Reliability Analysis of Social and Emotional Learning Measures

Amaniel P. Mrutu
University of Minnesota

Ernest C. Davenport, Jr.
University of Minnesota

Mark L. Davison
University of Minnesota

Mohammed A. A. Abulela
University of Minnesota
South Valley University

The current study examines the score reliability of survey measures of social and emotional learning constructs. In accordance with the Standards for Educational and Psychological Testing, this study collects validity evidence based on internal structure by conducting a confirmatory factor analysis to estimate reliability with coefficient hierarchical omega. This provides one estimate for the multidimensional construct of social and emotional learning and estimates for each subscale.

Keywords: social and emotional learning, survey, psychometric, reliability, alpha, omega

INTRODUCTION

The United States has historically relied on a deficit model to track child and adolescent development (Moore et al., 2004). This model asks: what are students doing wrong and how can we fix it? Programs that apply this deficit model such as Scared Straight (Petrosino et al., 2003) and DARE (Pan & Bai, 2009) have generally proved uninfluential in correcting negative behaviors. Positive youth development (PYD) offers a new paradigm for educational research by focusing on positive aspects of development in young people (Leman et al., 2017). This model which is rooted in positive psychology asks: what are students doing well and how can we intentionally encourage such positive behaviors? PYD is an approach that can be layered onto existing programs or curricula by emphasizing sociocultural aspects in the structuring of youth engagement activities (Moore, 2017; Smith et al., 2017).

PYD occurs in the context of Social and Emotional Learning (SEL). Taylor et al. (2017) provide meta-analytic evidence that SEL interventions positively influence "social-emotional skills, attitudes, and indicators of well-being" for students across demographics of race/ethnicity and socioeconomic status. SEL skills are positively associated with academic achievement from pre-kindergarten through grade 12 (Durlak

et al., 2011; Davis et al., 2014; Curby et al., 2013; Scales et al., 2006) and are particularly beneficial for disadvantaged youth because the development of social skills is deemed critical in young people demonstrating resilience (Oshri et al., 2017).

The transition from middle to high school is a particularly sensitive period for SEL skills. This is partially due to the neurocognitive development that youth undergo in the Dorsal Medial Prefrontal Cortex (DMPC) during adolescence (Sebastian et al., 2008). This biological change in the brain is mirrored by adolescents becoming more conscious of social comparisons with their peers. Scales et al. (2011) identify adolescent thriving, which is the opportunities and relationships that support youth in pursuit of their passions and empower them with the voice to make meaningful contributions to their surroundings, as a means for this age group to attain a healthy self-concept. SEL is one strategy for facilitating adolescent thriving in the school setting.

Measurement of Social and Emotional Learning

SEL is qualitative in nature describing positive social and emotional behaviors exemplified by youth. Researchers have created several instruments to convert the qualitative themes of SEL into measurable constructs such as the 5 C's (Lerner et al., 2005), CASEL model (Collaborative for Academic, Social, and Emotional Learning, 2012), and Developmental Assets Profile (DAP; Search Institute, 2013). The instrument in this study consists of items from the Minnesota Student Survey[1] (MSS), which reflect six constructs (i.e., Commitment to Learning [CtL], Positive Identity [PI], Social Competence [SC], Empowerment [EMP], Family/Community Support [FCS], Teacher/School Support [TSS]) derived from Search Institute's Developmental Assets Framework (DAF). The scales PI, SC, and EMP correspond directly to the DAP assessment (Search Institute, 2013), the CtL scale resembles a similar measure from the DAP, and FCS and TSS are based on "supports" concepts in the DAF.

From a confirmatory factor analysis of the MSS measures, Rodriguez (2017) concluded that the Developmental Skills three-factor model (i.e., CtL [based on the DAF], PI [a DAP measure], SC [a DAP measure]) and the Developmental Supports three-factor model (i.e., EMP [a DAP measure], FCS [based on the DAF], TSS [based on the DAF]) have a close to adequate fit based on full or partial results from 330,767 students who participated in the 2013 or 2016 MSS administrations.

