft skills measurement approaches (Abrahams et al., 2019; Duckworth & Yeager, 2015; Galloway et al., 2017; Kautz & Moore, 2018; Soland et al., 2013; Stecher & Hamilton, 2014). The most widely used method for measuring soft skills, self-reported measures, while relatively inexpensive and easy to use, present several potential sources of error that come from response-style effects, reference group bias, and social desirability bias (Abrahams et al., 2019; Duckworth & Yeager, 2015; Kyllonen, 2015). Moreover, because measures of soft skills intend to capture a respondent’s particular behavior, they may reflect factors indicated by the behavior not related to the skill (Kautz & Moore, 2018). For example, a lack of financial and social resources may constrain individuals’ self-regulation skills (Mullainathan & Shafir, 2013). Given that most soft skills measures have been developed, tested, and

validated in developed countries, they likely reflect certain assumptions about individuals' access to resources.

Several innovations have been proposed in response to the potential error introduced by Likert-style self-reports. Kyllonen (2015) identifies forced-choice (or ranking) methods, anchoring vignettes, situational judgment tests, and performance tasks as offering potential solutions. Forced choice methods, by presenting youth with two or more equally desirable options, and situational judgment tests, by presenting youth with scenarios and a series of choices on how best to respond, may reduce the likelihood of faking compared to rating scales. Anchoring vignettes require respondents to rate hypothetical situations or people on the same response scale on which respondents rate themselves; this allows adjusting (rescaling) individual responses with respect to the anchors, potentially mitigating response bias of unadjusted self-ratings. Task-based measures, such as games (Alaoui & Fons-Rosen, 2021), simulations (Hao et al., 2019; Martin-Raugh et al., 2020), or real-effort tasks (Charness et al., 2018) may help to reduce faking and the biases that may arise through self-reports, but also cost more and take more time to administer. These methods can also be plagued by inconsistencies across raters, tasks, and youth performance (Soland et al., 2013).

Some authors propose the use of multiple methods and respondents to minimize the weakness of self-reports alone. Ratings by others, such as program staff or teachers, might be assessed through simple questionnaires or through more complex methods such as observational assessments. Evidence indicates that, on average, ratings by others have more predictive validity for educational and job success than self-ratings (Connelly & Ones, 2010; Oh et al., 2011). However, these ratings are susceptible to *halo* and *horn* effects, whereby observers rate youth overall negatively or positively rather than providing a nuanced assessment. Researchers have also explored the use of proxy behaviors, such as academic performance, participation in student activities, and other observations by teachers and school administrators to complement achievement tests when evaluating students and

schools, with some evidence indicating that this is a promising approach (Heckman & Kautz, 2013). Soland et al. (2019) propose viewing survey meta-data as a source of performance task measures, finding that two survey behaviors, careless answering and item non-response rates, demonstrate low but significant correlations with various skills and personality factors (Barry & Finney, 2016; Zamarro, et al., 2018) as well as education outcomes (Hitt et al., 2016).

All of these approaches are potentially useful and may supplement or replace traditional rating scale measures. However, rating scales still have advantages: they are widely used and understood both from a respondent perspective and from a psychometrics perspective; they are inexpensive to develop, administer, and score; and with low stakes uses, as proposed here, some of the biases (e.g., social desirability) are not as problematic as they would be if the test were administered in a high-stakes setting. Therefore, we chose to develop a rating scale approach with the understanding that this would be more easily accessible and deployable by the target populations. In this study we supplement this measure with anchoring vignettes and several proxy measures.

Need for a Soft Skills Tool for Evaluating Youth Outcomes

The growth in soft skills-focused interventions has resulted in an urgent need among youth development programs for soft skill measures that can be used for program implementation and evaluation. However, currently, there is no single existing instrument that: encompasses key skills identified as important to multiple youth outcomes—workforce, violence prevention, and sexual and reproductive health (SRH); is suitable for measuring change over time in skill levels among youth in an international development program context operating in a low-to-middle income country; and meets other key criteria for use by youth programs (including ease of administration, validity, and reliability).

Building on the recent measurement advancements we sought to develop an instrument to measure key soft skills predictive of positive youth outcomes: youth employability, sexual and reproductive health (SRH), and violence prevention as well as be (a) targeted to youth ages 15-19; (b)

low cost and easy to administer, (c) sensitive to change over time, and (d) consistent across cultural contexts. This paper presents the development and validation of the YouthPower Action Youth Soft Skills Assessment (YAYSSA), a new instrument to measure youth soft skills in low-resource environments.

Development of the YAYSSA tool

Several steps were taken to construct and test the YAYSSA tool. We identified potential promising items from existing and openly available tools; we adapted items that were hypothesized to measure the constructs of interest; we carried out qualitative cognitive interviewing and revised the draft instrument accordingly; we administered pilot versions of the assessment in two countries across two points in time, and we carried out analyses examining the validity of its scale structure and predictive power as related to key youth outcomes in employment, sexual and reproductive health, and violence prevention. The YAYSSA validation took place in two lower resource country contexts, with youth programs in Uganda and Guatemala. The programs were selected because they offered geographic and cultural diversity and were willing to participate in the research.

Initial instrument development followed a conceptual framework drawn from the literature on key skills crucial for cross-sectoral youth outcomes (Gates et al., 2016; Lippman et al., 2015). Two literature reviews were carried out in sequence as part of a series of investments by USAID. The first review (Lippman et al., 2015) analyzed the literature on youth soft skills and workforce success. Following the first review's conceptual framework and using comparable methods, the second review (Gates et al., 2016) analyzed the literature on youth soft skills and violence prevention and sexual and reproductive health (SRH), incorporating results from the first review. Taken together, the reviews identified a common set of cross-cutting skills that foster positive youth outcomes across the domains of sexual and reproductive health, violence prevention, and workforce success: positive self-concept, self-control, and higher-order thinking skills.

Based on the first literature review, Lippman et al. (2015) created definitions for the soft skills terms. Gates et al. (2016) revised these definitions, incorporating findings from the violence prevention and sexual and reproductive health literature as needed.

Positive self-concept is defined as “a realistic awareness of oneself and one’s abilities that reflects an understanding of his/her strengths and potential (and hence, is, positive)” (Lippman et al., 2015, p. 84). Positive self-concept in this framework includes such constructs as self-efficacy, self-esteem, self-confidence, self-awareness, and self-belief (five sub-scales).

Self-control refers to one’s ability to “delay gratification, control impulses, direct and focus attention, manage emotions, and regulate behaviors” (Lippman et al., 2015, p. 81). Self-control includes aspects of delayed gratification, impulse control, attention, emotional regulation, behavior regulation, and thrill-seeking behavior (six sub-scales).

Higher-order thinking skills are skills encompassing problem solving, critical thinking, and decision making (Lippman et al., 2015, p. 77). These skills may reflect the same underlying skill set, which is the ability to take in information from multiple sources, identify the issue(s), evaluate potential options, and reach an appropriate conclusion (Stein, 2000). Higher-order thinking skills includes problem solving, critical thinking, and decision making (three sub-scales).

In addition to the three groups of skills noted above, *social skills and communication skills* were included in tool development and validation, bringing the total number of sub-scales to sixteen. Figure 1 illustrates the conceptual framework employed for the initial construction of the soft skills assessment.

Figure 1

Key Skills for Cross-Sectoral Youth Development: Top Supported Skills Across Fields



Our first set of hypotheses relates to the psychometric qualities of the YAYSSA instrument. First, we propose that we can identify the dimensions drawn from the literature review as depicted in Figure 1. Second, we propose that the dimensions we identify can be measured reliably with rating scale items. Third, we propose that we can find measurement equivalence across country and sex subgroups, as well as longitudinally. Third, we propose that the supplementary methods we develop, anchoring vignettes, and proxy measures will be correlated with the rating scale measures and may add to the prediction of outcomes beyond the rating scale measures due to their being less susceptible to rating scale biases.

Our second set of hypotheses relates to the relationships between YAYSSA scales and youth outcomes. Our theoretical framework posits that the soft skills are predictive of youth outcomes such as sexual and reproductive health, employment, and conflict resolution without violence. In line with this framework, we hypothesized that youth coming from lower socioeconomic backgrounds and experiencing greater economic challenges would demonstrate lower levels of positive self-concept. We also hypothesized that youth with higher social and communication skills would be more likely to be employed or be engaged in an income-generating activity. Finally, the framework proposes that youth

with stronger soft skills would be more likely to engage in non-violent conflict resolution, and more likely to make safer sexual and reproductive health choices.

Study 1: Instrument Development and Cognitive Interviews

This study concerns the initial development of the YAYSSA instrument and cognitive interview studies based on that instrument. We assembled items from sources, then conducted a cognitive interview study by administering the instrument to a small group of Uganda youth to get feedback.

Method

Instrument Item Development

As an initial step to develop the tool, the research team constructed a repository of more than 300 existing assessments that focus on measuring soft skills or life skills, paying particular attention to instruments that were applied outside high-income economies and with youth 15 or more years old (Galloway et al., 2017).¹ Drawing from this repository the initial version of the youth assessment included items that measured the sixteen sub-skills or sub-domains of skills that fall within the conceptual framework described in the Introduction. The items from the repository were revised to simplify wording and adapt to the country context. In the process of adaptation, we also removed references to items and situations that were not readily available to most beneficiaries of USAID youth programs, as well as items referencing school and being a student to make the tool more broadly applicable.

The initial version of the YAYSSA tool can be found in the first column of Table A. 1 in the Supplementary Materials. Below we describe the sources from which we adapted the tool and provide a sample item from the initial version (see Table A. 1). Many of the original source tools provide several scales, so we used some of the same original source tools for multiple subscales.

¹ More information on the repository and the tools included it in is available online at YouthPower.org.

Positive Self-Concept: Self-efficacy (6 items). This scale was adapted from the California Healthy Kids Survey: Social and Emotional Health Module, a 33-item strength-based assessment intended for children in grades 5 to 12 (e.g., “How often do you believe that: you can do most things if you try?”). The survey measures empathy, problem solving, self-efficacy, self-awareness, cooperation and communication, goals and aspirations, and resilience (Furlong et al., 2014).

Positive Self-Concept: Self-esteem (5 items). This scale was adapted from the Jamaica Youth Survey, a survey designed to assess five core competencies that Gardner et al. (2011) found were linked to adjustment and youth problem behaviors, namely, positive sense of self, self-control, decision-making skills, moral system of belief, and prosocial connectedness (e.g., “How often have you felt that: you have a number of good characteristics?”). The survey was developed for use with teenage boys and girls (ages 12-18) in urban Jamaica to evaluate the impact of youth development and violence prevention programs (Gardner et al., 2011).

Positive Self-Concept: Self-confidence and self-awareness (5 items). This scale was adapted from the Child and Adolescent Wellness Scale, a 150-item scale divided into 10 dimensions: Adaptability, Connectedness, Conscientiousness, Emotional self-regulation, Empathy, Initiative, Mindfulness, Optimism, Self-efficacy, and Social competence (e.g., “How often in the past 7 days have you felt sure of yourself?” and “How often do you what you are feeling in the moment?”). The Child and Adolescent Wellness scale was originally validated with students in grades 6 to 12 in the Rocky Mountain region of the United States (Copeland et al., 2010).

Positive Self-Concept: Self-belief (4 items). This scale was adapted from the Chinese Positive Youth Development Scale, a 90-item scale that assesses 15 aspects of positive youth development, including beliefs in the future and clear and positive identity (e.g., “How often do you see that your future will be pleasant/happy?”). The scale was originally validated in 2007 with 322 Chinese youth from Hong Kong (Shek et al., 2007).

Self-Control: Delayed gratification (6 items). This scale was adapted from the California Healthy Kids Survey: Social and Emotional Health Module and the Jamaica Youth Survey (e.g., “How often in the past 7 days have you been unable to wait for something you want?”).

Self-Control: Impulse control (4 items). This scale was adapted from the Jamaica Youth Survey, the Strengths and Difficulties Questionnaire for youth (SDQ), and the Strengths and Difficulties Questionnaire for parents and teachers (SDQ PorT) (e.g., “How often do you think before you do things?”). The SDQ is a 25-item self-report survey for youth ages 11-18 that measures emotional problems, conduct problems, hyperactivity, peer problems, and prosocial behavior (Goodman et al., 1998). The SDQ PorT is a 25-item report that asks observers, such as parents and teachers, to rate youth on the same qualities that the SDQ assesses (Goodman, 1997).

Self-Control: Direct and focus attention (5 items). This scale was adapted from the SDQ, the SDQ PorT, and the Conscientiousness Facets Tool, which assesses a set of conscientiousness facets—industriousness, perfectionism, tidiness, procrastination, refrainment, control, caution, task planning, and perseverance (e.g., “How often in the past 7 days have you had difficulty starting tasks?”). The tool was validated among students ages 13 to 19 from a large private high school on the East Coast of the U.S. (MacCann et al., 2009).

Self-Control: Managing emotions (4 items). This scale was adapted from the Jamaica Youth Survey; the Child and Adolescent Wellness Scale; the Social Emotional Skills, School Quality Improvement Index (SQII) developed by the California Office to Reform Education (CORE); and the Passport to Success Trainer Observation Tool (e.g., “How often in the past 7 days have you done things to calm down when you were angry?”). The SQII is a 29-item survey designed to help teachers and school administrators assess grade 5-12 students’ growth mindset, self-efficacy, self-management, and social awareness (CORE Districts, n.d.). The Passport to Success Trainer Observation Tool is an observational assessment designed to assess the skills of youth who complete the International Youth

Foundation's Passport to Success life skills program. Specifically, the tool assesses time management/punctuality, communication skills and listening skills, problem solving/ conflict management, emotional management, response to authority, responsibility, ability to get along well with others, and ability to adapt to change/adaptability (International Youth Foundation, n.d.).

Self-Control: Regulate behaviors (5 items). This scale was adapted from the Child and Adolescent Wellness Scale, the SQII, the Passport to Success Trainer Observation Tool, and the Personal Potential Index (PPI) (e.g., "How often were you able to stop yourself when you were going to do something you would regret?"). The PPI is a web-based admissions tool developed by Educational Testing Service (ETS) to assess core personal attributes of undergraduate students that graduate, and business school deans have identified as important factors for success in graduate and professional school. Specifically, the PPI assesses creativity, communication, teamwork, resilience, planning and organization, and ethics and integrity (Educational Testing Service, 2014).

Self-Control: Thrill-seeking and risk-taking. (4 items). This scale was adapted from the Communities that Care Survey, a self-reported questionnaire for youth that measures risk and protective factors that predict substance use, delinquency, and related problem behaviors among youth 11-18 years old items (e.g., "How often have you done what feels good no matter what?"). The tool was originally validated with children and youth in grades 5 to 10 in the U.S. (Arthur et al., 2002).

Higher-Order Thinking Skills (HOTS): Problem solving (5 items). This scale was adapted from the Chinese Positive Youth Development Scale and Responses to Stress Questionnaire (e.g., "How often did you develop a plan to solve the problem?"). The Responses to Stress Questionnaire comprises 57 self-reported items that measure coping and involuntary stress responses to a range of stressful situations among adolescents and children over 9 years old (Connor-Smith et al., 2000). We adapted items from a version of the Responses to Stress Questionnaire that assess responses to peer stress. The

tool was originally validated among adolescents 16-19 and 12-18 years old in New England in the U.S (Connor-Smith et al., 2000).

Higher-Order Thinking Skills (HOTS): Critical thinking (5 items). This scale was adapted from the Chinese Positive Youth Development Scale, the Child and Adolescent Wellness Scale, the PISA 2015 Problem Solving Experiences Scale (section D of the PISA student questionnaire), and the Optional Work Readiness Tool for the Department of Labor's WIA Youth Program (e.g., "When answering these next three questions, think about the last few times someone told you an interesting story. How often did you separate the true and false parts of the story?"). The PISA Problem Solving Experiences Scale was developed by the Organisation for Economic Co-operation and Development (OECD) to assess the drive and motivation among 15 years old cross-nationally, including perseverance and openness to problem solving (OECD, 2014). The Optional Work Readiness Tool is an observational evaluation tool intended for use by young adults' supervisors on the job. It measures 10 foundational skills, including problem-solving/critical thinking (U. S. Department of Labor, n.d.).

Higher-Order Thinking Skills (HOTS): Decision making (4 items). This scale was adapted from the Jamaica Youth Survey (Gardner et al., 2011) (e.g., "When answering these next three questions, think about the last major decisions you made. Please tell me how often you did the following things before making the decisions. How often did you collect a lot of information before making the decisions?").

Social Skills (8 items). This scale was adapted from the adolescent-reported items from Child Trends' Flourishing Children Survey, Social Competence Scale (e.g., "How often do you do your part when you work in a group?"). The Social Competence Scale measures "a set of positive skills necessary to get along well with others and function constructively in groups, including: respecting and expressing appreciation for others; being able to work well with others, present ideas and listen to others' ideas, and work and cooperate in heterogeneous groups; demonstrating context-appropriate behavior and the

ability to behave according to social norms; and using a range of skills or processes aimed at resolving conflict” (Lippman et al., 2014, p. 12). It was validated among children ages 12-17 in the U.S. (Lippman et al., 2014).

Communication Skills (7 items). This scale was adapted from the Senna 2.0 survey and the Responses to Stress Questionnaire (e.g., “How often can you discuss a problem with a friend without making things worse?”). The Senna 2.0 survey, which expands on the Senna 1.0 survey, is used to evaluate social and emotional skills in the school environment. The final tool is a six-factor structure that resembles the Big Five personality dimensions, plus a negative self-evaluation factor. The Senna 2.0 tool has been validated with youth between ages 12 and 19 in Brazil (Primi et al., 2016a).

Response Options.

We tested four response options formats, as shown by Figure A. 1 in the Supplementary Materials. The first three response options were designed as a 5-point behavioral frequency scale that asked youth to assess “how often” they acted out a behavior associated with a particular skill (e.g., “How often do you think things through before you do them?”), where the first option only showed “Almost Never” and “Almost Always” in the two extremes of a continuous line, with the numbers 1 and 5 on top of it, respectively; the second option added the numbers 2, 3 and 4 along the line, and the third option contained “1. Almost Never” / “2. Rarely” / “3. Sometimes” / “4. Often” / “5. Almost Always” displayed along a continuous line (see Figure A. 1). The fourth option was a 5-point scale containing “1. Not at all like you”, “2. A little like you”, “3. Somewhat like you”, “4. A lot like you”, and “5. Exactly like you” displayed along a continuous line. Frequency options were expected to provide greater precision and clarity in the responses compared to agreement options. The initial version of the tool also included a group of “importance items,” one for each scale, intended to assess how youth valued each of the measured skills (e.g., “How important is this to you: Being confident in yourself?”), with the hypothesis that youth who value a certain skill were more likely to work to improve that skill. Response options

were “Extremely Important” / “Very Important” / “Important” / “A Little Important” / “Not at all Important”.

Participants

Cognitive interviews were conducted with fifty youth randomly selected from each of 6 schools located in the cities of Kampala, Jinja, and Gulu, with schools randomly selected from the sample of schools participating in the Educate! Program in Uganda. The program is a non-US Government program for youth in secondary schools².

Procedure

Five trained interviewers administered cognitive interviews to participants from October to November 2017. Items were presented one at a time, and respondents were asked to describe what they thought was the meaning of the items. Interviewers also asked about the response categories and how respondents interpreted them. Notes of the sessions were recorded and analyzed later by the research team. Further, there was a pretest of the youth tool with 23 youth at six schools (three urban and three rural).

Human Subjects Approval. The study was approved by FHI 360's Protection of Human Subjects Committee. In Uganda, this study was approved by the AIDS Support Organization (TASO) and the Uganda National Council for Science and Technology.

Results

In most schools, comprehension was good. Participants seemed to understand about 75-80% of the items. In one school, comprehension was extremely low. We learned that showing a visual of the scale seemed to help the tool administration go more quickly and we use this technique in all

² Educate! is an after-school soft skills development program sponsored by non-government donors that provides entrepreneurship and life skills training for youth. Educate! has operated in Uganda since 2010 and gradually expanded to nearly 600 schools. Their model integrates training modules as after-school activities for a select group of eligible youth (called scholars).

subsequently data collection. Data from the pretest showed respondents took one hour on average to complete the instrument.

Most participants preferred the 5-point scale containing “1. Almost Never”, “2. Rarely”, “3. Sometimes”, “4. Often”, and “5. Almost Always”, explaining that having the words and number labels helped them to understand the meaning of the response options. We revised option 1 to “1. Never or Almost Never” and option 5 to “5. Almost Always or Always”. Some participants said they preferred the 5-point scale containing “1. Not at all like you”, “2. A little like you”, “3. Somewhat like you”, “4. A lot like you”, and “5. Exactly like you”, but their responses to the questions that used this scale indicated that they did not comprehend the response options. Most participants had trouble using this scale correctly either because they could not identify if “you” referred to the interviewer or the respondent or because they did not understand what the options meant overall.

The interviews also revealed some cultural values/norms that might influence participant responses. We noticed an emphasis—especially for girls—on following the rules, being “good,” and “keeping quiet,” even when someone was being abusive. We theorized that this might bias respondents’ answers toward the extreme for some of the self-control questions. We also noticed that participants placed high value on being in school and learning. Some participants used “going to school” as an example of “what feels good” for the question “how often do you do what feels good no matter what?”, an item that was intended to assess their behavioral regulation skills.

Several participants understood questions about money or gifts differently from what we had intended. Participants conceptualized “saving” in a very different way—they explained that they save for immediate needs for school and for their health, like pens, paper, shoes, and, for girls, menstrual pads. Most participants did not think about waiting for something in terms of delayed gratification, but rather as foolish, explaining that they did not think it would be wise to wait for something that may never come.

Finally, we noticed a few word choice problems. Comprehension issues came up with words such as demonstrate, directions, shove, and pleasant, among others. In response, we revised the instrument to simplify the language and use participants' preferred terms. Then, we revised items that did not "translate" as we had intended, with a focus on items measuring delayed gratification, emotional regulation, and communication skills.

Discussion

The qualitative cognitive interviewing was carried out as preparation for the full pilot administration in Study 2. The qualitative cognitive interviews served to assess the content validity of the items, various response options, and the overall structure of the tool; and to ensure a smooth process for the Study 2 including electronic data capture and analysis. This process yielded useful lessons about the tool, relating to participants' preferences on the response options and participants' understanding of the items—in particular, how their context affected their interpretation of the items. As a result of these lessons, the team finalized response options, revised item wording, revised the content of some questions (for example, by adding more specific examples), and dropped some questions that were not relevant in the context.

Study 2: Uganda Baseline Study

This study involved the administration of the revised instrument to a larger sample of Uganda participants to determine the psychometric properties of the assessment, which was slightly revised after Study 1.

Method

YAYSSA Instrument

Based on participant feedback and results from the cognitive interview study (Study 1) we revised the items for this study, as described previously. Supplementary materials (Table A. 1) shows how items changed from those initially developed.

Participants

The tool was administered in March 2018 to 1,098 youth. Fifty-nine schools were sampled for the study, with an average of 19 youth sampled per school, including both Educate! scholars and age-eligible non-scholars. Schools were randomly selected from a list of Educate! participant schools; youth between 15 and 20 years old were surveyed (average age 17), randomly selected within each school. Of the youth sampled, 53% were female; 86% took the survey in English (as opposed to local languages), 20% reported being sexually active, 98% were in school full time, 19% were getting job training, 15% were working full or part time, and 15% were looking for jobs.

Procedure

YAYSSA Administration. The YAYSSA tool was administered by tablet in a one-on-one interview format, and the duration of the interview was 50 minutes on average. The assessment was conducted by 30 enumerators at baseline between March 12 – March 28 of 2018. The tool was translated into eight local languages, to ensure full comprehension by the youth.

Enumerator Training. Prior to the baseline data collection, a two-person team conducted a five-day training for the enumerators on the tool structure, study design, and study administration, including recruitment and sampling. The enumerators were provided with Samsung tablet with the electronic version of the instrument. Enumerators were also trained on research ethics in accordance with approved protocol.

Study Coordination. For each round of data collection, study coordinators contacted the school or program site ahead of time to arrange the interview; upon arrival at each school or program site, a team lead first met with the head teacher or program coordinator, who would gather all eligible youth for random selection. Before each interview, the enumerator explained the purpose of the study and the interview and asked for the participant's informed consent. The requirement for parental consent was waived by all review boards due to the low-risk nature of the study. After providing informed consent, the youth assessment was administered in a face-to-face interview form as a survey, with behavioral outcomes offered for self-administration by youth on a tablet.

Human Subjects Approval. The study was approved by FHI 360's Protection of Human Subjects Committee. The Uganda study was approved by the AIDS Support Organization (TASO) and the Uganda National Council for Science and Technology.

Results

Given that item development arose from a hypothesized factor structure we first tested the hypothesized scale structure using a confirmatory factor analysis (CFA) modeling approach.

Confirmatory Factor Analysis

We performed CFA to test whether the data fit the hypothesized measurement model, with ordinal response scales. For model fitting with categorical data, different estimators can be used (Muthén et al., 2015); here we used the diagonally weighted least squares estimator (WLSMV in Mplus 8). Factor variances were set to 1 in the estimations to identify the model and provide a scale for the latent factor variables. Factors were allowed to freely intercorrelate. Indicator items were allowed to load freely on one factor only (according to the hypothesized structure) and the model included independent item error terms for the indicator items. The statistics employed to evaluate model fit were Chi-square, Tucker-Lewis Index (TLI), comparative fit index (CFI), root mean square residual fit index (RMSEA), and the standardized root mean square residual (SRMR) (Bollen 1989, Hoyle 1995, Hu &

Bentler 1999). Because the Chi-square is sensitive to large sample sizes and may reject well-fitting models, our model fit assessment placed more emphasis on the other statistics. Following the recommendations by Hu and Bentler (1999), we adopted the following cut-off values as a guide for establishing if the models fit the data well: SRMR ideally below 0.08 and at most 0.10; RMSEA ideally below 0.06 and at most 0.10; and CFI and TLI ideally above 0.95, with a minimum of 0.90.

We fit the CFA model described above to the data from Uganda Baseline (N=1098), with factors being the sixteen sub-scales, each with at least three or four indicator items (79 items total)³. Table 1 shows that the model fit the data poorly, based both on the CFI (0.804) and TLI (0.791) being below the acceptable thresholds.⁴ The model returned 126 modification indexes with value above 10.⁵

Table 1

CFA Goodness-of-fit indicators Study 1

	Uganda Baseline
Chi Sq.	7369.004
df	2882
p > Chi Sq.	<0.001
CFI	0.804
TLI	0.791
SRMR	0.058
RMSEA	0.038
90% CI, lower bound	0.037
upper bound	0.077
Observations	1098

Note: CFA ran using WLSMV estimator in Mplus 8 with factor variances set to 1. Following Hu & Bentler (1999), we adopt the following cut-off values to determine if the model fits the data well: SRMR ideally below 0.08 and at most 0.10; RMSEA ideally below 0.06 and at most 0.10; and CFI and TLI ideally above 0.95, with a minimum of 0.90.

³ The 16 sub-scales are: self-efficacy, self-esteem, self-confidence, self-awareness, self-believe, gratification, impulses, attention, emotions, regulate behaviors, thrill seeking, problem solving, critical thinking, decisions, social skills, and communication skills.

⁴ CFA loadings are shown in Table A. 2 in the Supplementary Materials.

⁵ We also ran another specification of the model defining five scales (positive self-confidence, self-control, higher-order thinking skills, social skills, and communication skills), which also returned inadequate fit (CFA=0.467, TLI=0.452, SRMR=0.093, RMSEA=0.061) and 171 modification indexes with value above 10.

Exploratory Factor Analysis

Given the inadequate results from the CFA, we conducted an exploratory factor analysis (EFA) to gain insight into the structure of the response data from the YASSA tool. We conducted an EFA treating the response data as ordinal categorical using the diagonally weighted least squares estimator (WLSMV in Mplus 8) and oblique (geomin) rotation. To determine the number of factors to extract for rotation, we computed solutions for 1 to 20 factors, then compared solutions with respect to Kaiser's (eigenvalue greater than one) criterion, Cattell's scree test, Horn's (1965) parallel analysis, Bayesian Information criterion, and factor interpretability. Eigenvalues for the Kaiser's criterion and Cattell scree test were obtained in Stata/SE 15, using principal axis factoring. We conducted two versions of the Parallel analysis, one based on eigenvalues from the observed correlation matrix (using Mplus 8); and the other based on the polychoric correlation matrix (using the software FACTOR (<http://psico.fcep.urv.es/utilitats/factor/>, Lorenzo-Seva & Ferrando, 2006, 88-91).

The best fitting solution considering these criteria was a 4-factor solution, accounting for 80% of the variance.⁶ The eigenvalue for factor 1 was 9.2 and accounted for 50% of the variance; the second factor presented an eigenvalue of 2.5 and accounted for 13.6% of the variance; and factor 3 and 4 presented eigenvalues of 1.6 and 1.3 and accounted for 8.9% and 7.0% of the variance, respectively. Factor loadings from this 4-factor solution are shown in the Supplementary Materials.

In the rotated solution, Factor 1 comprises most of the items defined as "Positive self-concept" in our original framework, while Factor 3 comprises most of the items from the original "Higher order thinking skills (HOTS)" scale and "Social and Communication skills" scales. Interestingly, Factor 4 "Negative self-concept," comprises all the negatively worded items. Factor 2 comprises the group of

⁶ Figure A. 2, Table A. 3, and Table A. 4 in the Supplementary Materials show results for the Parallel Analysis, Cattell scree test, and Bayesian Information Criterion used in the factor extraction decision. The eigenvalues, the two Parallel Analysis, and Cattell scree test suggested the extraction of 4 factors. The lowest BIC was found in the 4-factor solution. We decided to retain 4 factors.

items we refer to as “importance items,” intended to assess how youth valued each of the measured skills, with the hypothesis that youth who value a certain skill were more likely to work to improve that skill. We decided to drop the 16 “importance items”, as it appeared they loaded on a separate factor due to the grammatical structure, not adding relevant information to the model. We also dropped 14 items with low loading or cross loadings⁷ or that were poorly formulated and difficult to understand. For instance, two of the delayed gratification items presented low loadings and we decided to drop the 3rd item as well. Youth had already shown some difficulty with these items during the cognitive interviews.

Study 3

Study 3 had several purposes. First, we wanted to validate the structure of the revised instrument, after revisions were made following the findings in Uganda from Study 2. We did this with baseline data from a second country, Guatemala, using exploratory factor analysis. Second, we wanted to compare the means, factor structure, and reliability of the instrument for Guatemala and Uganda, and between baseline and a second, endline data collection in Guatemala. We did this using confirmatory factor analysis and internal consistency and test-retest reliability analysis. (Details on a measurement invariance analysis of these data are presented as a separate study in Cunha et al., 2021.) We also wanted to examine correlations between YAYSSA instrument scales and youth outcomes.

⁷ The 14 items dropped were: self_efficacy1 (How often do you believe that: it is hard for you to solve your problems?), self_awareness4 (How often do you know how you make other people feel?), gratification1 (How often do you save your money for something you want to buy later?), gratification2 (How often do you find it challenging to wait for something?), gratification3 (How often would you prefer to get one pen now rather than many pens later?), impulses2 (In the past month, how often have you interrupted your agemate when they were telling a story?), attention1 (In the past month, how often have you finished the work that you set out to do despite challenges?), attention3 (In the past month, how often have you kept doing something that you should do even if you didn't like it, such as homework?), regulate_behaviors3 (In the past month, how often have you got your work done immediately instead of waiting until the last minute?), thrill_seeking1 (How often do you do crazy things, such as drinking alcohol, even if they are a little dangerous?), critical_thinking1 (How often did you separate the true and false parts of the story?), social_skills3 (How often do you do your part when you work in a group?), social_skills7 (How often do you respect views that differ from your own?), communication2 (How often do you listen to your agemates' ideas?).

Finally, we wanted to evaluate whether anchoring vignettes (King & Wand, 2017) might help increase comparability between the two countries.

Method

Table 2 presents an overview of the data collections from all 3 studies.

Table 2.
Overview of Data Collections from All 3 Studies

	Study 1	Study 2	Study 3
Instrument	Initial	1st revision	2nd (final) revision
Data collection (N/k)	Uganda Cognitive Interviews (50/6 schools)	Uganda Baseline (1098/59 schools)	Uganda Endline (1010/59 schools) Guatemala Cognitive Interviews (55) Guatemala Baseline (794/59 sites) Guatemala Endline (784/67sites)
Testing dates	Oct-Nov 2017	Mar 2018	Aug 2018-Feb 2019
Number of items	80	79 + 10 Anchoring vignette items	50 + 10 Anchoring vignette items
Number of scales	16 (plus an "importance" scale)	16 (plus an "importance" scale)	4
Response categories	Frequency	Frequency	Agreement
Analyses	Qualitative	CFA, EFA, Reliability	CFA, EFA, Reliability, Longitudinal Change, Correlations with outcomes, Anchoring vignettes

Participants

Uganda Endline. Data collection took place in November 2018 with 1,010 youth, with a split sample of youth that had been tested at baseline (Study 2), and new participants, to test any prior participation effects. The same 59 schools participating in Study 1 also participated in Study 2, with an average of 18 youth sampled per school, and as with Study 1, included both Educate! participants and nonparticipants. Repeat administration (test-retest) was performed for a subsample of 57 randomly selected youth within two-weeks after endline administration. Participants were youth between 15 and 20 years old (average age 17.4), 54% female; 95% took the survey in English, 48% lived in rural areas,

23% reported being sexually active, 99% were in school full time, 30% were getting job training, 26% were working full or part time, and 27% were looking for jobs.

Guatemala Cognitive Interviewing. Fifty-five youth were randomly selected from sites participating in the USAID-funded Proyecto Puentes program in three communities in the Western Highlands in Guatemala during August of 2018⁸.

Guatemala Baseline. 794 youth participants were sampled from 59 program sites in the Western Highlands of the country at baseline (average of 15 youth per site, although large variation per site). The sites were selected randomly from a list of program sites provided by the Proyecto Puentes team. At baseline, youth between 15 and 19 years old were interviewed (average age 16.4), 53% of the respondents were female, 62% lived in rural areas, 69% spoke Spanish at home, 12% reported being sexually active, 80% were in school full time, 11% were getting job training, 43% were working full or part time, and 26% were looking for jobs.

Guatemala Endline. 784 youth were sampled across 67 sites at endline in January 2019 (average of 7 youth per site). Youth were selected randomly from a list of program sites provided by the Proyecto Puentes team. More sites were sampled at endline because the data collection firm had challenges tracking down youth participants because the program had ended. Therefore, several replacement sites were selected. A sub-sample of 126 participants was randomly selected for retest at endline within two weeks of full-scale administration.

Instruments

Revised YAYSSA. The initial round of data collection from the Uganda baseline (Study 2) pointed to issues with comprehension and interpretability of some of the items, as manifest in the low EFA loadings. Therefore, we revised and simplified items after this first round of EFA. First, we transformed

⁸ Proyecto Puentes is a youth development program implemented by World Vision in the Western Highlands that delivers life skills and technical and vocational training among 15-24 years-old, both in-school and out-of-school.

items from asking youth to evaluate how often they engaged in a certain behavior indicative of a skill to simple self-assessments about their perceptions of their skills. Additionally, we simplified the items by shortening them and removing extraneous language and unclear references. Some of the items are only slightly different (for example, “I can do most things if I make an effort” from the California Healthy Kids Survey was modified to “I can do most things if I try”). We also added one item to the self-esteem scale. Finally, we replaced the 5-point behavioral frequency response options⁹ with a simple 4-point agreement (endorsement) scale, with the categories “Strongly Disagree”/“Disagree”/“Agree”/“Strongly Agree.” For many items in the survey, we noted that agreement is a more natural, comprehensible, and interpretable response than is frequency. For example, to “agree” to the statement “I can do most things if I try” is understandable to a respondent and interpretable to an audience, whereas to respond “often” to the statement, “I can do most things if I try,” was harder to comprehend by the respondents and interpret by the youth.

Overall, the version of YAYSSA used in Study 3 is quite different from any one of the tools that it was adapted from, with a total of 50 items plus 10 anchoring vignettes items. Table A. 9 in the Supplementary Materials shows how the original version of the tool changed after the revisions¹⁰.

Additional Measures. In line with our conceptual framework, which posits that youth with stronger soft skills are more likely to have more positive outcomes, the YAYSSA assessment included questions about employment or job-seeking status of the youth participants, whether or not they had had sex (as a measure of sexual and reproductive health for this age group), as well as their recent experiences with conflict situations. In addition, the instrument included demographic and background

⁹ The response options for each item were initially designed as a 5-point behavioral frequency scale that asked youth to assess “how often” they act out a behavior that is associated with a particular skill (e.g., “How often do you think things through before you do them?”). Response options ranged from “Always or Almost Always” to “Never or Almost Never”.

¹⁰ The final version of the tool can be accessed online through the following link:
https://www.youthpower.org/sites/default/files/YouthPower/files/resources/YP%20Action%20Youth_Tool_English-v5.pdf

measures, such as disability status, to examine relationships between the measured soft skills and the youth characteristics. Using these measures, we construct four validation outcomes to correlate with the four soft skills scales: 1) “Employment Score”, a variable indicating success in the workforce; 2) “Ever had sex”, a variable indicating if the youth ever had sex; 3) “Any Disability”, a variable indicating if the youth reported having any disability; and 4) “Any Violence”, a variable indicating if the youth reported perpetrating any kind of violence (physical, verbal, or emotional). We also developed a module of socioeconomic status proxies, a module on disability status (following the Washington Group Short Set on Functioning developed (2017) by the Washington Group on Disability Statistics). All items, including the outcomes, were designed as self-reported scales, with an ordinal scale measuring the frequency/presence of the behavior, or the level to which the respondent agreed with a statement. Table A. 11 in the Supplementary Materials describes these variables in more details and the items can be found in the online version of the tool.¹¹

Anchoring Vignettes. The anchoring vignettes (AV) technique is used to improve comparability among assessments of attitudes and preferences in self-report questionnaires (King & Wand, 2017). The method aims to assess response styles using short descriptions of hypothetical persons (vignettes) that vary systematically in the latent traits represented in the inventory. Respondents are requested to rate the persons described in the vignettes on an item similar to those used for the respondents’ self-descriptions, adopting the same response format and rating scale (Primi et. al 2016b). A secondary use of anchoring vignettes is to assess participants’ comprehension of and engagement with the survey—if participants rate the vignettes correctly, this indicates good comprehension and engagement.