A second round of factor analysis by Smith et al. (2019) involved a comparison of bifactor models on the 2016 MSS data which suggested that a structure with a general factor and six specific factors, instead of two (i.e., Developmental Skills, Developmental Supports), yield a better model fit. Rodriguez et al. (2019) examined validity evidence based on relations to other variables for this instrument and reported that SEL measures have positive associations with healthy behaviors (e.g., college goals, healthy diet) and negative associations with risky behaviors (e.g., skipped school, sent to office for discipline).

The Present Study

Although Rodriguez et al. (2019) evaluated the interpretations and proposed uses of the MSS measures, the reliability of the instrument was not examined in detail. The objectives of the current study are to ascertain: (1) To what extent are the scores from the MSS measures (based on the DAF) deemed reliable? (2) How does the reliability of the MSS scores compare to the original DAP scales? The current study contributes to the literature for supporting the interpretations and uses of the MSS measures through an examination of the reliability of scores based on the six-factor bifactor model and reliability of scale scores relative to Search Institute's (2013) initial field test with the original DAP model.

METHODS

Cronbach (1951) is the most cited publication for measuring reliability despite criticisms from Sijtsma (2009, 2015), McNeish (2018), and Green and Yang (2015) among other psychometricians who claim the assumptions underlying coefficient alpha (i.e., tau equivalence, uncorrelated error structure) are unreasonable. Davenport et al. (2015) explain that alpha is not strictly speaking a measure of internal consistency because alpha is a function of the average inter-item correlation and number of items. This

means that adding more items to a scale of measurement increases coefficient alpha, which is a parameter distinct from internal consistency (average inter-item correlation as defined by Cronbach [1951]). Due to the dependence on number of items, Schmitt (1996) notes that coefficient alpha does not have a specific threshold for what constitutes an acceptable estimate of reliability.

Raykov and Marcoulides (2019) responded to McNeish's (2018) claim about coefficient alpha being obsolete by noting that this phrase is too severe as there may be cases where the assumption of an uncorrelated error structure is met. However, this counterargument is conditional on the assumption of tau equivalence (i.e., all items on the scale contribute equally to the total score) which is seldom obtained in practice (e.g., Sijtsma, 2009, 2015; Schmitt 1996). Therefore, Raykov and Marcoulides (2019) state an exception case that, although plausible, is improbable leading one to ask if coefficient alpha is the most cited estimate of reliability due to psychometric properties or convention.

Coefficient omega, developed by McDonald (1999), is suggested as a preferable alternative to alpha because of its more robust properties such as being a congeneric measure (i.e., does not assume that all items contribute equally to the total score). This study conducts a confirmatory factor analysis in the R package lavaan (Rosseel, 2012) with full information maximum likelihood (FIML) estimation on the six-factor bifactor model proposed by Smith et al. (2019). Model fit indices for Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA) were examined with the following guidelines for acceptable fit: CFI > 0.90; TLI > 0.90; and RMSEA < 0.08 (Garver & Mentzer, 1999; Bentler, 1990). According to the *Standards for Educational and Psychological Testing*, "If the rationale for a test score interpretation for a given use depends on premises about the relationships among test items or among parts of the test, evidence concerning the internal structure of the test should be provided" (American Educational Research Association et al., 2014, p. 26-27). Rios and Wells (2014) suggest collecting validity evidence based on internal structure for bifactor models with coefficient hierarchical omega (ω_H) and omega subscales (ω_S) as defined in Equation 1 and Equation 2, respectively, where λ_{iG} denotes the general factor loading for item *i*; λ_{iS_n} denotes the specific factor loading for item *i* on the n^{th} subscale, and ${\theta_i}^2$ denotes the residual variance for item *i*.

$$\omega_H = \frac{(\sum \lambda_{iG})^2}{(\sum \lambda_{iG})^2+(\sum \lambda_{iS_1})^2+(\sum \lambda_{iS_2})^2+\ldots+(\sum \lambda_{iS_n})^2+\sum {\theta_i}^2} \quad (1)$$

$$\omega_S = \frac{(\sum \lambda_{iS_n})^2}{(\sum \lambda_{iG})^2+(\sum \lambda_{iS_n})^2+\sum {\theta_i}^2} \quad (2)$$