The application of anchoring vignettes usually involves attributing at least one vignette per item, if not more. Previous applications have used as many as 12 vignettes per self-assessment questions

¹¹ The final version of the tool can be accessed online through the following link:
https://www.youthpower.org/sites/default/files/YouthPower/files/resources/YP%20Action%20Youth_Tool_English-v5.pdf

(King & Wand, 2017). However, the addition of one or more questions for each self-assessment may increase the application costs considerably. Moreover, it can substantially increase the total time of the survey, adding to respondent burden. For these reasons, we opted to match two sets of AVs to many items, which may have impacted the validity of the AV-adjusted scores.

A set of anchoring vignettes (high and low on each scale) was included in the tool for each scale, for subsequent adjustment analysis (see Table A. 15). For AV adjustment, each item is rescored based on their position relative to the AV scoring: higher than High AV (score=5), same as High AV (score=4), between High and Low AV (score=3), same as Low AV (score=2), and lower than Low AV (score=1). The scale for the raw scores is 1 to 4; the scale for the AV adjusted scores is 1 to 5. This allows for a non-parametric adjustment that is not based on assumptions of an underlying distribution for each item. There are two independent AV adjustments for the Social and Communication Skills scale, due to having two sets of AVs for that scale. We explore the AV response patterns, the Cronbach's alpha statistics for the scales obtained from the AV-adjusted scores, and how the AV-adjusted scores correlate with the outcomes (employment, ever had sex, disability, and violence). See Table A. 15 in the Supplementary Materials for a description of the AV items.

Procedure

Cognitive Interview Studies in Guatemala. The process of qualitative cognitive interviewing in Guatemala was shorter than the process in Uganda, given the instrument went through detailed revision in Uganda based on Studies 1 and 2. Through a rapid analysis of the cognitive interview data, the team identified problematic words and phrasing and revised the tool accordingly. The instrument was administered in Spanish. A translation from English to Spanish, and backtranslation were administered prior to cognitive testing, to generate a working version prior to pilot administration.

Main Study Administrations. The revised YAYSSA tool was administered by tablet in a one-on-one interview format in both countries, and the duration of the interview was 50 minutes on average.

Uganda followed the same protocol as in Study 2. Youth were asked to fill out the behavioral items individually on a tablet, to aid in their comfort level in responding to potentially sensitive questions.

Uganda endline data collection was conducted by 30 enumerators between November 10 – 26, 2018. In Guatemala, at baseline, 20 enumerators conducted the assessment between September 4 – 22, 2018, and 17 enumerators administered the tool at endline between January 21 - February 7, 2019.

Enumerator Training. A refresher training was conducted in Uganda, and a full training was conducted in Guatemala prior to data collection. The training covered tool structure, study design, recruitment and sampling, as well as the use of the electronic platform for data collection, and research ethics.

Study Coordination. Similar to Study 2, study coordinators from each country contacted the program site ahead of time to arrange the interviews. A team lead met with the head teacher or program coordinator, who would gather all eligible youth for random selection. Enumerators explained the purpose of the study to all participants, including those re-testing, and obtained their informed consent and administered the assessment in an interview format

Human Subjects Reviews. The Guatemala study was approved by FHI 360's Protection of Human Subjects Committee and in Guatemala through a local context review. The approval procedure for Uganda is described in Study 2's Method section.

Results

We conducted exploratory factor analysis using data from Guatemala baseline to identify an appropriate scale structure for the revised instrument, followed by confirmatory factor analysis determining the final structure of the instrument, using endline data from Uganda and Guatemala. We then computed scale averages for participants used in concurrent and predictive validity analyses, including correlations with relevant socioeconomic and demographic characteristics, as well as proxies of youth outcomes of interest, such as self-reported employment status, self-reported ever having sex,

and self-reported perpetration of violent or abusive behaviors. Confirmatory factor analysis and exploratory factor analysis were conducted using Mplus 8 software; internal consistency, test-retest reliability, change over time, and predictive validity were performed using Stata/SE 15 software, while anchoring vignettes adjustments were conducted using the R package (R Core Team, 2021) anchors (Wand et al., 2007).¹²

Exploratory Factor Analysis

The revised 50 items of the Guatemala Baseline data collection (N= 794) were subjected to a second exploratory factor analysis. The EFA returned a 4-factor solution as the best fit for the baseline data, accounting for 86% of the variance.¹³ The eigenvalue of factor 1 was 7.3 and accounted for 55% of the variance; the second factor presented an eigenvalue of 1.7 and accounted for 13% of the variance; and factor 3 and 4 presented eigenvalues of 1.4 and 0.9 and accounted for 10.9% and 6.7% of the variance. Factor loadings are shown in Table A. 8 of the Supplementary Materials. Overall, items presented high loadings and only two items with low loadings were dropped.¹⁴

After rotation, the 4-factor solution is similar to the solution obtained from the Study 2 EFA and similar to the original theoretical framework, with the exception that now items classified as “HOTS” and Social and Communication Skills no longer load on the same factor. The items that loaded onto Factor 1 fall under our originally hypothesized construct “positive self-concept.” The items that loaded onto Factor 2 include items from our original conceptualizations of social skills and communication and items

¹² Item descriptive information and percent of response by response category can be found in the Supplementary Materials for Uganda baseline data (Table A. 18) and Guatemala baseline data (Table A. 19); and in Cunha et al. (2021) for Uganda and Guatemala endline data.

¹³ Figure A. 3, Table A. 6, and Table A. 7 in the Supplementary Materials shows results for the two Parallel Analysis, Cattell scree test, and Bayesian Information Criterion used in the factor extraction decision. The eigenvalues and Cattell scree test suggested the extraction of 3, while the two Parallel Analysis suggested the extraction of 3 and 4 factors respectively. The lowest BIC was found in the 3-factor solution. Given the 4-factor solution is very similar to the original theoretical proposed structure, we decided to retain 4 factors.

¹⁴ The item emotion3 (If a friend tells me I did something wrong, I can stay calm) and social_skills2 (I find a way of working things out if two of my friends quarrel) were dropped due to low loadings.

that measure how well youth ask for help when they try to solve problems. Examples include: “I get along well with people from different backgrounds” and “I write well.” Our initial distinction of communication skills from social skills came from the way communication skills was defined in the workforce literature as a separate skill (Lippman et al., 2015). However, it is not surprising that these two factors converged, given the overlap in the behaviors associated with these skills, such as participating in a team, asserting oneself appropriately to resolve a conflict, and complimenting others.¹⁵ Items loading onto Factor 3, “Higher order thinking skills” scale, includes decision making, impulse control, and problem-solving.

The only notable departure from the originally hypothesized structure is the emergence of Factor 4, “Negative self-concept”, which reflects the negative self-concept as distinct from the positive self-concept continuum measured by Factor 1. This factor aggregates all the negatively framed items. This distinction between negative self-concept and positive self-concept is supported by evidence from the psychology literature, which consistently finds two separate personality dimensions in negative and positive affect (distress and general negativity, including anger, disgust, fear, and nervousness versus enthusiasm, activeness, and alertness) (Watson et al., 1988; Watson & Tellegen, 1985). Similarly, Tellegen and Waller’s (1987) analysis finds seven major personality dimensions, which include two dimensions that reflect how one feels about oneself: positive valence (for example, “excellent” vs. “ordinary”) and negative valence (for example, “evil” vs “decent.”). We’ve summarized how the items from the 5-factor structure map onto each of the 4 factors in Table A. 9 in the Supplementary Materials.

Confirmatory Factor Analysis

Following the scale structure that emerged through the exploratory factor analysis, we examined the fit of the new scales through a confirmatory factor analysis (CFA) using baseline data from Uganda (N=1010) and Guatemala (N=784) where the 48-item revised version of the instrument was

¹⁵ See Harvard’s [Explore SEL Taxonomy](#) for more examples.

implemented. The results of the CFA confirmed strong model fit in both countries. The comparative fit index (0.947 for Uganda and 0.916 for Guatemala); the Tucker-Lewis fit index (0.944 for Uganda and 0.912 for Guatemala); the standardized root mean square residual (0.045 for Uganda and 0.057 for Guatemala); and the root mean square residual fit index (0.03 for Uganda and 0.041 for Guatemala) indicate a good fit between the model and the observed data, as shown in Table 3.¹⁶

Table 3*CFA Goodness-of-fit indicators Study 2*

	Uganda	Guatemala
Chi Sq.	2082.661	2480.613
df	1074	1074
p > Chi Sq.	<0.001	<0.001
CFI	0.947	0.916
TLI	0.944	0.912
SRMR	0.045	0.057
RMSEA	0.03	0.041
RMSEA 90% CI, lower bound	0.029	0.039
RMSEA 90% CI, upper bound	0.032	0.043
Observations	1010	784

Note: CFA ran using WLSMV estimator in Mplus 8 with factor variances set to 1. Following Hu & Bentler (1999), we adopt the following cut-off values to determine if the model fits the data well: SRMR ideally below 0.08 and at most 0.10; RMSEA ideally below 0.06 and at most 0.10; and CFI and TLI ideally above 0.95, with a minimum of 0.90.

Internal consistency

We use the Cronbach's alpha and Omega as an initial measure of scale internal consistency, or the strength of the relationships between items within a scale. Table 4 shows Cronbach's alpha and Omega statistics for the 4-scales at Baseline Guatemala and Endline Guatemala and Uganda. For Uganda, Cronbach's alpha values ranged from 0.66 to 0.81 at endline; and for Guatemala, Cronbach's alpha values were from 0.70 to 0.83 at baseline and from 0.70 to 0.86 at endline. Omega reliability

¹⁶ CFA loadings are shown in Table A. 10 in the Supplementary Materials.

results are very similar to Cronbach's alpha statistics. Cronbach's alpha statistics and Omega indicate a substantial level of internal consistency.¹⁷

Table 4*Cronbach's alpha and Omega reliability for Study 3*

Dimension	Uganda Endline		Guatemala Baseline		Guatemala Endline	
	Alpha	Omega	Alpha	Omega	Alpha	Omega
Positive self-concept	0.81	0.82	0.83	0.83	0.86	0.86
Negative self-concept	0.66	0.65	0.71	0.70	0.76	0.76
HOTS	0.69	0.69	0.70	0.70	0.70	0.70
Social & communication skills	0.66	0.66	0.71	0.72	0.72	0.71
Sample Size	1010		794		784	

Test-Retest Reliability

Test-retest reliability analysis was performed as gauge of measurement stability over a short period of time. In this study, repeat administration was performed for a subsample of 57 youth at Uganda endline and 126 youth at Guatemala endline over an interval of two weeks, with the same enumerators. Table 5 shows that the test-retest correlation coefficients ranged between .5 and .7 between Uganda and Guatemala for the four scales.

Table 5*Test-retest reliability coefficients – Study 3*

Dimension	Uganda	Guatemala
Positive self-concept	0.70	0.70
Negative self-concept	0.58	0.65
HOTS	0.65	0.64
Social & communication skills	0.56	0.70
Sample Size	57	126

The test-retest statistics displayed at the scale level are comparable to those reported on other soft skills measurement tools, such as the General Self-Efficacy Scale (.45 - .75 in Schwarzer & Jerusalem

¹⁷ There are no universal standards for internal consistency reliability, and test use tends to be a moderating consideration—using tests for high stakes decisions requires higher reliability (Kane, 2011). However, values of .70 and above are generally considered acceptable (Cortina, 1993; Taber, 2018).

1995); the Responses to Stress Questionnaire (.49 to .76 for the 19 parcels and .69 to .81 for the five factors (Connor-Smith et al. 2000); and the Strengths and Difficulties Questionnaire: Parent or Teacher version (for the parent version: .54 - .81; for the teacher version .58 - .8 per Stone et al. 2015). The knowledge base on reasonable test-retest reliability statistics to expect on soft skills assessment is still limited; however, it appears that the levels reported on the YAYSSA tool follow the trend of these prior assessments.

Change Over Time

A crucial element of the tool validation is the assessment of its ability to capture change in soft skills over time, plausibly affected by an intervention. While neither the Uganda nor Guatemala administrations were designed for a program evaluation (although program activities took place in the interim between initial and later testing), a simple before-and-after measurement between points in time is the first step towards understanding whether the tool has the potential of capturing a program or time effect. Because the tool changed from baseline to endline in Uganda, it is only possible to look at change over time using Guatemala data.

Table 6 shows how the scales changed over time in Guatemala for the sample of 450 repeating youth who were surveyed both at baseline and endline (panel sample). The skills levels at endline are lower when compared to the baseline levels, except for negative self-concept, which are higher. The differences range from 0.03 to 0.11, although differences are not significant for HOTS and social and communication skills.¹⁸ **Pre-post changes in the opposite direction from what might be expected are not uncommon in self-report studies.** As Bhanji et al (2012) point out, interpretation of changes on a self-assessment depend on a common metric, which can be violated when participants' standards or understanding of the dimension changes during the treatment period, which is sometimes referred to as

¹⁸ Table A. 11 in the Supplementary Materials shows the means at baseline and endline for the non-repeating sample of youth who were interview only once, either at baseline or endline. Results are similar to the panel sample.

response shift bias (Drennan & Hyde, 2008). This may have occurred here. One approach to address this is through the use of retrospective pretest designs (Bhanji et al, 2012; Drennan & Hyde, 2008), but that falls outside the scope of the present study.

The correlations between the baseline and endline scores varies from 0.36 to 0.45 (Table A. 12 in the Supplementary Materials). Because there is no comparison group, it is not possible to attribute this difference to the program. However, the conclusion remains that the scales on positive and negative self-concept have registered change between the baseline and endline administrations.

Table 6

Mean differences and change over time for Guatemala baseline and endline – Panel Sample

Factor	Base. Mean	Base. SD	End Mean	End SD	Pooled SD	SE	t- score	Diff (E- B)
Positive Self Concept	3.38	0.31	3.28	0.34	0.33	0.022	-4.47	-0.10*
Negative Self Concept	2.04	0.30	2.15	0.34	0.32	0.021	5.00	0.11*
HOTS	2.99	0.32	2.96	0.31	0.31	0.021	-1.73	-0.04
Communication & Social Skills	2.98	0.37	2.95	0.35	0.36	0.024	-1.19	-0.03
Observations	450		450					

Note. Panel Sample contains the youth in Guatemala who were interviewed both at baseline and endline.

* $p < .05$.

Correlations of Scale Scores with Outcomes

We performed Pearson Correlations in Stata/SE 15 to examine the relation between the YAYSSA tool and the outcomes. These are shown in Table 7. Overall, although the magnitude of the correlation is modest for some of the outcomes, they are all in the expected direction. Youth with higher positive self-concept, HOTS, and social and communication skills were more likely to have a higher employment score, while youth with a higher negative self-concept were less likely to have a higher employment score (although the latter relationship is not significant). In Uganda, youth with higher negative self-concept were more likely to have ever had sex, while in Guatemala youth with higher social and communication skills were more likely to have ever had sex. Looking at disability status, youth who reported having “any disability” also reported lower positive skills (positive self-concept, HOTS, and communication and social skills) and higher negative self-concept. Finally, youth who reported having

perpetrated “any violence” also reported higher negative self-concept and lower positive skills (although the relationship was not significant for positive self-concept in Guatemala). In sum, the scales appear to have been predictive of the youth behaviors important for the validation, and these relationships were similar in magnitude across the two countries.

Table 7

Correlations between YAYSSA scales and selected outcome variables

	Uganda (endline)				Guatemala (baseline)			
	Pos SC	Neg SC	HOTS	SC Skills	Pos SC	Neg SC	HOTS	SC Skills
Employ	0.07*	-0.05	0.09*	0.07*	0.09*	-0.07	0.07*	0.08*
Had Sex	-0.05	0.08*	-0.04	-0.01	0.04	-0.06	0.01	0.10*
Disability	-0.07*	0.12*	-0.08*	-0.09*	-0.11*	0.10*	-0.16*	-0.16*
Violence	-0.12*	0.20*	-0.12*	-0.20*	-0.06	0.10*	-0.12*	-0.13*

Note. Correlations at baseline showed for Guatemala and at endline for Uganda. Table A. 13 in the Supplementary Materials contains description and definitions of the selected outcome variables. Pos SC = positive self-concept; Neg SC = negative self-concept; HOTS = higher order thinking skills; SC Skills = social and communication skills. Table A. 14 in the Supplementary Materials shows descriptive statistics for the selected outcome variables.

* $p < 0.05$

Anchoring Vignettes Adjustment

Table A. 16 in the Supplementary Materials includes information on the scale means for the raw and anchoring vignette adjusted scores and Figure A. 4 in the Supplementary Materials shows the response patterns for each set of AVs. The first row shows the frequencies for Uganda endline, and the second and third rows show frequencies for Guatemala baseline and endline, respectively. The first bar, labeled “1,2” shows the frequency of responses when the respondents ordered the High AV higher than the Low AV; the second bar, labeled “{1,2}”, shows when responses were tied, which means respondents gave the same rating for the Low and the High AV; and the third bar shows the cases where the Low AV was rated higher than the High AV. Overall, responses were in the direction we would expect, where respondents rated the High AV higher than the Low AV for more than 80% of the cases, with the exception of Communication for Uganda endline. We can take this as an indication that the

level of engagement and comprehension was very high among respondents, and the AVs were working as a good diagnostic that respondents were taking the survey seriously.

Table A. 17 in the Supplementary Materials shows Cronbach's alpha statistics for the 4-scales obtained from the AV-adjusted scores. Internal consistency values were stronger for AV-adjusted scores. However, Cronbach's alpha is not an adequate measure to evaluate AVs reliability, as pointed out by von Davier et al. (2017). According to the authors, the AV approach relies on the assumptions that the vignettes are supposed to be invariant across respondents and the response to vignette prompts are supposed to be without error and strictly ordered. They show these assumptions are not always met and that higher Cronbach's alpha are obtained regardless of whether the assumptions are met or not.

Correlating the AV adjusted score with outcome variables is a more appropriate method of exploring the validity of the vignettes (He et al. 2017; Kyllonen & Bertling 2014; Primi et al. 2016b). Table 8 shows correlations of the AV-adjusted scores with external variables. Although the direction of the correlations was similar to the ones obtained from the non-adjusted scores, the magnitude of the correlation was considerably lower, indicating the AV-adjusted scores were not out-performing the non-adjusted scores.

Because the AV-adjusted scores did not improve the scales, we recommend the use of the non-adjusted scores for the main analysis. We did not find support for the need to use AVs for cross-country comparisons between the two sites where the pilot was implemented. However, users may want to

perform an AV analysis if they are conducting cross-country comparison between countries with contexts dissimilar to our two sites. We also recommend using the AVs to investigate engagement and comprehension of the survey as shown by Figure A. 4 in the Supplementary Materials, as we mentioned above.

Table 8

Correlations between YAYSSA scales and selected outcome variables - AV-adjusted scores

	Uganda (endline)				Guatemala (baseline)			
	Pos SC	Neg SC	HOTS	SC Skills	Pos SC	Neg SC	HOTS	SC Skills
Employ	0	-0.02	0.03	0.01	0.03	0	-0.01	0
Had Sex	-0.04	0.03	-0.01	-0.02	0.05	0.06	-0.08*	-0.01
Disability	-0.03	0.04	-0.06	-0.13*	-0.06	0.04	-0.04	-0.11*
Violence	-0.03	0.08*	-0.05	-0.14*	-0.07	0.05	-0.04	-0.06

Note: Correlations at baseline showed for Guatemala and at endline for Uganda. Table A. 13 in the Supplementary Materials contains description and definitions of the selected outcome variables. Pos SC stands for positive self-concept; Neg SC stands for negative self-concept; HOTS stands for higher order thinking skills; and SC Skills stands for social and communication skills. Table A. 14 shows descriptive statistics for the selected outcome variables.

* $p < .05$.

Discussion

Results from two validation sites indicate that the YAYSSA tool was internally consistent and valid, as demonstrated through EFA, CFA, Cronbach's alpha statistics, test-retest statistics, and correlation coefficients between soft skills and outcomes. We also found that the tool was invariant by country, gender, and socio-economic status (see Cunha et al., 2021). Finally, we found that the tool measures change over time; however, there is no comparison group, and we cannot attribute this change to the partner programs.

The originally hypothesized structure changed somewhat during the process of the tool development—negative self-concept emerged as a construct distinct from positive self-concept, and social skills and communication skills clustered together. Our empirically determined scale structure demonstrated a good fit with the data.

Validation of the tool revealed several interesting findings. First, violence and disability status were both predictive of soft skills. Youth who reported having engaged in any kind of violence reported lower levels of skills across the board (and higher levels of negative self-concept). Similarly, youth who reported having any disability reported lower levels of all skills (and higher negative self-concept). Employment status, and to some extent, ever having sex, were also predictive of soft skills. Youth's employment outcomes were associated with small but significant changes in the expected direction for all skills, except for negative self-concept. Ever having had sex was significantly associated with higher negative self-concept in Uganda and higher social and communication skills in Guatemala.

In addition to these psychometric tests, we performed several other analyses that revealed useful information for future development of soft skills measures. Our integration of anchoring vignettes into the tool shows the utility of anchoring vignettes as a technique for assessing respondents' engagement with and comprehension of an instrument. However, because the AV-adjusted scores did not improve the scales, we recommend the use of the non-adjusted scores for the main analysis.

In sum, the YAYSSA tool presents a validated assessment of youth soft skills such as positive self-concept, negative self-concept, higher order thinking skills, and social and communication skills. The assessment can be used to predict youth outcomes in key areas and measure change in the level of soft skills over time, in as short as a few months. Further testing is necessary to determine validity in contexts substantially different than the youth programs in which the assessment was validated, with other youth outcomes, and in a causal inference framework.

Our experience in developing this instrument reveals several key lessons for tool development, implementation, and future research. First, contextualization is a critical first step. The cognitive testing that we conducted in Uganda and Guatemala revealed critical information that informed our initial revisions and helped us to interpret several confounding findings. For example, data from the cognitive interviews in Uganda showed that items measuring delayed gratification were not necessarily

contextually appropriate. This informed our decision to ultimately remove these items from the instrument. Second, our analysis of the Uganda baseline data in Study 2 showed that the most low-performing items were often also the wordiest. Overall, the tool performed better once we simplified the item wording and structure. Because youth were not using all five response options provided to them in the first version of the tool, we reduced the number of response options to four.

In addition, implementers seeking to measure program participants' soft skills levels and progress should build in sufficient time and resources in the program cycle to conduct baseline (and endline, where relevant) assessments—and to contextualize the assessment for the program and country context. The time required to administer a survey will depend primarily on the readiness of the tool for administration, the number and geographic spread of youth to be surveyed; and the number of staff available to administer the survey.

We recommend that future research studies focus on the following gaps we identified during the process of developing this tool. First, testing this tool in the context of an aligned soft skills intervention, with a control group to more clearly isolate where change in skill levels may be coming from would be a valuable addition to the literature. Future studies could also focus on differences in skill levels by intervention dosage or modality to discern how much of an intervention is required to “move the needle” on a certain skill, and whether certain intervention modalities work better than others, especially by cultural context. Moreover, further research on the pathways linking skills and outcomes is needed. Our analysis found that higher communication and social skills among youth were linked to a higher likelihood that youth had ever had sex in Guatemala. Our earlier literature review (Gates et al., 2016) found several studies linking higher skill levels to violent behavior and risky sexual behaviors. Research exploring the relationship between youth soft skills and specific employment outcomes, such as on-the-job performance, retention, and wages, would also be useful for understanding how youth's skills support them in the workforce. Finally, additional testing and validation of this tool in other

contexts would help us to better understand cross-cultural differences in how skills change over time, how skills relate to key outcomes, and how skills relate to other key demographics, such as socioeconomic status (SES), gender, and disability—and whether other environmental factors need to be taken into consideration when planning soft skills interventions would also be a valuable addition to the literature. Testing of the tool in other countries may also reveal different response styles among youth, in which case the AVs might be applied to compare data across contexts.

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.
- Alaoui, L. & Fons-Rosen, C. (2021). Know when to fold 'em: The flip side of grit. *European Economic Review*, 136. DOI: 10.1016/j.eurocorev.2021.103736.
- Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality psychology and economics. In *Handbook of the Economics of Education* (Vol. 4, pp. 1-181). Elsevier.
- Arthur, M. W., Hawkins, J. D., Pollard, J. A., Catalan, R. F., & Baglioni, A. J. (2002). Measuring risk and protective factors for substance use, delinquency, and other adolescent problem behaviors: The Communities That Care Youth Survey. *Evaluation Review*, 26, 2002, pp. 575–601.
- Barry, C. L., & Finney, S. J. (2016). Modeling change in effort across a low-stakes testing session: A latent growth curve modeling approach. *Applied Measurement in Education*, 29(1), 46-64.
- Bhanji, F., Gottesman, R., de Grave, W., Steinert, Y., & Winer, L. R. (2012). The retrospective pre–post: a practical method to evaluate learning from an educational program. *Academic emergency medicine*, 19(2), 189-194. <https://doi.org/10.1111/j.1553-2712.2011.01270.x>
- Bollen, K. A. (1989). Structural equations with latent variables. New York, NY: Wiley.
- Carneiro, P., Crawford, C., & Goodman, A. (2007). The impact of early cognitive and non-cognitive skills on later outcomes, Center for the Economics of Education Discussion Paper 92.
- Charness, G., Gneezy, U., & Henderson, A. (2018). Experimental methods: Measuring effort in economics experiments. *Journal of Economic Behavior and Organization*, 149, 74-87.
- Connelly, B. S., & Ones D. S. (2010). Another perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136(6), 1092–1122. doi:10.1037/a0021212
- Connor-Smith, J. K., Compas, B. E., Wadsworth, M. E., Thomsen, A. H., & Saltzman, H. (2000). Responses to stress in adolescence: measurement of coping and involuntary stress responses. *Journal of consulting and clinical psychology*, 68(6), 976-992.
- Copeland, E. P., Nelson, R., & Traughber, M. C. (2010). Wellness dimensions relate to happiness in children and adolescents. *Advances in School Mental Health Promotion*, 3(4), 25–37.
- CORE Districts. (n.d.). School Quality Improvement Index, Student Social-Emotional Learning. Retrieved from <https://coredistricts.org/our-improvement-data/improvement-measures/>.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, 78(1), 98. <https://doi.org/10.1037/0021-9010.78.1.98>

- Cunha, N., Martinez, A., Kyllonen, P., Gates, S. (2021). Cross-Country Comparability of a Social-Emotional Skills Assessment Designed for Youth in Low-Resource Environments. *International Journal of Testing*, 21:3-4, 182-219.
- Davies, M., Shin, H., Khorramdel, L., & Stankov, L. (2017). The effects of vignette scoring on reliability and validity of self-reports. *Applied Psychological Measurement* 42(4):291–306.
- Deming, D. J. (2017). The value of soft skills in the labor market. *NBER Reporter*, 4, 7-11.
- Drennan, J., & Hyde, A. (2008). Controlling response shift bias: the use of the retrospective pre-test design in the evaluation of a master's programme. *Assessment & Evaluation in Higher Education*, 33(6), 699-709.
- Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44(4), 237-251.
- Educational Testing Service (ETS). (2014). ETS Personal Potential Index: Evaluator User's Guide. www.ets.org/ppi.
- Furlong, M. J., You, S., Renshaw, T. L., Smith, D. C., & O'Malley, M. D. (2014). Preliminary development and validation of the Social and Emotional Health Survey for secondary school students. *Social Indicators Research*, 117(3): 1011–1032.
- Galloway, T., Lippman, L., Burke, H., Diener, O., & Gates, S. (2017). Measuring Soft Skills in International Youth Development Programs: A Review and Inventory of Tools, Washington, DC: USAID's YouthPower: Implementation, YouthPower Action.
- Gardner, J. M., Williams, K. R., Guerra, N. G., & Walker, I. (2011). The Jamaica youth survey: assessing core competencies and risk for aggression among Jamaican youth. *Caribbean Quarterly*, 57(1), 35-53.
- Gates, S., Lippman, L., Shadowen, N., Burke, H., Diener, O., & Malkin, M. (2016). Key Soft Skills for Cross-Sectoral Youth Outcomes. Washington, DC: USAID's YouthPower: Implementation, YouthPower Action.
- Goodman R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.
- Goodman R., Meltzer H., Bailey V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child and Adolescent Psychiatry*, 7, 125-130.
- Hao, J., Liu, L., Kyllonen, P., Flor, M., & von Davier, A. A. (2019). *Psychometric considerations and a general scoring strategy for assessments of collaborative problem solving* (Research Report No. RR-19-41). Princeton, NJ: Educational Testing Service. <https://doi.org/10.1002/ets2.12276>
- He, J., Buchholz, J., & Klieme, E. (2017). Effects of anchoring vignettes on comparability and predictive validity of student self-reports in 64 cultures. *Journal of Cross-Cultural Psychology*, 48, 319-334.
- Heckman, J. J., & Kautz, T. (2013). *Fostering and measuring skills: Interventions that improve character and cognition* (No. w19656). National Bureau of Economic Research.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, 24(3), 411-482.
- Hitt, C., Trivitt, J., & Cheng, A. (2016). When you say nothing at all: The predictive power of student effort on surveys. *Economics of Education Review*, 52, 105-119.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185
- Hoyle, R. H. (1995). Structural equation modeling: Concepts, issues, and applications. Thousand Oaks, CA: Sage Publications.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1e55.
- International Youth Foundation. (n.d.) Passport to Success: Trainer Observation Tool.

- Kane, M. (2011), The Errors of Our Ways. *Journal of Educational Measurement*, 48(1) 12-30.
<https://doi.org/10.1111/j.1745-3984.2010.00128.x>
- Kautz, T., & Moore, Q. (2018). "Measuring Self-Regulation Skills in Evaluations of Employment Programs for Low-Income Populations: Challenges and Recommendations." OPRE Report 2018-83. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- King, G., & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis*, 15, 46–66.
- Kyllonen, P. C., & Bertling, J. J. (2014). Innovative questionnaire assessment methods to increase cross-country comparability. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 277-286). Boca Raton, FL: CRC Press.
- Kyllonen, P. C. (2015). Designing Tests to Measure Personal Attributes and Noncognitive Skills. In Suzanne Lane, Mark R. Raymond, Thomas M. Haladyna (Eds.), *Handbook of Test Development*. Abingdon: Routledge.
- Lippman, L., Moore, K.A., Guzman, L., Ryberg, R., McIntosh, H., Ramos, M., Caal, S., Carle, A., and Kuhfeld, M. (2014). *Flourishing children: Defining and testing indicators of positive development*. Springer.
- Lippman, L., Ryberg, R., Carney, R. & Moore, K.A. (2015). Key “Soft Skills” that Foster Youth Workforce Success: Toward a Consensus Across Fields. Washington, DC: USAID, FHI 360, Child Trends. Published through the Workforce Connections project managed by FHI 360 and funded by USAID.
- Lorenzo-Seva, U., & Ferrando, P. J. 2006. FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods* 38(1): 88–91.
- MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences*, 19(4), 451-458.
- Martin-Raugh, M. P., Kyllonen, P. C., Hao, J., Bacall, A., Becker, D., Kurzum, C., Yang, Z., Yan, F., & Barnwell, P. (2020). Negotiation as an interpersonal skill: Generalizability of negotiation outcomes and tactics across contexts at the individual and collective levels. *Computers in Human Behavior*, 104, 105966. <https://doi.org/10.1016/j.chb.2019.03.030>
- Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*. New York, NY: Times Books.
- Muthén, B., Muthén, L., & Asparouhov, T. (2015). Estimator choices with categorical outcomes. Mplus paper. Downloaded from <https://www.statmodel.com/download/EstimatorChoices.pdf>.
- OECD (2014). PISA 2012 Results: Creative Problem Solving: Students’ Skills in Tackling Real-Life Problems (Volume V), PISA, OECD Publishing.
- Oh, I. S., Wang, G., and Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*, 96(4), 762–773.
- Primi, R., Santos, D., John, O. P., & De Fruyt, F. (2016a). Development of an inventory assessing social and emotional skills in Brazilian youth. *European Journal of Psychological Assessment*.
- Primi, R., Zanon, C., Santos, D., De Fruit, F. & John, O. P. (2016b). Can they make adolescent self-reports of social-emotional skills more reliable, discriminant, and criterion-valid? *European Journal of Psychological Assessment*, 32, 39–51.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
<https://pdfs.semanticscholar.org/19d5/ba23b50c86d0c46895664bb2cd49b09c993e.pdf>

- Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston, Measures in health psychology: A user's portfolio. Causal and control beliefs (pp. 35-37). Windsor, England: NFER-NELSON.
- Shek, D. T., Siu, A. M., & Lee, T. Y. (2007). The Chinese positive youth development scale: A validation study. *Research on Social Work Practice*, 17(3), 380-391.
- Soland, J., Hamilton, L. S., & Stecher, B. M. (2013). Measuring 21st Century Competencies. Global Cities Education Network.
- Soland, J., Zamarro, G., Cheng, A., & Hitt, C. (2019). Identifying Naturally Occurring Direct Assessments of Social-Emotional Competencies: The Promise and Limitations of Survey and Assessment Disengagement Metadata. *Educational Researcher*, 48(7), 466-478.
- Stecher, B. M., & Hamilton, L. S. (2014). Measuring Hard-to-Measure Student Competencies: A Research and Development Plan. RAND Corporation. Santa Monica, CA.
- Stein, S. G. (2000). Equipped for the future content standards: *What adults need to know and be able to do in the 21st century*. National Institute for Literacy.
- Stone, L. L., Janssens, J. M., Vermulst, A. A., Van Der Maten, M., Engels, R. C., & Otten, R. (2015). The Strengths and Difficulties Questionnaire: psychometric properties of the parent and teacher version in children aged 4–7. *BMC psychology*, 3(1), 4.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in science education*, 48(6), 1273-1296.
<https://doi.org/10.1007/s11165-016-9602-2>
- Tellegen, A., & Waller, N. G. (1987). Re-examining basic dimensions of natural language trait descriptors. In *95th Annual Convention of the American Psychological Association, New York*.
- U. S. Department of Labor (n.d.). Attachment D: Optional Work Readiness Tool for the WIA Youth Program. Author. <https://wdr.doleta.gov/directives/attach/TEGL/TEGL07-10a4.pdf>.
- Wand, J., King, G., & Lau, O. (2011). Anchors: Software for anchoring vignette data. *Journal of Statistical Software*, 42(1), 1-25.
- Washington Group on Disability Statistics. (2017). The Washington Group Short Set on Functioning: Question Specifications. <http://www.washingtongroup-disability.com/wp-content/uploads/2016/12/WG-Document-4-The-Washington-Group-Short-Set-on-Functioning-Question-Specifications.pdf>.
- Watson, D., Clark, L. A., Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*. 54 (6), 1063–1070. doi:10.1037/0022-3514.54.6.1063
- Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological bulletin*, 98(2), 219.
- Zamarro, G., Cheng, A., Shakeel, M. D., & Hitt, C. (2018). Comparing and validating measures of non-cognitive traits: Performance task measures and self-reports from a nationally representative internet panel. *Journal of Behavioral and Experimental Economics*, 72, 51-60.

Supplementary Materials

Figure A. 1

Response options tested during cognitive interviews in Study 1

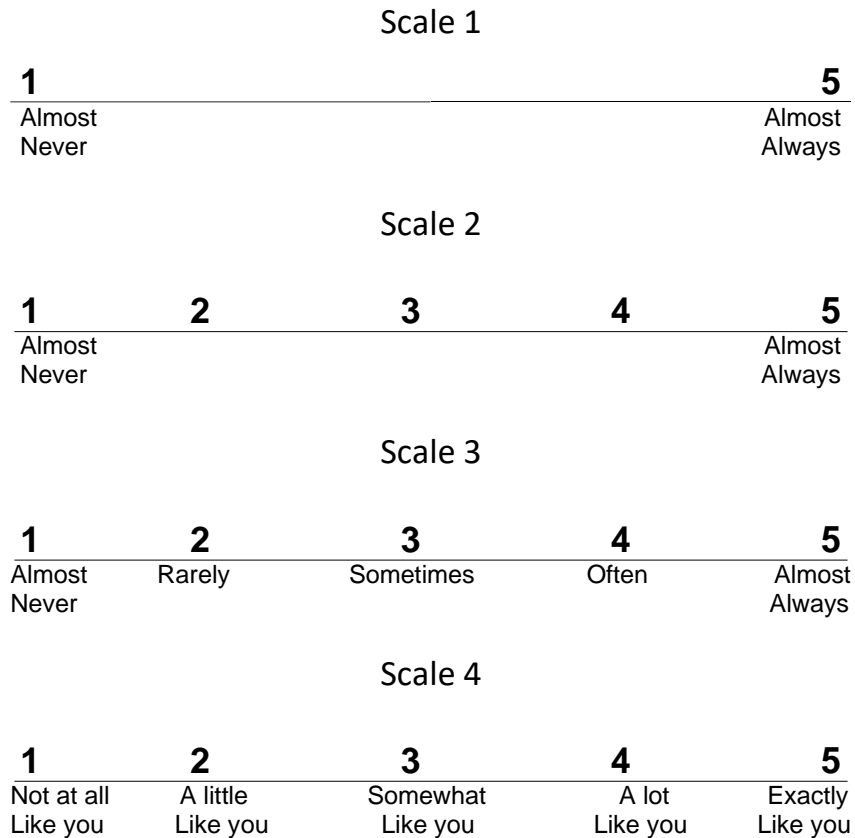


Table A. 1

Item revision after Study 2 EFA (5 versus 4 factor structure)

Original Item	Scale	Subscale	Decision	Revised Item
When answering these questions, please think about times when you were healthy and had your basic needs, including school fees, met.				