Additionally, the alpha() function with listwise deletion in the R package psych was used for two purposes: computing coefficient alpha and corrected item-total correlations. Next, the alpha.ci() function from the R package psych (Revelle, 2020) was used to find the 95% confidence interval for the estimate of coefficient alpha according to the procedure by Feldt et al. (1987). Although coefficient alpha is not an ideal estimate of score reliability, this was the reliability reported by Search Institute (2013) in their initial field test for the original DAP measures. Coefficient alpha is calculated in this study for the purpose of comparing the reliabilities from the MSS measures and the original DAP scales on a common metric. In the case that one or more scales have relatively low score reliabilities, the items on the MSS measure will be examined in greater detail. Since the scales are congeneric measures, the index, alpha if item deleted, is not considered appropriate for analyzing the impact of individual items on score reliability (Raykov, 2007). Instead, corrected item-total correlations are reported for understanding how each item contributes to internal consistency of the scale. This statistic is calculated by correlating person item scores to person scale scores without that item. If scale items have corrected item-total correlations greater than .30, the item meaningfully correlates to the aggregated score of remaining items (De Vaus, 2013).

Data

The current study is based on anonymous data from the 2016 MSS, which was administered to 126,868 students in grades 8, 9, and 11. Overall, approximately 65% of Minnesota students from the selected grade

levels participated in the MSS (Minnesota Department of Education, 2016). The data are presumed missing at random. This study used 37 four-point rating scale items (i.e., strongly agree, agree, disagree, strongly disagree) from the MSS to measure three scales adopted from and three scales adapted from Search Institute's DAP assessment, which are identified in this study as CtL, PI, SC, EMP, FCS, and TSS. CtL is a scale of six items regarding students' perceptions about their interest in school activities and student responsibilities (e.g., How often do you care about doing well in school?). PI is a scale of six items that address student self-concept (e.g., I feel in control of my life and future.). SC is a scale of eight items regarding how the student relates with their peers (e.g., I build friendships with other people.). EMP is a scale of six items describing the student's perceived level of safety and status in their surrounding environment including their school, home, and neighborhood (e.g., I feel safe at school.). FCS is a scale of five items referring to the extent of care that the student reports receiving from adults in their non-school environment (e.g., How much do you feel adults in your community care about you?). TSS is a scale of six items that gauge the level of perceived attentiveness shown by schoolteachers and staff to students (e.g., Overall, adults at my school treat students fairly.).

RESULTS

The bifactor model is reported in Table 1. Even with FIML, 1.65% of cases were removed from the factor analysis due to empty cells. Corresponding fit indices for the bifactor model were CFI of .864, TLI of .847, and RMSEA of .064. The RMSEA has acceptable fit (RMSEA < .08), CFI has close to acceptable fit (CFI > .90), and TLI has close to acceptable fit (TLI > .90) (Garver & Mentzer, 1999; Bentler, 1990). Regarding the close to acceptable fit of the CFI and TLI, Brown (2015) suggested to not depend solely on fit indices since they are only descriptives for ascertaining the lack of fit and can be impacted by several factors such as model misspecification. Furthermore, this model was retained for analysis because, based on the models examined by Smith et al. (2019), the six-factor bifactor model was the relatively best-fitting model for the dataset.

TABLE 1
MSS MEASURES OF SEL BIFACTOR MODEL

Item	λ_{SEL}	λ_{CtL}	λ_{PI}	λ_{SC}	λ_{EMP}	λ_{FCS}	λ_{TSS}	θ^2
Y18	.387	.866						.266
Y19	.325	.735						.236
Y20r	.168	.256						.518
Y21a	.187	.264						.258
Y21b	.328	.811						.309
Y21c	.413	1.000						.357
Y60a	.697		1.000					.276
Y60b	.748		.854					.344
Y60f	.765		.942					.244
Y60g	.635		.456					.512
Y60h	.758		.544					.365
Y60n	.535		-.195					.646
Y60c	.572			.737				.422
Y60d	.698			-.061				.429
Y60e	.756			.083				.411
Y60i	.707			.332				.348
Y60j	.629			1.000				.211
Y60k	.634			.427				.394