			Added	<i>I'm going to ask you some questions about how you solve problems and do tasks, and some questions about how you feel about yourself. Remember there are no right or wrong answers. I just want to learn about you and what you think.</i>
How often do you believe that: it is hard for you to work your problems?	Positive Self-Concept	Self-Efficacy	Revised	How often do you believe that: it is hard for you to solve your problems?
How often do you believe that: there are many things that you do not do well?	Positive Self-Concept	Self-Efficacy	Revised	How often do you believe that: there are many things that you do poorly?
How often do you believe that: you are good at learning?	Positive Self-Concept	Self-Efficacy	Revised	How often do you believe that: you are good at learning something new?
How often do you believe that: you can do most things if you try?	Positive Self-Concept	Self-Efficacy	Revised	How often do you believe that: you can do most things if you make the effort?
How often do you believe that: you can do what you need to do to succeed in life?	Positive Self-Concept	Self-Efficacy	Revised	How often do you believe that: you can do something that will help you succeed in life?
How important is this to you: Being capable of doing most things?	Positive Self-Concept	Self-Efficacy	Unchanged	
How often have you felt that you are not good at all?	Positive Self-Concept	Self-Esteem	Revised	How often do you feel that you are not good at all?
How often have you felt that: members of your household appreciate you?	Positive Self-Concept	Self-Esteem	Revised	How often have you felt that: the people you live with at home value you?
How often have you felt that: you are just as important as anybody else?	Positive Self-Concept	Self-Esteem	Excluded	
	Positive Self-Concept	Self-Esteem	New Item Added	How often do you feel that you are a valued member of your community?

How often have you felt that: you have a number of good characteristics?	Positive Self-Concept	Self-Esteem	Revised	How often have you felt that: you have a number of good qualities?
How important is this to you: Liking yourself just the way you are?	Positive Self-Concept	Self-Esteem	Unchanged	
How often in the past 7 days have you felt confident in your abilities?	Positive Self-Concept	Self-Confidence	Revised	How often do you feel good about your skills?
How often in the past 7 days have you felt unsure that you can be successful?	Positive Self-Concept	Self-Confidence	Revised	How often do you feel not sure that you can be successful?
How often in the past 7 days have you felt that you don't trust your abilities?	Positive Self-Concept	Self-Confidence	Revised	How often do you feel that you don't trust your skills?
How often in the past 7 days have you felt sure of yourself?	Positive Self-Concept	Self-Confidence	Revised	How often do you feel confident in yourself?
How important is this to you: Being confident in yourself?	Positive Self-Concept	Self-Confidence	Unchanged	
How often do you find it hard to know what you are feeling?	Positive Self-Concept	Self-Awareness	Revised	How often do you find it hard to know how you are feeling?
How often do you know what you are good at and not good at?	Positive Self-Concept	Self-Awareness	Revised	How often do you know what you are good at?
How often do you know what you are feeling in the moment?	Positive Self-Concept	Self-Awareness	Revised	How often do you know how you are feeling inside at any particular moment?
How often are you aware of how you make other people feel?	Positive Self-Concept	Self-Awareness	Revised	How often do you know how you make other people feel?
How important is this to you: Knowing how you feel about something?	Positive Self-Concept	Self-Awareness	Unchanged	
How often do you see that your future will be pleasant/happy?	Positive Self-Concept	Self-Belief	Revised	How often do you see that your future will be happy?

How often do you believe that you will achieve what you want?	Positive Self-Concept	Self-Belief	Revised	How often do you believe that you will reach your future goals?
How often do you know that you are going to be okay?	Positive Self-Concept	Self-Belief	Revised	How often do you know that you are going to be fine?
How often do you believe you can make things happen that will improve your life?	Positive Self-Concept	Self-Belief	Unchanged	
How important is this to you: Believing in yourself?	Positive Self-Concept	Self-Belief	Unchanged	
How often do you save your money for something you want to buy later?	Self-control	Gratification	Unchanged	
How often in the past 7 days have you been unable to wait for something you want?	Self-control	Gratification	Revised	How often do you find it challenging to wait for something?
How often would you prefer to get a small gift today than wait to get a big gift later?	Self-control	Gratification	Revised	How often would you prefer to get one pen now rather than many pens later?
How often would you choose to get paid less now if it meant you would get paid more later?	Self-control	Gratification	Excluded	
How important is this to you: Being able to wait for what you want?	Self-control	Gratification	Unchanged	
How often do you do things without thinking?	Self-control	Impulses	Revised	How often do you do things without thinking about what you're doing?
How often in the past 7 days have you allowed others to speak without interruption?	Self-control	Impulses	Revised	In the past month, how often have you interrupted your friend when they were telling a story?
How often do you think before you do things?	Self-control	Impulses	Revised	How often do you think through things before you do them?
How important is this to you: Controlling your behavior?	Self-control	Impulses	Unchanged	

How often in the past 7 days have you finished the work that you were doing?	Self-control	Attention	Revised	In the past month, how often have you finished the work that you set out to do despite challenges?
How often in the past 7 days were you easily distracted?	Self-control	Attention	Revised	In the past month, how often have you been unable to pay attention?
How often in the past 7 days have you kept doing something that you needed to do even if you didn't like it?	Self-control	Attention	Revised	In the past month, how often have you kept doing something that you should do even if you didn't like it, such as homework?
How often in the past 7 days have you had difficulty starting tasks?	Self-control	Attention	Revised	In the past month, how often have you found it difficult to start your work?
How important is this to you: Focusing your attention in order to complete something?	Self-control	Attention	Unchanged	
How often in the past 7 days have you done things to calm down when you were angry?	Self-control	Emotions	Revised	In the past month, how often have you done things to control your anger or temper, for example when you have quarreled with your friend?
How often in the past 7 days have you been upset by little things?	Self-control	Emotions	Revised	In the past month, how often have you been annoyed by little things, like if someone steps on your shoe?
How often in the past 7 days have you remained calm when criticized?	Self-control	Emotions	Revised	In the past month, how often have you remained calm when a friend tells you that you did something poorly?
How important is this to you: dealing with your feelings?	Self-control	Emotions	Unchanged	
How often were you able to stop yourself when you were going to do something you would regret?	Self-control	Regulate Behaviors	Revised	In the past month, how often were you able to stop yourself when you were going to do something you would regret?
How often in the past 7 days have you chosen not to follow directions?	Self-control	Regulate Behaviors	Revised	In the past month, how often have you refused to follow instructions?

How often in the past 7 days have you got your work done right away instead of waiting until the last minute?	Self-control	Regulate Behaviors	Revised	In the past month, how often have you got your work done immediately instead of waiting until the last minute?
How important is this to you: Doing your best, even when you don't want to do something?	Self-control	Regulate Behaviors	Unchanged	
How often have you done crazy things even if they are a little dangerous?	Self-control	Thrill Seeking	Revised	How often do you do crazy things, such as drinking alcohol, even if they are a little dangerous?
How often have you done what feels good no matter what?	Self-control	Thrill Seeking	Revised	How often do you do what feels good to you without thinking about its results?
How often have you done something dangerous because someone dared you to do it?	Self-control	Thrill Seeking	Revised	How often do you do something risky because of peer pressure?
How important is this to you: Avoiding risky behavior?	Self-control	Thrill Seeking	Unchanged	
<i>When answering these next four questions, think about the last few problems you have had in the past month, like when an object breaks</i>				<i>When answering these next four questions, think about the last few problems you have had in the past month, like when an object breaks</i>
How often did you try to fix the problems or take action to change things?	HOTS	Problem-solving	Revised	In the past month, how often did you take action to solve the problems?
How often did you ask other people for help or for ideas about how to solve the problem?	HOTS	Problem-solving	Revised	In the past month, how often did you ask other people for help with the problems?
How often did you try to think of different ways to change the problems or fix the situation?	HOTS	Problem-solving	Revised	In the past month, how often did you try to think of different ways to solve the problems?
How often did you develop a plan to solve the problem?	HOTS	Problem-solving	Revised	In the past month, how often did you make a plan to solve the problems?

How important is this to you: Knowing how to solve your problems?	HOTS	Problem-solving	Unchanged	
<i>When answering these next three questions, think about the last few times someone told you an interesting story</i>				<i>When answering these next three questions, think about the last few times someone told you an interesting story</i>
How often did you separate the true and false parts of the story?	HOTS	Critical Thinking	Unchanged	
How often did you find new ways of looking at the situation?	HOTS	Critical Thinking	Revised	How often did you question why someone in the story did what they did?
How often did you link facts together?	HOTS	Critical Thinking	Revised	How often did you connect pieces of evidence together?
How important is this to you: Being able to see how pieces of information relate to each other?	HOTS	Critical Thinking	Unchanged	
<i>When answering these next three questions, think about the last few times you made a decision.</i>				<i>When answering these next three questions, think about the last few times you made a decision.</i>
How often did you collect a lot of information before making the decisions?	HOTS	Decisions	Revised	Before making the decisions, how often did you collect a lot of information?
How often did you think about how it would affect others before making the decisions?	HOTS	Decisions	Revised	Before making the decisions, how often did you think about how others would be affected?
How often did you consider different choices before making the decisions?	HOTS	Decisions	Revised	Before making the decisions, how often did you consider different options?
How important is this to you: Knowing how to make good decisions?	HOTS	Decisions	Unchanged	
How often do you avoid making your agemates look bad?	Social skills	Social Skills	Revised	How often do you avoid making your friends look bad?

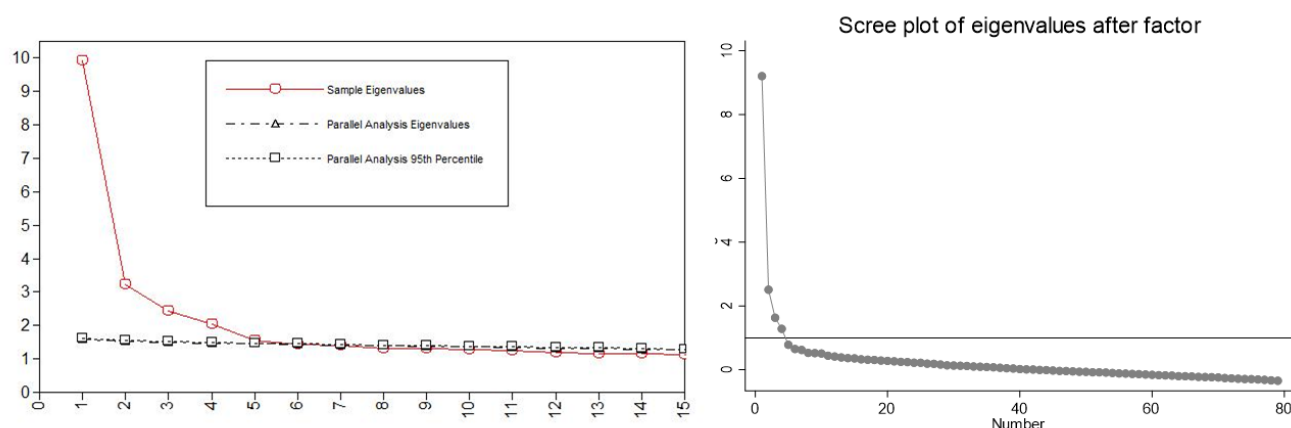
How often do you find a way to work things out if two of your friends disagree?	Social skills	Social Skills	Revised	How often do you find a way to work things out if two of your friends quarrel?
How often do you do your part when you work in a group?	Social skills	Social Skills	Unchanged	
How often do you get along well with people of different backgrounds?	Social skills	Social Skills	Revised	How often do you relate well with people of different backgrounds?
How often do you find it easy to make friends?	Social skills	Social Skills	Unchanged	
How often do you control your anger when you have a disagreement with a friend?	Social skills	Social Skills	Revised	How often do you control your anger when you have a misunderstanding with a friend?
How often do you respect other points of view, even if you disagree?	Social skills	Social Skills	Revised	How often do you respect views that differ from your own?
How important is this to you: Getting along well with others?	Social Skills	Social Skills	Revised	How important is this to you: Relating well with others?
How often do you write a story or composition well?	Communication	Communication	Revised	How often do you write a story or letter well?
How often do you listen to other students' ideas?	Communication	Communication	Revised	How often do you listen to your classmates' ideas?
How often can you discuss a problem with a friend without making things worse?	Communication	Communication	Unchanged	
How often are you embarrassed to ask questions in class?	Communication	Communication	Revised	How often are you uncomfortable to ask questions in class?
How often are you rude to others?	Communication	Communication	Unchanged	
How often do you let others know how you feel?	Communication	Communication	Revised	How often do you tell others how you feel?
How important is this to you: Being a good communicator?	Communication	Communication	Revised	How important is this to you: Being a good communicator?

Table A. 2*CFA estimation Study 1 - Baseline Uganda*

Item name	Item description	Estimate	S.E.	Est./S.E.	P V
self_efficacy1	18. How often do you believe that: it is hard for you to solve your problems?	-0.164	0.033	-4.911	0
self_efficacy2	19. How often do you believe that: there are many things that you do poorly?	-0.085	0.033	-2.577	0
self_efficacy3	20. How often do you believe that: you are good at learning something new?	0.463	0.03	15.57	0
self_efficacy4	21. How often do you believe that: you can do most things if you make the effort?	0.567	0.031	18.511	0
self_efficacy5	22. How often do you believe that: you can do something that will help you succeed in life?	0.507	0.03	17.103	0
self_efficacy_imp	23. How important is this to you: Being capable of doing most things?	0.486	0.029	16.796	0
self_esteem1	24. How often do you feel that you are not good at all?	-0.18	0.03	-5.945	0
self_esteem2	25. How often have you felt that: the people you live with at home value you?	0.412	0.031	13.329	0
self_esteem5	26. How often do you feel that you are a valued member of your community?	0.411	0.03	13.643	0
self_esteem4	27. How often have you felt that: you have a number of good qualities?	0.43	0.029	14.688	0
self_esteem_imp	28. How important is this to you: Liking yourself just the way you are?	0.388	0.032	12.261	0
self_confidence1	29. How often do you feel good about your skills?	0.533	0.03	18.064	0
self_confidence2	30. How often do you feel not sure that you can be successful?	-0.37	0.031	-11.962	0
self_confidence3	31. How often do you feel that you don't trust your skills?	-0.403	0.03	-13.308	0
self_confidence4	32. How often do you feel confident in yourself?	0.633	0.027	23.806	0
self_confidence_imp	33. How important is this to you: Being confident in yourself?	0.521	0.028	18.614	0
self_awareness1	34. How often do you find it hard to know how you are feeling?	-0.065	0.035	-1.865	0
self_awareness2	35. How often do you know what you are good at?	0.432	0.031	14.012	0
self_awareness3	36. How often do you know how you are feeling inside at any particular moment?	0.406	0.033	12.475	0
self_awareness4	37. How often do you know how you make other people feel?	0.308	0.032	9.667	0
self_awareness_imp	38. How important is this to you: Knowing how you feel about something?	0.479	0.032	15.071	0
self_belief1	39. How often do you see that your future will be happy?	0.553	0.03	18.159	0
self_belief2	40. How often do you believe that you will reach your future goals?	0.623	0.027	23.21	0
self_belief3	41. How often do you know that you are going to be fine?	0.518	0.029	18.031	0
self_belief4	42. How often do you believe you can make things happen that will improve your life?	0.587	0.028	20.788	0
self_belief_imp	43. How important is this to you: Believing in yourself?	0.563	0.03	18.695	0
gratification1	44. How often do you save your money for something you want to buy later?	0.346	0.04	8.607	0
gratification2	45. How often do you find it challenging to wait for something?	0.133	0.033	3.976	0
gratification3	46. How often would you prefer to get one pen now rather than	0.155	0.035	4.371	0

	many pens later?				
gratification_imp	47. How important is this to you: Being able to wait for what you want?	0.464	0.045	10.377	0
impulses1	48. How often do you do things without thinking about what you're doing?	-0.199	0.031	-6.357	0
impulses2	49. In the past month, how often have you interrupted your agemate when they were telling a story?	0.011	0.033	0.347	0
impulses3	50. How often do you think through things before you do them?	0.558	0.034	16.206	0
impulses_imp	51. How important is this to you: Controlling your behavior?	0.519	0.035	14.91	0
attention1	52. In the past month, how often have you finished the work that you set out to do despite challenges?	0.29	0.034	8.518	0
attention2	53. In the past month, how often have you been unable to pay attention?	-0.198	0.031	-6.447	0
attention3	54. In the past month, how often have you kept doing something that you should do even if you didn't like it, such as homework?	0.217	0.031	6.895	0
attention4	55. In the past month, how often have you found it difficult to start your work?	-0.044	0.027	-1.635	0
attention_imp	56. How important is this to you: Focusing your attention in order to complete something?	0.395	0.038	10.356	0
emotions1	61. In the past month, how often have you done things to control your anger or temper, for example when you have quarreled with your friend?	0.471	0.033	14.327	0
emotions2	62. In the past month, how often have you been annoyed by little things, like if someone steps on your shoe?	-0.174	0.03	-5.823	0
emotions3	63. In the past month, how often have you remained calm when a friend tells you that you did something poorly?	0.399	0.032	12.637	0
emotions_imp	64. How important is this to you: dealing with your feelings?	0.393	0.032	12.4	0
regulate_behaviors1	65. In the past month, how often were you able to stop yourself when you were going to do something you would regret?	0.408	0.033	12.446	0
regulate_behaviors2	66. In the past month, how often have you refused to follow instructions?	-0.272	0.032	-8.533	0
regulate_behaviors3	67. In the past month, how often have you got your work done immediately instead of waiting until the last minute?	0.41	0.032	12.985	0
regulate_imp	68. How important is this to you: Doing your best, even when you don't want to do something?	0.379	0.031	12.369	0
thrill_seeking1	69. How often do you do crazy things, such as drinking alcohol, even if they are a little dangerous?	0.35	0.046	7.597	0
thrill_seeking2	70. How often do you do what feels good to you without thinking about its results?	0.277	0.035	7.811	0
thrill_seeking3	71. How often do you do something risky because of peer pressure?	0.373	0.034	10.967	0
thrill_seeking_imp	72. How important is this to you: Avoiding risky behavior?	-0.603	0.041	-14.795	0
problemsolving1	74. In the past month, how often did you take action to solve the problems?	0.405	0.032	12.791	0
problemsolving2	75. In the past month, how often did you ask other people for help with the problems?	0.305	0.032	9.417	0

problemsolving3	76. In the past month, how often did you try to think of different ways to solve the problems?	0.501	0.029	17.418	0
problemsolving4	77. In the past month, how often did you make a plan to solve the problems?	0.5	0.028	18.039	0
problemsolving_imp	78. How important is this to you: Knowing how to solve your problems?	0.573	0.029	20.032	0
critical_thinking1	80. How often did you separate the true and false parts of the story?	0.323	0.034	9.591	0
critical_thinking2	81. How often did you question why someone in the story did what they did?	0.476	0.033	14.356	0
critical_thinking3	82. How often did you connect pieces of evidence together?	0.448	0.032	13.842	0
critical_thinking_imp	83. How important is this to you: Being able to see how pieces of information relate to each other?	0.476	0.03	15.75	0
decisions1	89. Before making the decisions, how often did you collect a lot of information?	0.416	0.029	14.275	0
decisions2	90. Before making the decisions, how often did you think about how others would be affected?	0.415	0.029	14.412	0
decisions3	91. Before making the decisions, how often did you consider different options?	0.409	0.029	14.288	0
decisions_imp	92. How important is this to you: Knowing how to make good decisions?	0.519	0.03	17.527	0
social_skills1	97. How often do you avoid making your agemates look bad?	0.33	0.03	11.058	0
social_skills2	98. How often do you find a way to work things out if two of your friends quarrel?	0.439	0.027	16.02	0
social_skills3	99. How often do you do your part when you work in a group?	0.469	0.028	16.476	0
social_skills4	100. How often do you relate well with people of different backgrounds?	0.429	0.028	15.448	0
social_skills5	101. How often do you find it easy to make friends?	0.331	0.029	11.342	0
social_skills6	102. How often do you control your anger when you have a misunderstanding with a friend?	0.452	0.027	16.686	0
social_skills7	103. How often do you respect views that differ from your own?	0.332	0.029	11.379	0
social_skills_imp	104. How important is this to you: Relating well with others?	0.515	0.026	19.784	0
communication1	109. How often do you write a story or letter well?	0.324	0.033	9.788	0
communication2	110. How often do you listen to your agemates' ideas?	0.418	0.032	13.155	0
communication3	111. How often can you discuss a problem with a friend without making things worse?	0.358	0.031	11.542	0
communication4	112. How often are you uncomfortable to ask questions in class?	-0.32	0.032	-9.939	0
communication5	113. How often are you rude to others?	-0.295	0.033	-8.988	0
communication6	114. How often do you tell others how you feel?	0.246	0.033	7.414	0
communication_imp	115. How important is this to you: Being a good communicator?	0.534	0.032	16.531	0

Figure A. 2*Parallel Analysis (Mplus) and Cattell scree test for EFA Study 2 - Uganda Baseline***Table A. 3***Parallel Analysis Study 2 (FACTOR)*

Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
16.9783*	2.6203	2.7306
5.7921*	2.5298	2.6344
3.8574*	2.466	2.5624
3.0987*	2.4074	2.4949
2.1344	2.3575	2.4357

Note: parallel analysis implemented using FACTOR, following Lorenzo-Seva & Ferrando (2006), based on minimum rank factor analysis as recommended by Timmerman & Lorenzo-Seva (2011). Polychoric correlation matrices used. Permutation of the raw data was performed to obtain random correlation matrices as suggested by Buja & Eyuboglu (1992). * Advised number of dimensions.

Table A. 4*Bayesian and other statistics for factor extraction EFA Study 2 - Baseline Uganda*

	Number of factors				
	1	2	3	4	5
Akaike (AIC)	245656.142	244079.599	243435.944	242937.255	242805.843
Bayesian (BIC)	246841.437	245654.991	245396.433	245277.838	245521.519
Sample-Size Adjusted BIC	246088.669	244654.477	244151.349	243791.361	243796.824
Chi-Square Test of Model Fit	7263.006	5530.463	4732.809	4082.12	3800.707
Degrees of Freedom	3002	2924	2847	2771	2696
P-Value	0	0	0	0	0
RMSEA	0.038	0.026	0.020	0.015	0.014

CFI	0.787	0.908	0.948	0.969	0.974
TLI	0.782	0.903	0.943	0.965	0.970
SRMR	0.065	0.046	0.039	0.034	0.032

Note: AIC, BIC and Chi-Square statistics were performed under the assumption items are continuous, while RMSEA, CFI, TLI, and SRMR were performed under the assumption items are categorical.

Table A. 5*EFA factor loadings Study 2 - Uganda Baseline*

Item name	Item description	Pos.	Import.	HOTS + SOC/COM	Neg.
self_efficacy1	18. How often do you believe that: it is hard for you to solve your problems?	-0.068	-0.074	0.055	0.229
self_efficacy2	19. How often do you believe that: there are many things that you do poorly?	-0.039	0.122	-0.006	0.324
self_efficacy3	20. How often do you believe that: you are good at learning something new?	0.407	0.117	0.044	-0.02
self_efficacy4	21. How often do you believe that: you can do most things if you make the effort?	0.465	0.118	0.146	0.03
self_efficacy5	22. How often do you believe that: you can do something that will help you succeed in life?	0.589	0.104	-0.063	0.021
self_efficacy_imp	23. How important is this to you: Being capable of doing most things?	0.161	0.609	-0.159	-0.026
self_esteem1	24. How often do you feel that you are not good at all?	-0.108	0.013	0.093	0.406
self_esteem2	25. How often have you felt that: the people you live with at home value you?	0.444	-0.017	0.083	-0.003
self_esteem5	26. How often do you feel that you are a valued member of your community?	0.43	-0.041	0.087	-0.051

self_esteem4	27. How often have you felt that: you have a number of good qualities?	0.402	0.034	0.102	-0.016
self_esteem_imp	28. How important is this to you: Liking yourself just the way you are?	0.113	0.485	-0.061	0.027
self_confidence1	29. How often do you feel good about your skills?	0.479	0.024	0.111	-0.017
self_confidence2	30. How often do you feel not sure that you can be successful?	-0.192	-0.041	-0.022	0.348
self_confidence3	31. How often do you feel that you don't trust your skills?	-0.308	-0.097	0.142	0.393
self_confidence4	32. How often do you feel confident in yourself?	0.544	0.072	0.088	-0.068
self_confidence_imp	33. How important is this to you: Being confident in yourself?	0.159	0.592	-0.111	0.005
self_awareness1	34. How often do you find it hard to know how you are feeling?	-0.068	0.028	0.175	0.38
self_awareness2	35. How often do you know what you are good at?	0.434	-0.076	0.11	-0.07
self_awareness3	36. How often do you know how you are feeling inside at any particular moment?	0.312	0.096	0.106	0.029
self_awareness4	37. How often do you know how you make other people feel?	0.142	-0.021	0.252	-0.003
self_awareness_imp	38. How important is this to you: Knowing how you feel about something?	0.037	0.528	0.04	-0.008
self_belief1	39. How often do you see that your future will be happy?	0.728	-0.071	-0.12	-0.045
self_belief2	40. How often do you believe that you will reach your future goals?	0.69	-0.042	0.027	0.033
self_belief3	41. How often do you know that you are going to be fine?	0.539	-0.026	0.012	-0.036
self_belief4	42. How often do you believe you can make things happen that will	0.5	0.087	0.073	0.021

	improve your life?				
self_belief_imp	43. How important is this to you: Believing in yourself?	0.142	0.674	-0.162	0.013
gratification1	44. How often do you save your money for something you want to buy later?	0.338	-0.073	0.165	-0.009
gratification2	45. How often do you find it challenging to wait for something?	0.036	0.162	0.095	0.202
gratification3	46. How often would you prefer to get one pen now rather than many pens later?	0.066	0.04	0.167	0.133
gratification_imp	47. How important is this to you: Being able to wait for what you want?	-0.005	0.549	0.059	-0.022
impulses1	48. How often do you do things without thinking about what you're doing?	-0.053	0.01	-0.015	0.369
impulses2	49. In the past month, how often have you interrupted your agemate when they were telling a story?	0.09	-0.021	0.048	0.197
impulses3	50. How often do you think through things before you do them?	0.258	0.129	0.297	-0.092
impulses_imp	51. How important is this to you: Controlling your behavior?	0.018	0.667	-0.021	-0.021
attention1	52. In the past month, how often have you finished the work that you set out to do despite challenges?	0.162	-0.022	0.262	-0.065
attention2	53. In the past month, how often have you been unable to pay attention?	0.061	-0.054	-0.098	0.403
attention3	54. In the past month, how often have you kept doing something that you should do even if you didn't like it, such as homework?	0.201	-0.045	0.162	-0.008
attention4	55. In the past month,	0.01	0.099	0.046	0.394

	how often have you found it difficult to start your work?				
attention_imp	56. How important is this to you: Focusing your attention in order to complete something?	-0.033	0.596	0.044	-0.012
emotions1	61. In the past month, how often have you done things to control your anger or temper, for example when you have quarreled with your friend?	0.178	0.013	0.378	-0.055
emotions2	62. In the past month, how often have you been annoyed by little things, like if someone steps on your shoe?	-0.042	0.039	-0.036	0.337
emotions3	63. In the past month, how often have you remained calm when a friend tells you that you did something poorly?	0.217	0.031	0.237	-0.044
emotions_imp	64. How important is this to you: dealing with your feelings?	-0.093	0.508	0.108	-0.009
regulate_behaviors1	65. In the past month, how often were you able to stop yourself when you were going to do something you would regret?	0.199	0.09	0.283	0.048
regulate_behaviors2	66. In the past month, how often have you refused to follow instructions?	0.071	0.008	-0.142	0.545
regulate_behaviors3	67. In the past month, how often have you got your work done immediately instead of waiting until the last minute?	0.316	0.058	0.16	-0.016
regulate_imp	68. How important is this to you: Doing your best, even when you don't want to do something?	-0.042	0.471	0.09	-0.011

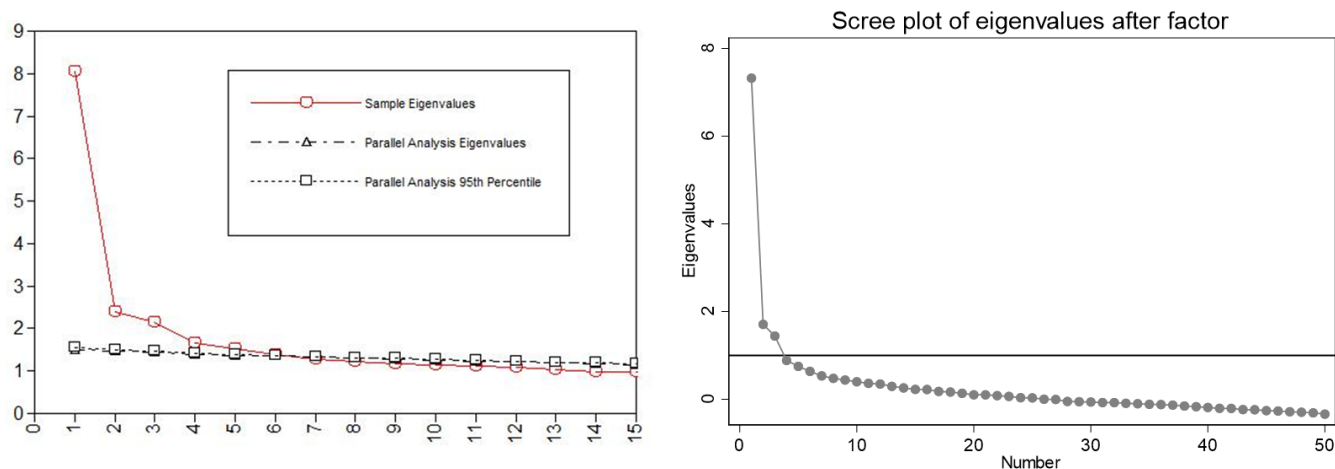
thrill_seeking1	69. How often do you do crazy things, such as drinking alcohol, even if they are a little dangerous?	-0.001	-0.102	-0.053	0.397
thrill_seeking2	70. How often do you do what feels good to you without thinking about its results?	-0.021	0.01	-0.089	0.335
thrill_seeking3	71. How often do you do something risky because of peer pressure?	-0.154	-0.024	-0.024	0.331
thrill_seeking_imp	72. How important is this to you: Avoiding risky behavior?	-0.013	0.551	0.054	-0.074
problemsolving1	74. In the past month, how often did you take action to solve the problems?	-0.031	0.044	0.439	-0.034
problemsolving2	75. In the past month, how often did you ask other people for help with the problems?	0.045	-0.025	0.345	0.031
problemsolving3	76. In the past month, how often did you try to think of different ways to solve the problems?	0.167	0.053	0.389	0.051
problemsolving4	77. In the past month, how often did you make a plan to solve the problems?	0.03	0.057	0.5	0.013
problemsolving_imp	78. How important is this to you: Knowing how to solve your problems?	0.044	0.627	0.03	0.027
critical_thinking1	80. How often did you separate the true and false parts of the story?	0.12	-0.019	0.263	0.016
critical_thinking2	81. How often did you question why someone in the story did what they did?	0.155	0.074	0.35	0.083
critical_thinking3	82. How often did you connect pieces of evidence together?	0.148	-0.015	0.396	0.07
critical_thinking_imp	83. How important is this to you: Being able to see how pieces of information	-0.096	0.484	0.2	0.064

	relate to each other?				
decisions1	89. Before making the decisions, how often did you collect a lot of information?	0.195	-0.068	0.404	-0.028
decisions2	90. Before making the decisions, how often did you think about how others would be affected?	0.103	0.045	0.389	-0.043
decisions3	91. Before making the decisions, how often did you consider different options?	0.118	0.089	0.345	-0.009
decisions_imp	92. How important is this to you: Knowing how to make good decisions?	0.06	0.645	-0.004	0.007
social_skills1	97. How often do you avoid making your agemates look bad?	-0.013	0.12	0.325	0.006
social_skills2	98. How often do you find a way to work things out if two of your friends quarrel?	0.096	0.058	0.407	0.001
social_skills3	99. How often do you do your part when you work in a group?	0.233	0.155	0.221	-0.001
social_skills4	100. How often do you relate well with people of different backgrounds?	0.099	0.006	0.426	-0.018
social_skills5	101. How often do you find it easy to make friends?	0.221	-0.045	0.24	0.005
social_skills6	102. How often do you control your anger when you have a misunderstanding with a friend?	0.035	0.013	0.424	-0.19
social_skills7	103. How often do you respect views that differ from your own?	0.194	0.078	0.162	0.014
social_skills_imp	104. How important is this to you: Relating well with others?	-0.056	0.539	0.164	-0.057
communication1	109. How often do you write a story or letter well?	0.051	-0.044	0.324	-0.065

communication2	110. How often do you listen to your agemates' ideas?	0.094	-0.009	0.355	-0.065
communication3	111. How often can you discuss a problem with a friend without making things worse?	0.001	-0.059	0.403	-0.104
communication4	112. How often are you uncomfortable to ask questions in class?	0.013	-0.057	-0.154	0.338
communication5	113. How often are you rude to others?	0.007	-0.02	-0.148	0.355
communication6	114. How often do you tell others how you feel?	-0.066	-0.003	0.323	-0.038
communication_imp	115. How important is this to you: Being a good communicator?	-0.091	0.522	0.18	-0.038

Figure A. 3

Parallel Analysis and Cattell scree test for EFA Study 3 - Guatemala Baseline

**Table A. 6**

Parallel Analysis Study 3 (FACTOR)

Real-data % of variance	Mean of random % of variance	95 th percentile of random % of variance
23.0757	4.1236	4.3524
5.9652	3.9375	4.1505
4.9830*	3.8	3.99
3.6034	3.6877	3.872
3.4248	3.5732	3.7323

Note: parallel analysis implemented using FACTOR, following Lorenzo-Seva & Ferrando (2006), based on minimum rank factor analysis as recommended by Timmerman & Lorenzo-Seva (2011). Polychoric correlation matrices used. Permutation of the raw data was performed to obtain random correlation matrices as suggested by Buja & Eyuboglu (1992).

** Advised number of dimensions when 95 percentile is considered. * Advised number of dimensions when mean is considered.

Table A. 7

Bayesian and other statistics for factor extraction EFA Study 3 - Baseline Guatemala

	Number of factors				
	1	2	3	4	5
Akaike (AIC)	74314.157	73761.425	73241.236	73061.194	72940.996
Bayesian (BIC)	75018.897	74696.38	74401.708	74442.485	74538.407
Sample-Size Adjusted BIC	74542.559	74064.437	73617.337	73508.861	73458.707
Chi-Square Test of Model Fit	3569.609	2918.876	2302.688	2028.646	1816.448
Degrees of Freedom	1175	1126	1078	1031	985
P-Value	0	0	0	0	0
RMSEA	0.051	0.043	0.037	0.034	0.031
CFI	0.838	0.889	0.924	0.936	0.95
TLI	0.831	0.879	0.913	0.924	0.938
SRMR	0.072	0.06	0.049	0.045	0.041

Note: AIC, BIC and Chi-Square statistics were performed under the assumption items are continuous, while RMSEA, CFI, TLI, and SRMR were performed under the assumptions items are categorical.

Table A. 8

EFA factor loadings Study 3 - Guatemala Baseline

Item name	Item description	Pos.	Soc & Com	HOTS	Neg
self_efficacy2	4. There are many things that I can't do very well	0.093	-	0.026	0.288
self_efficacy3	5. I'm good at learning new things	0.499	0.1	0.032	-
self_efficacy4	6. I can do most things if I make an effort	0.427	0.027	0.12	0.086
self_efficacy5	7. I can do things that will help me succeed in life	0.639	0.042	0.001	0.098
self_esteem1	8. I think I am no good at all.	0.187	0.048	0.082	0.568
self_esteem2	9. I feel valued by the people I live with at home	0.556	0.069	0.051	0.122
self_esteem3	10. I'm a valued member of my community	0.308	0.231	0.035	0.024
self_esteem4	11. I have a number of good qualities.	0.499	0.14	-0.08	0.128
self_esteem5	12. I like myself just the way I am	0.731	-	-	-

			0.001	0.015	0.009
self_confidence1	13. I feel good about my skills	0.645	0.024	0.077	0.008
self_confidence2	14. I'm not sure I can be successful	-	-	-	-
self_confidence3	15. I'm not confident about my skills	0.225	0.004	0.054	0.415
self_confidence4	16. I feel confident in myself	-	-	-	-
self_awareness1	17. It is hard to know what I'm feeling	0.287	0.063	0.062	0.542
self_awareness2	18. I know what I'm good at	-	-	-	-
self_awareness3	19. I know how I'm feeling inside at any particular moment	0.666	0.043	0.113	0.021
self_belief1	20. My future will be happy	-	-	-	-
self_belief2	21. I can achieve most of my future goals.	0.181	0.128	0.023	0.312
self_belief3	22. I know I'm going to be fine	0.471	0.098	0.113	0.064
self_belief4	23. I can make things happen that will improve my life	0.262	0.193	0.129	0.04
impulses1	24. I do things before I think through them	0.622	0.132	0.04	0.109
impulses2	25. I think carefully before doing anything	-	-	-	-
attention1	26. I have a hard time concentrating on one thing.	0.629	0.009	0.076	0.012
attention2	27. I have difficulty starting tasks	0.587	0.211	0.139	0.124
emotions1	30. When things go wrong for me, I'm good controlling my temper	-	-	-	-
emotions2	31. I'm easily annoyed by little things (like if someone steps on my shoe)	0.557	0.029	0.158	0.005
emotions3	32. If a friend tells me I did something wrong, I can stay calm	-	-	-	-
regulate_behaviors1	33. If I'm doing something that I know I would regret, I'm able to stop before it is too late	0.007	0.055	0.224	0.561
regulate_behaviors2	34. I'm good at following instructions	0.182	0.06	0.383	0.143
thrill_seeking1	35. I do whatever feels good to me, without thinking about the results	-	-	-	-
thrill_seeking2	36. If my friends are doing something risky, I will do it with them	0.139	0.282	0.03	0.391
problemsolving1	37. I took action to solve the problems	-	-	-	-
problemsolving2	38. I asked other people for help to solve the problems	0.144	0.072	0.168	0.498
problemsolving3	39. I tried to think of different ways to solve the problems	-	-	-	-
problemsolving4	40. I made a plan to solve the problems	0.029	0.28	0.323	0.035
		-	-0.02	0.074	0.451
		-	-	-	-
		0.057	0.048	0.271	0.034
		-	-	-	-
		0.25	0.033	0.204	0.215
		-	-	-	-
		0.161	0.275	0.203	0.189
		-	-	-	-
		0.008	0.014	0.218	0.555
		-0.19	0.175	0.096	0.421
		-	-	-	-
		0.036	0.055	0.571	0.047
		0.034	0.001	0.399	0.005
		0.059	0.052	0.603	0.019
		-	0.09	0.529	0.075

		0.009	-	-	
critical_thinking1	41. I questioned why someone in the story did what they did	0.057	0.206	0.396	0.024
critical_thinking2	42. I connected pieces of evidence together	0.058	0.057	0.347	0.102
decisions1	45. I collected a lot of information before making the decision	0.082	0.052	0.486	0.05
decisions2	46. I thought about how other people would be affected before making the decision	-	-	-	-
decisions3	47. I considered different options before making the decision	0.014	0.044	0.461	0.001
social_skills1	50. I'm able to stand up for myself without putting others down	0.084	0.027	0.493	0.104
social_skills2	51. I find a way of working things out if two of my friends quarrel	0.106	0.283	0.258	0.133
social_skills3	52. I get along well with people from different backgrounds	0.126	0.223	0.13	0.132
social_skills4	53. I find it easy to make friends	0.155	0.449	0.088	0.048
social_skills5	54. I can control my anger when I have a misunderstanding with a friend	0.094	0.538	0.043	0.043
communication1	57. I write well.	-	-	-	-
communication2	58. I am good at resolving disagreements.	0.036	0.305	0.297	0.067
communication3	59. It is easy for me to ask questions in a public setting.	0.169	0.31	0.085	0.019
communication4	60. I am rude to others.	0.028	0.482	0.218	0.003
communication5	61. It is easy for me to share my feelings with others.	0.053	0.429	0.021	0.016
		-	-	-	-
		0.037	0.129	0.093	0.435
		0.02	0.46	0.011	0.068

Table A. 9*Item revision after Study 2 EFA (5 versus 4 factor structure)*

Original Item	Factor according to 5-factor theoretical structure	Subscale	Decision	Revised Item	Factor according to 4-factor EFA-determined structure
When answering these questions, please think about times when you were healthy and had your basic needs, including school fees, met.				When answering these questions, please think about times when you were healthy and had your basic needs, including school fees, met.	

I'm going to ask you some questions about how you solve problems and do tasks, and some questions about how you feel about yourself. Remember there are no right or wrong answers. I just want to learn about you and what you think.				Please tell us how much you agree with each statement. There are no right or wrong answers, just tell us what you think.	
How often do you believe that: it is hard for you to solve your problems?	Positive Self-Concept	Self-Efficacy	Excluded		
How often do you believe that: there are many things that you do poorly?	Positive Self-Concept	Self-Efficacy	Revised	There are many things that I can't do very well.	Negative Self-Concept
How often do you believe that: you are good at learning something new?	Positive Self-Concept	Self-Efficacy	Revised	I'm good at learning new things.	Positive Self-Concept
How often do you believe that: you can do most things if you make the effort?	Positive Self-Concept	Self-Efficacy	Revised	I can do most things if I make an effort.	Positive Self-Concept
How often do you believe that: you can do something that will help you succeed in life?	Positive Self-Concept	Self-Efficacy	Revised	I can do things that will help me succeed in life.	Positive Self-Concept
How important is this to you: Being capable of doing most things?	Positive Self-Concept	Self-Efficacy	Excluded		
How often do you feel that you are not good at all?	Positive Self-Concept	Self-Esteem	Revised	I think I am no good at all.	Negative Self-Concept
How often have you felt that: the people you live with at home value you?	Positive Self-Concept	Self-Esteem	Revised	I feel valued by the people I live with at home.	Positive Self-Concept
How often do you feel that you are a valued member of your community?	Positive Self-Concept	Self-Esteem	Revised	I'm a valued member of my community.	Positive Self-Concept
How often have you felt that: you have a number of good qualities?	Positive Self-Concept	Self-Esteem	Revised	I have a number of good qualities.	Positive Self-Concept
How important is this to you: Liking yourself just the way you are?	Positive Self-Concept	Self-Esteem	Excluded		
			New Item Added	I like myself just the way I am.	Positive Self-Concept

How often do you feel good about your skills?	Positive Self-Concept	Self-Confidence	Revised	I feel good about my skills.	Positive Self-Concept
How often do you feel not sure that you can be successful?	Positive Self-Concept	Self-Confidence	Revised	I'm not sure I can be successful.	Negative Self-Concept
How often do you feel that you don't trust your skills?	Positive Self-Concept	Self-Confidence	Revised	I'm not confident about my skills.	Negative Self-Concept
How often do you feel confident in yourself?	Positive Self-Concept	Self-Confidence	Revised	I feel confident in myself.	Positive Self-Concept
				Now I will read some statements about you. Tell me how you agree with the statements.	
How important is this to you: Being confident in yourself?	Positive Self-Concept	Self-Confidence	Excluded		
How often do you find it hard to know how you are feeling?	Positive Self-Concept	Self-Awareness	Revised	It is hard to know what I'm feeling.	Negative Self-Concept
How often do you know what you are good at?	Positive Self-Concept	Self-Awareness	Revised	I know what I'm good at.	Positive Self-Concept
How often do you know how you are feeling inside at any particular moment?	Positive Self-Concept	Self-Awareness	Revised	I know how I'm feeling inside at any particular moment.	Positive Self-Concept
How often do you know how you make other people feel?	Positive Self-Concept	Self-Awareness	Excluded		
How important is this to you: Knowing how you feel about something?	Positive Self-Concept	Self-Awareness	Excluded		
How often do you see that your future will be happy?	Positive Self-Concept	Self-Belief	Revised	My future will be happy.	Positive Self-Concept
How often do you believe that you will reach your future goals?	Positive Self-Concept	Self-Belief	Revised	I can achieve most of my future goals.	Positive Self-Concept
How often do you know that you are going to be fine?	Positive Self-Concept	Self-Belief	Revised	I know I'm going to be fine.	Positive Self-Concept
How often do you believe you can make things happen that will improve your life?	Positive Self-Concept	Self-Belief	Revised	I can make things happen that will improve my life.	Positive Self-Concept

How important is this to you: Believing in yourself?	Positive Self-Concept	Self-Belief	Excluded		
How often do you save your money for something you want to buy later?	Self-control	Gratification	Excluded		
How often do you find it challenging to wait for something?	Self-control	Gratification	Excluded		
How often would you prefer to get one pen now rather than many pens later?	Self-control	Gratification	Excluded		
How important is this to you: Being able to wait for what you want?	Self-control	Gratification	Excluded		
How often do you do things without thinking about what you're doing?	Self-control	Impulses	Revised	I do things before I think through them.	Negative Self-Concept
In the past month, how often have you interrupted your friend when they were telling a story?	Self-control	Impulses	Excluded		
How often do you think through things before you do them?	Self-control	Impulses	Revised	I think carefully before doing anything.	HOTS
How important is this to you: Controlling your behavior?	Self-control	Impulses	Excluded		
In the past month, how often have you finished the work that you set out to do despite challenges?	Self-control	Attention	Excluded		
In the past month, how often have you been unable to pay attention?	Self-control	Attention	Revised	I have a hard time concentrating on one thing.	Negative Self-Concept
In the past month, how often have you kept doing something that you should do even if you didn't like it, such as homework?	Self-control	Attention	Excluded		
In the past month, how often have you found it difficult to start your work?	Self-control	Attention	Revised	I have difficulty starting tasks.	Negative Self-Concept
How important is this to you: Focusing your attention in order to complete something?	Self-control	Attention	Excluded		

In the past month, how often have you done things to control your anger or temper, for example when you have quarreled with your friend?	Self-control	Emotions	Revised	When things go wrong for me, I'm good controlling my temper.	HOTS
In the past month, how often have you been annoyed by little things like if someone	Self-control	Emotions	Revised	I'm easily annoyed by little things (like if someone steps on my shoe)	Negative Self-Concept
In the past month, how often have you remained calm when a friend tells you that you did something poorly?	Self-control	Emotions	Revised	If a friend tells me I did something wrong, I can stay calm.*	n/a
How important is this to you: dealing with your feelings?	Self-control	Emotions	Excluded		
In the past month, how often were you able to stop yourself when you were going to do something you would regret?	Self-control	Regulate Behaviors	Revised	If I'm doing something that I know I would regret, I'm able to stop before it is too late.	Positive Self-Concept
In the past month, how often have you refused to follow instructions?	Self-control	Regulate Behaviors	Revised	I'm good at following instructions.	Soc. & Com. Skills
In the past month, how often have you got your work done immediately instead of waiting until the last minute?	Self-control	Regulate Behaviors	Excluded		
How important is this to you: Doing your best, even when you don't want to do something?	Self-control	Regulate Behaviors	Excluded		
How often do you do crazy things, such as drinking alcohol, even if they are a little dangerous?	Self-control	Thrill Seeking	Excluded		
How often do you do what feels good to you without thinking about its results?	Self-control	Thrill Seeking	Revised	I do whatever feels good to me, without thinking about the results	Negative Self-Concept
How often do you do something risky because of peer pressure?	Self-control	Thrill Seeking	Revised	If my friends are doing something risky, I will do it with them	Negative Self-Concept
How important is this to you: Avoiding risky behavior?	Self-control	Thrill Seeking	Excluded		
When answering these next four questions, think about the last few problems you have had in the past month,				When answering these next four questions, think about the last few problems you have had and tell us how	

like when an object breaks				much you agree with each statement.	
In the past month, how often did you take action to solve the problems?	HOTS	Problem-solving	Revised	I took action to solve the problems.	HOTS
In the past month, how often did you ask other people for help with the problems?	HOTS	Problem-solving	Revised	I asked other people for help to solve the problems.	HOTS
In the past month, how often did you try to think of different ways to solve the problems?	HOTS	Problem-solving	Revised	I tried to think of different ways to solve the problems.	HOTS
In the past month, how often did you make a plan to solve the problems?	HOTS	Problem-solving	Revised	I made a plan to solve the problems.	HOTS
How important is this to you: Knowing how to solve your problems?	HOTS	Problem-solving	Excluded		
When answering these next three questions, think about the last few times someone told you an interesting story				When answering these next two questions, think about the last few times someone told you an interesting story and tell us how much you agree with each statement.	
How often did you separate the true and false parts of the story?	HOTS	Critical Thinking	Excluded		
How often did you question why someone in the story did what they did?	HOTS	Critical Thinking	Revised	I questioned why someone in the story did what they did.	HOTS
How often did you connect pieces of evidence together?	HOTS	Critical Thinking	Revised	I connected pieces of evidence together.	HOTS
How important is this to you: Being able to see how pieces of information relate to each other?	HOTS	Critical Thinking	Excluded		
When answering these next three questions, think about the last few times you made a decision.				When answering these next three questions, think about the last few times you made a decision and tell us how much you agree with each statement.	
Before making the decisions, how often did you collect a	HOTS	Decisions	Revised	I collected a lot of information before making	HOTS

lot of information?				the decision.	
Before making the decisions, how often did you think about how others would be affected?	HOTS	Decisions	Revised	I thought about how other people would be affected before making the decision.	HOTS
Before making the decisions, how often did you consider different options?	HOTS	Decisions	Revised	I considered different options before making the decision.	HOTS
How important is this to you: Knowing how to make good decisions?	HOTS	Decisions	Excluded		
How often do you avoid making your friends look bad?	Social skills	Social Skills	Revised	I'm able to stand up for myself without putting others down.	Soc. & Com. Skills
How often do you find a way to work things out if two of your friends quarrel?	Social skills	Social Skills	Revised	I find a way of working things out if two of my friends quarrel.*	n/a
How often do you do your part when you work in a group?	Social skills	Social Skills	Excluded		
How often do you relate well with people of different backgrounds?	Social skills	Social Skills	Revised	I get along well with people from different backgrounds.	HOTS
How often do you find it easy to make friends?	Social skills	Social Skills	Revised	I find it easy to make friends.	HOTS
How often do you control your anger when you have a misunderstanding with a friend?	Social skills	Social Skills	Revised	I can control my anger when I have a misunderstanding with a friend.	Soc. & Com. Skills
How often do you respect views that differ from your own?	Social skills	Social Skills	Revised		
How important is this to you: Relating well with others?	Social Skills	Social Skills	Excluded		
How often do you write a story or letter well?	Communication	Communication	Revised	I write well.	Soc. & Com. Skills
How often do you listen to your agemates' ideas?	Communication	Communication			
How often can you discuss a problem with a friend without making things worse?	Communication	Communication	Revised	I am good at resolving disagreements.	Soc. & Com. Skills
How often are you uncomfortable to ask	Communication	Communication	Revised	It is easy for me to ask questions in public.	Soc. & Com. Skills

questions in class?					
How often are you rude to others?	Communi- cation	Communi- cation	Revised	I am rude to others.	Negative Self- Concept
How often do you tell others how you feel?	Communi- cation	Communi- cation	Revised	It is easy for me to share my feelings with others.	Soc. & Com. Skills
How important is this to you: Being a good communicator?	Communi- cation	Communi- cation	Excluded		

Note: The response options for each item were initially designed as a 5-point behavioral frequency (“1.Never or Almost Never”, “2.Rarely”, “3.Sometimes”, “4.Often”, and “5.Almost Always or always”) and was changed to a 4-point endorsement scale (“1.Strongly Disagree” “2.Disagree”, “3.Agree”, “4.Strongly Agree”) after the tool revision. Two items indicated with an asterisk (*) were dropped after the EFA analysis of Study 2 (baseline Guatemala) and were not included in the final version of the tool implemented at endline in both countries. The final version of the tool can be accessed online through the following link:

https://www.youthpower.org/sites/default/files/YouthPower/files/resources/YP%20Action%20Youth_Tool_English-v5.pdf

Table A. 10

CFA estimation Study 3 - Endline Uganda and Guatemala

CFA Estimation Study 3 - Example Uganda and Guatemala									
Item Name	Item Description	Load	S.E.	Est./ S.E.	P- Val	Load	S.E.	Est./ S.E.	P- Val
		Uganda				Guatemala			
Positive Self-Concept									
self_efficacy3	I'm good at learning new things	0.608	0.025	24.81	0	0.615	0.028	21.79	0
self_efficacy4	I can do most things if I make an effort	0.536	0.029	18.28	0	0.552	0.028	19.93	0
self_efficacy5	I can do things that will help me succeed in life	0.647	0.025	25.86	0	0.701	0.025	27.95	0
self_esteem2	I feel valued by the people I live with at home	0.585	0.026	22.81	0	0.585	0.027	21.31	0
self_esteem3	I'm a valued member of my community	0.526	0.026	20.01	0	0.495	0.033	15.18	0
self_esteem4	I have a number of good qualities.	0.552	0.026	20.84	0	0.659	0.028	23.96	0
self_esteem5	I like myself just the way I am	0.649	0.027	24.38	0	0.749	0.02	37.92	0
self_confidence1	I feel good about my skills	0.589	0.026	22.9	0	0.722	0.022	33.12	0
self_confidence4	I feel confident in myself	0.683	0.023	29.55	0	0.685	0.023	29.77	0
self_awareness2	I know what I'm good at	0.609	0.023	26.8	0	0.703	0.022	31.33	0
self_awareness3	I know how I'm feeling inside at any particular moment	0.507	0.028	18.4	0	0.471	0.03	15.45	0
self_belief1	My future will be happy	0.521	0.029	17.68	0	0.684	0.024	28.48	0
self_belief2	I can achieve most of my future goals.	0.486	0.03	16.31	0	0.68	0.023	29.5	0
self_belief3	I know I'm going to be fine	0.538	0.028	19.46	0	0.6	0.026	22.68	0
self_belief4	I can make things happen that will improve my life	0.606	0.026	23.36	0	0.663	0.024	27.62	0

regulate_behaviors1	If I'm doing something that I know I would regret, I'm able to stop before it is too late	0.563	0.026	21.81	0	0.445	0.031	14.21	0
Negative Self-Concept									
self_efficacy2	There are many things that I can't do very well	0.057	0.04	1.422	0.155	0.251	0.043	5.863	0
self_esteem1	I cannot do anything well	0.483	0.034	14.38	0	0.701	0.03	23.33	0
self_confidence2	I'm not sure I can be successful	0.611	0.029	20.93	0	0.741	0.026	29.03	0
self_confidence3	I'm not confident about my skills	0.563	0.031	18.17	0	0.664	0.029	22.97	0
self_awareness1	It is hard to know what I'm feeling	0.179	0.036	4.983	0	0.103	0.045	2.285	0.022
impulses1	I do things before I think through them	0.662	0.029	22.58	0	0.597	0.03	19.69	0
attention1	I have a hard time concentrating on one thing.	0.224	0.037	6.111	0	0.379	0.039	9.671	0
attention2	I have difficulty starting tasks	0.247	0.037	6.644	0	0.557	0.034	16.5	0
emotions2	I'm easily annoyed by little things (like if someone steps on my shoe)	0.455	0.035	13.01	0	0.489	0.037	13.29	0
thrill_seeking1	I do whatever feels good to me, without thinking about the results	0.454	0.034	13.47	0	0.571	0.034	16.57	0
thrill_seeking2	If my friends are doing something risky, I will do it with them	0.598	0.033	18.28	0	0.644	0.033	19.54	0
communication4	I am rude to others.	0.508	0.035	14.36	0	0.488	0.038	12.87	0
HOTS									
impulses2	I think carefully before doing anything	0.67	0.024	27.86	0	0.623	0.029	21.78	0
emotions1	When things go wrong for me, I'm good controlling my temper	0.531	0.029	18.64	0	0.588	0.029	20.56	0
problemsolving1	I took action to solve the problems	0.328	0.031	10.55	0	0.564	0.031	18.4	0
problemsolving2	I asked other people for help to solve the problems	0.467	0.03	15.78	0	0.452	0.033	13.75	0
problemsolving3	I tried to think of different ways to solve the problems	0.537	0.027	20.03	0	0.578	0.03	19.4	0
problemsolving4	I made a plan to solve the problems	0.506	0.027	18.41	0	0.455	0.031	14.84	0
critical_thinking1	I questioned why someone in the story did what they did	0.431	0.029	14.97	0	0.212	0.04	5.318	0
critical_thinking2	I connected pieces of evidence together	0.426	0.031	13.83	0	0.167	0.037	4.542	0
decisions1	I collected a lot of information before making the decision	0.493	0.027	17.99	0	0.634	0.027	23.11	0
decisions2	I thought about how other people would be affected before making the decision	0.413	0.03	13.64	0	0.445	0.033	13.56	0
decisions3	I considered different options before making the decision	0.58	0.025	23.43	0	0.641	0.027	23.41	0
Social and Communication Skills									
regulate_behaviors2	I'm good at following instructions	0.574	0.027	21.08	0	0.683	0.026	26.74	0

social_skills1	I'm able to stand up for myself without putting others down	0.536	0.027	19.74	0	0.579	0.03	19.02	0
social_skills3	I get along well with people from different backgrounds	0.464	0.029	16.27	0	0.563	0.03	18.76	0
social_skills4	I find it easy to make friends	0.553	0.028	19.74	0	0.588	0.028	21.24	0
social_skills5	I can control my anger when I have a misunderstanding with a friend	0.503	0.029	17.16	0	0.571	0.027	21.25	0
communication1	I write well.	0.455	0.031	14.51	0	0.503	0.035	14.39	0
communication2	I am good at resolving disagreements.	0.435	0.032	13.77	0	0.495	0.03	16.28	0
communication3	It is easy for me to ask questions in a public setting.	0.514	0.027	19.05	0	0.44	0.034	12.88	0
communication5	It is easy for me to share my feelings with others.	0.365	0.03	12.34	0	0.31	0.035	8.804	0

Table A. 11

Mean differences and change over time for Guatemala baseline and endline – Non-Repeating Sample – Study 3

	Base.Mean	Base. SD	End Mean	End SD	Pooled SD	SE	t-score	Diff (E-B)
Positive Self Concept	3.38	0.31	3.25	0.31	0.31	0.024	-5.68	-0.13*
Negative Self Concept	2.04	0.32	2.21	0.30	0.31	0.024	7.04	0.17*
HOTS	2.97	0.36	2.92	0.30	0.33	0.025	-1.91	-0.05
Communication & Social Skills	2.95	0.39	2.90	0.36	0.38	0.029	-1.70	-0.05
Observations	344		334					

Notes: Significance is denoted as: * $p < 0.05$. Non-Repeating Sample contains youth who were sampled only once, either at baseline or at endline.

Table A. 12

Correlation baseline and endline scores for the panel sample in Guatemala – Study 3

Positive Self Concept	0.36*
Negative Self Concept	0.40*
HOTS	0.36*
Communication & Social Skills	0.45*

Note: 450 observations. Significance is denoted as: * $p < 0.05$. Panel Sample contains the youth in Guatemala who were interviewed both at baseline and endline.

Table A. 13

Selected outcome variables description – Study 3

Variable	Composition
Employ (Employment Score)	Average of the following binary variables, coded 1 for “Yes” and 0 for “No”: 1) In the last 3 months, did you receive payment for any work that you did? 2) Have you ever been interviewed for a job? 3) Have you ever received a job offer? 4) Are you running my own business?
Disability	Following the Washington Group Short Set of Questions, the list of disability questions included: difficulty seeing, even if wearing glasses; difficulty hearing, even if using a hearing aid; difficult walking or climbing steps; difficulty remembering or concentrating; difficulty (with self-care such as) bathing or dressing; and difficulty communicating, for example understanding or being understood. Response options included no difficulty; some difficulty; a lot of difficulty; cannot do at all. “Any Disability” variable assumes value 1 if students answered having at least some difficult to any of the questions and zero if they answered no difficulty to all of the questions.
Had Sex	Indicates whether the youth reported ever having had sex.
Violence	Students were asked how many times in the past month they engaged in the following actions: insulted someone else’s family (i.e. said something bad about them); made fun of or mocked someone else to make them angry; shamed or embarrassed someone to their face; not let someone be a part of your group anymore because you were upset or angry at them; and said mean things about someone to make others laugh. “Any Violence” variable assumes value 1 if students answered once or more for any of the questions, and zero if they answered zero for all of the questions.

Table A. 14*Descriptive statistics of selected outcome variables – Study 3*

	Guatemala			Uganda		
	N	Mean	SD	N	Mean	SD
Employ	1,010	0.31	0.3	795	0.24	0.24
Had Sex	1,005	0.23	0.42	747	0.12	0.33
Disability	1,010	0.43	0.5	794	0.51	0.5
Violence	1,004	0.68	0.47	793	0.23	0.49

Note: Statistics at baseline showed for Guatemala and at endline for Uganda. Table A. 13 contains description and definitions of the selected outcome variables.

Table A. 15*Description of Anchoring vignette items*

Self-Efficacy AV High	Mary is usually a good student. She often enjoys learning, and usually gets good marks at school. Most of the time, she believes she will do well on assignments. Based on this information, please tell me how much you agree with this statement: Mary believes she can be successful in school.
Self-Efficacy AV Low	Betty often dislikes learning. She does not always get good marks at school, and often does not believe she will do well on assignments. Based on this information, please tell me how much you agree with this statement: Betty believes she can be successful in school.

Emotions AV High	John was criticized by his friend for something John thought he did well. John did not like the criticism, but remained calm and did not get annoyed by it. John continued the conversation and asked questions to learn more about what he might do differently. Based on this information, please tell me how much you agree with the following statement: John is good at controlling his emotions.
Emotions AV Low	Moses was criticized by his friend for something Moses thought he did well. Moses did not like the criticism, and got angry and shouted at his friend. Based on this information, please tell me how much you agree with the following statement: Moses is good at controlling his emotions.
Decisions AV High	When Grace was deciding where to work, she tried to learn about different jobs by gathering information, talking to others, and thinking about how the job would affect her and others. Based on this information, please tell me how much you agree with the following statement: Grace is good at decision-making.
Decisions AV Low	When Sarah was deciding where to work, she decided on the spot without gathering information, talking to others, or thinking about how the job would affect her and others. Based on this information, please tell me how much you agree with the following statement: Sarah is good at decision-making.
Social Skills AV High	David is usually understanding and kind to others. He is often trusting, and usually finds it easy to cooperate with others. Based on this information, please tell me how much you agree with the following statement: David will relate well with others on his team.
Social Skills AV Low	Sam is sometimes kind and sometimes rude to others. He does not always trust people, and sometimes he finds it difficult to cooperate with others. Based on this information, please tell me how much you agree with the following statement: Sam relates well with others on his team.
Communication AV High	Joy tends to agree with others, and she expresses her ideas well. Many people consider Joy a good listener. Based on this information, please tell me how much you agree with the following statement: Joy communicates well to her team members.
Communication AV Low	Diana often disagrees with others. She sometimes starts quarrels. Some people consider Diana rude. Based on this information, please tell me how much you agree with the following statement: Diana communicates well to her team members.
<i>Note.</i> Response options for the Anchoring vignette items were: " Strongly Agree" / "Agree" / "Disagree" / "Strongly Disagree". The AV item for each scale/subscale was asked before the first item of each scale	

Table A. 16*Scale means for raw and anchoring vignette (AV) adjusted scores – Study 3*

Positive self concept	Negative self concept	HOTS	Social & communication skills
--------------------------	--------------------------	------	----------------------------------

	N	Raw	AV	Raw	AV	Raw	AV	Raw	AV 1	AV 2
Uganda Endline	1010	3.48	3.69	2.03	2.55	3.24	3.51	3.23	4.32	3.44
Guatemala Baseline	794	3.38	3.86	2.04	2.52	2.98	3.55	2.97	3.52	3.38
Guatemala Endline	784	3.27	3.66	2.17	2.69	2.94	3.39	2.93	3.40	3.21
Uganda Retest 1	57	3.47	3.61	1.97	2.63	3.27	3.50	3.24	4.35	3.39
Uganda Retest 2	57	3.49	3.66	1.84	2.49	3.30	3.47	3.30	4.47	3.41
Guatemala Retest 1	126	3.33	3.64	2.09	2.64	2.98	3.40	2.97	3.56	3.27
Guatemala Retest 2	126	3.29	3.61	2.05	2.65	3.00	3.38	3.03	3.52	3.43

Note. The scale for the raw scores is 1 to 4; the scale for the AV adjusted scores is 1 to 5 (2 times the number of AVs plus 1). There are two independent AV adjustments for Factor 4 due to having two sets of AVs for that factor.

Figure A. 4

Anchoring Vignettes, Response Analysis – Study 3

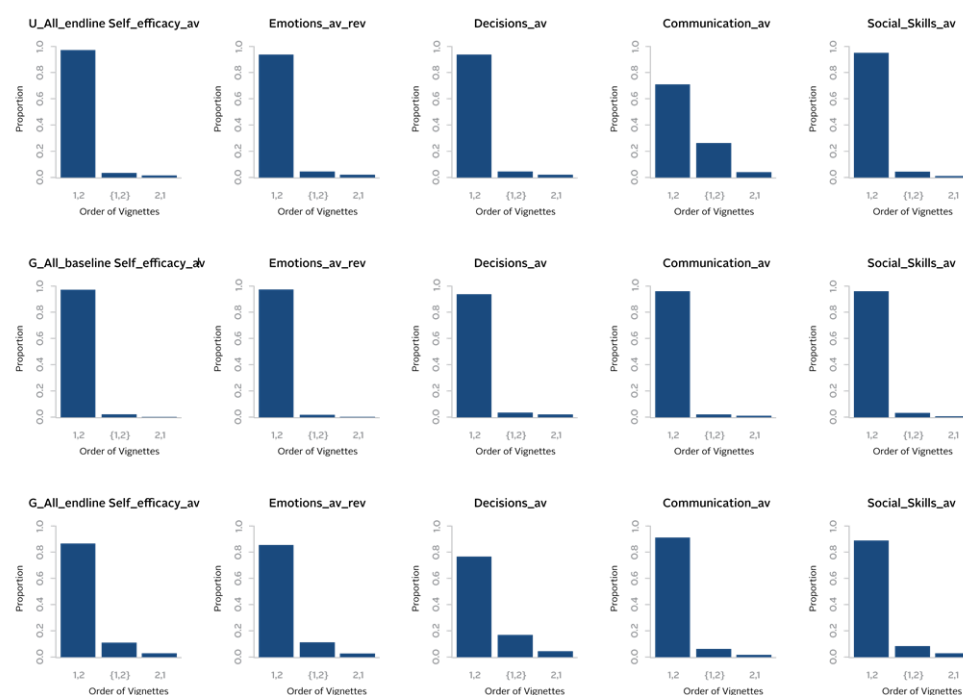


Table A. 17

Cronbach's alpha AV reliability using AV-adjusted scores – Study 3

Dimension	Uganda		Guatemala	
	Endline		Baseline	Endline

Positive self concept	0.94	0.94	0.95
Negative self concept	0.85	0.81	0.89
HOTS	0.89	0.86	0.92
Social & communication skills I	0.84	0.79	0.85
Social & communication skills II	0.86	0.81	0.87
Sample Size	1098	794	784

Table A. 18

Item descriptive information and percent of response by response category for Uganda baseline data - Study 3

Item name	Item descriptive information				Percent of responses by response category				
	Obs	Mean	Std. Dev.	Median	1	2	3	4	5
self_efficacy1	1097	2.69	1.05	3	14	24	48	5	8
self_efficacy2	1098	2.62	0.99	3	15	27	46	7	5
self_efficacy3	1096	4.23	0.99	5	1	6	18	19	56
self_efficacy4	1098	4.44	0.90	5	1	4	13	16	67
self_efficacy5	1097	4.43	0.91	5	1	3	13	17	65
self_efficacy_imp	1098	4.23	0.76	4	0	2	12	46	40
self_esteem1	1097	2.24	0.98	2	27	30	36	4	3
self_esteem2	1098	4.20	1.11	5	3	7	16	15	59
self_esteem5	1098	4.03	1.09	4	2	7	27	16	48
self_esteem4	1098	3.92	1.09	4	2	7	30	18	43
self_esteem_imp	1097	4.33	0.86	5	1	3	7	36	51
self_confidence1	1097	4.34	0.92	5	0	4	17	18	60
self_confidence2	1097	2.30	1.18	2	31	29	27	5	8
self_confidence3	1097	2.03	1.04	2	39	31	23	4	3
self_confidence4	1097	4.30	1.01	5	2	5	17	16	61
self_confidence_imp	1098	4.29	0.78	4	1	3	7	45	44
self_awareness1	1097	2.69	1.02	3	12	30	44	8	7
self_awareness2	1098	4.00	1.07	4	1	7	28	18	46
self_awareness3	1097	3.69	1.13	4	3	10	35	18	34
self_awareness4	1098	3.39	1.18	3	5	16	40	13	26
self_awareness_imp	1098	3.91	0.89	4	1	7	18	49	26
self_belief1	1097	4.45	0.90	5	1	4	13	14	68
self_belief2	1098	4.39	0.96	5	1	5	15	14	66
self_belief3	1098	3.95	1.08	4	2	6	31	15	45
self_belief4	1098	4.22	1.00	5	1	4	21	18	55
self_belief_imp	1098	4.43	0.70	5	0	2	6	40	53
gratification1	1098	3.86	1.12	4	2	9	33	14	42
gratification2	1097	2.93	1.19	3	11	25	38	10	15
gratification3	1096	3.38	1.38	3	11	17	25	15	32

gratification_imp	1098	4.07	0.98	4	2	7	12	41	39
impulses1	1098	2.26	1.06	2	27	37	26	6	5
impulses2	1098	2.91	1.28	3	16	22	36	8	18
impulses3	1097	3.89	1.14	4	3	8	30	15	44
impulses_imp	1098	4.39	0.70	4	0	1	5	43	49
attention1	1096	3.69	1.12	4	4	9	33	22	32
attention2	1098	2.67	1.23	3	18	31	31	7	13
attention3	1097	3.56	1.24	3	6	14	32	15	33
attention4	1098	2.66	1.10	3	16	26	42	8	8
attention_imp	1098	4.20	0.83	4	1	3	12	44	41
emotions1	1097	3.76	1.13	4	3	9	33	19	36
emotions2	1097	2.31	1.08	2	25	35	29	4	6
emotions3	1097	3.62	1.25	4	6	13	28	19	34
emotions_imp	1097	3.84	1.04	4	3	10	17	42	28
regulate_behaviors1	1094	3.62	1.29	4	9	10	26	20	35
regulate_behaviors2	1097	2.00	0.95	2	36	35	23	4	2
regulate_behaviors3	1097	3.93	1.18	4	4	9	24	18	46
regulate_imp	1097	3.83	1.04	4	3	10	18	40	29
thrill_seeking1	1097	1.32	0.74	1	79	13	5	2	1
thrill_seeking2	1097	2.58	1.14	3	19	30	36	7	9
thrill_seeking3	1097	2.00	0.98	2	38	34	23	3	3
thrill_seeking_imp	1097	4.13	1.09	4	6	4	7	38	45
problemsolving1	1097	3.54	1.22	3	6	12	35	15	31
problemsolving2	1097	3.71	1.17	4	4	10	32	17	36
problemsolving3	1097	3.89	1.15	4	3	8	28	17	43
problemsolving4	1096	3.83	1.15	4	3	8	30	18	41
problemsolving_imp	1097	4.32	0.76	4	1	2	7	45	45
critical_thinking1	1095	3.40	1.16	3	6	13	41	16	24
critical_thinking2	1097	3.45	1.22	3	6	16	35	15	29
critical_thinking3	1097	3.48	1.22	3	6	15	34	15	30
critical_thinking_imp	1096	3.80	0.93	4	1	8	24	43	23
decisions1	1095	3.92	1.13	4	3	8	29	16	44
decisions2	1097	3.57	1.22	3	6	13	33	16	32
decisions3	1097	3.64	1.20	4	5	11	33	17	34
decisions_imp	1097	4.43	0.71	5	0	1	6	39	53
social_skills1	1095	3.42	1.41	3	13	15	23	15	34
social_skills2	1095	3.78	1.18	4	5	8	29	20	38
social_skills3	1097	4.24	0.99	5	1	5	19	19	56
social_skills4	1097	3.86	1.15	4	2	11	27	17	42
social_skills5	1095	4.04	1.09	4	2	8	25	17	49
social_skills6	1097	3.76	1.18	4	5	8	33	16	38
social_skills7	1097	3.81	1.14	4	3	10	29	19	39
social_skills_imp	1097	4.24	0.76	4	0	3	9	48	40
communication1	1095	3.44	1.10	3	3	13	46	12	26
communication2	1096	3.94	1.11	4	2	7	28	17	44
communication3	1097	3.68	1.21	4	5	12	31	16	37
communication4	1096	2.55	1.15	3	20	30	36	5	10
communication5	1097	1.90	0.97	2	42	34	19	2	3

communication6	1096	3.27	1.10	3	4	18	46	12	20
communication_imp	1097	4.21	0.76	4	0	3	11	48	38

Table A. 19

Item descriptive information and percent of response by response category for Guatemala baseline data - Study 3

Item name	Item descriptive information				Percent of responses by response category			
	Obs	Mean	Std. Dev.	Median	1	2	3	4
self_efficacy2	811	2.32	0.69	2	10	50	37	3
self_efficacy3	811	3.35	0.58	3	1	3	57	39
self_efficacy4	811	3.38	0.61	3	1	5	51	44
self_efficacy5	811	3.55	0.53	4	0	0	44	56
self_esteem1	810	1.74	0.56	2	31	65	3	1
self_esteem2	811	3.52	0.59	4	1	2	41	56
self_esteem3	809	3.08	0.67	3	1	14	59	25
self_esteem4	811	3.31	0.54	3	0	3	62	34
self_esteem5	811	3.60	0.53	4	0	1	37	61
self_confidence1	811	3.48	0.56	4	0	2	46	51
self_confidence2	811	1.79	0.64	2	30	62	5	2
self_confidence3	811	1.85	0.56	2	24	69	7	1
self_confidence4	811	3.47	0.58	4	1	2	47	50
self_awareness1	811	2.50	0.71	2	6	45	42	7
self_awareness2	811	3.31	0.56	3	0	4	60	36
self_awareness3	810	3.08	0.69	3	1	15	57	26
self_belief1	808	3.45	0.57	3	0	2	49	48
self_belief2	811	3.51	0.53	4	0	1	46	52
self_belief3	806	3.35	0.60	3	1	5	54	40
self_belief4	811	3.51	0.54	4	0	1	46	52
impulses1	811	1.99	0.58	2	15	72	10	2
impulses2	811	3.16	0.61	3	1	9	64	27
attention1	811	2.49	0.70	2	6	45	43	6
attention2	811	2.34	0.70	2	9	52	34	4
emotions1	811	2.97	0.65	3	2	16	64	18
emotions2	811	2.12	0.69	2	15	61	20	4
emotions3	811	2.64	0.79	3	6	37	43	13
regulate_behaviors1	810	3.15	0.66	3	3	8	62	28
regulate_behaviors2	811	3.17	0.63	3	1	11	60	29
thrill_seeking1	810	1.89	0.62	2	24	65	10	2
thrill_seeking2	811	1.67	0.59	2	38	58	3	1
problemsolving1	810	3.01	0.66	3	2	13	64	20
problemsolving2	810	3.00	0.75	3	3	19	53	25
problemsolving3	809	3.16	0.59	3	1	7	66	25
problemsolving4	810	2.95	0.71	3	3	19	58	20
critical_thinking1	807	2.87	0.72	3	4	22	58	16

critical_thinking2	808	2.62	0.74	3	5	37	47	10
decisions1	810	3.09	0.64	3	2	11	63	24
decisions2	810	2.89	0.73	3	3	23	56	18
decisions3	811	3.12	0.65	3	2	8	64	25
social_skills1	810	3.13	0.63	3	2	9	64	25
social_skills2	811	2.94	0.75	3	3	21	54	22
social_skills3	811	3.18	0.69	3	1	13	52	34
social_skills4	811	3.13	0.70	3	1	15	53	31
social_skills5	811	2.91	0.75	3	3	24	52	21
communication1	811	3.10	0.59	3	0	12	66	22
communication2	811	2.82	0.69	3	2	27	56	14
communication3	811	2.52	0.77	2	7	45	38	10
communication4	810	1.79	0.58	2	29	63	7	0
communication5	811	2.75	0.77	3	4	32	47	16