Item	λ_{SEL}	λ_{CtL}	λ_{PI}	λ_{SC}	λ_{EMP}	λ_{FCS}	λ_{TSS}	θ^2
Y60m	.341			.192				.430
Y60q	.530			.193				.570
Y22b	.376				.597			.268
Y22c	.312				1.000			.061
Y22d	.293				.635			.145
Y60l	.865				-.031			.318
Y60o	.789				-.017			.370
Y60p	.767				-.056			.308
Y8r	.430					.303		.498
Y59a	.558					.700		.292
Y59b	.688					1.000		.195
Y59c	.674					.410		.584
Y59e	1.000					.399		.863
Y21d	.377						1.000	.242
Y21e	.354						1.000	.176
Y21f	.340						.802	.313
Y21g	.354						.971	.149
Y21h	.455						.858	.296
Y59d	.902						.820	.618

Note. λ_{SEL} denotes the general factor loading for the social and emotional learning measures; λ_{CtL} denotes the specific factor loading for the Commitment to Learning scale; λ_{PI} denotes the specific factor loading for the Positive Identity scale; λ_{SC} denotes the specific factor loading for the Social Competence scale; λ_{EMP} denotes the specific factor loading for the Empowerment scale; λ_{FCS} denotes the specific factor loading for the Family/Community Support scale; λ_{TSS} denotes the specific factor loading for the Teacher/School Support scale; θ^2 denotes item residual variance.

Next, hierarchical omega is computed from the confirmatory factor analysis results and was .82, which is a reliability estimate based on the general factor scores. This is interpreted as 82% of the variance being explained by the general factor. The omega subscale reliabilities for the six SEL measures are reported in Table 2. After accounting for the proportion of variance attributed to the SEL general factor, relatively little variation (i.e., ranging from 1% to 7%) is accounted for by each of the omega subscales. Therefore, the remaining variation after partialling out the general factor does not make a significant contribution to scale reliability and is considered negligible.

Coefficient alpha for the 2016 MSS adopted and modified DAP scales and corresponding Search Institute field test scales are also reported in Table 2.

TABLE 2
DEVELOPMENTAL ASSETS FRAMEWORK RELIABILITY ESTIMATES

SEL Measure	# of Items	MSS Scale (Omega Subscale)	SI Scale (alpha)	MSS Scale (alpha)	MSS Scale (95% CI for alpha)
Commitment to Learning	6	.03	.83	.696	(.694, .699)
Positive Identity	6	.03	.79	.844	(.842, .845)
Social Competence	8	.02	.79	.845	(.844, .847)
Empowerment	6	.01	.74	.810	(.808, .812)
Family/Community Support	5	.02	.80*	.794	(.792, .795)
Teacher/School Support	6	.07	.80*	.859	(.857, .860)

Note. SI Scale denotes Search Institutes DAP model field test; MSS Scale denotes the modified DAP model based on items from the Minnesota Student Survey 2016 data; CI denotes confidence interval; * denotes that field test reported reliability estimate for an aggregated Support scale. [The last two * measures do not exist in the DAP.]

As shown in Table 2, the 95% confidence intervals for coefficient alpha on the MSS scales are narrow due to large sample sizes such that the confidence intervals do not enclose the coefficient alpha estimate from the corresponding SI scale. A distinction is observed regarding the proximity of each point estimate to their corresponding confidence interval. Some confidence intervals for the MSS scale are relatively closer to enclosing the corresponding coefficient alpha point estimate from the SI scale. Of all measures, the CtL scale presents the greatest concern. Recall that the MSS version of CtL reflects a construct of different items resembling the assets framework and does not employ the exact same items as the DAP. In Search Institute's (2013) field test, the alpha of the CtL scale equals .83 whereas the corresponding alpha from the MSS version equals .696. Due to this relatively greater discrepancy in reliability between the MSS and SI models, the corrected item-total correlations for the CtL scale are further examined to ascertain if any scale items contribute to the lower estimate of coefficient alpha. Table 3 presents corrected item-total correlations for the CtL scale.

TABLE 3
MSS COMMITMENT TO LEARNING CORRECTED ITEM-TOTAL CORRELATIONS

Item	Corrected Item-Total Correlation
1. How often do you care about doing well in school?	.67
2. How often do you pay attention in class?	.64
3. How often do you go to class unprepared? (reversed)	.23
4. If something interests me, I try to learn more about it.	.34
5. I think things I learn at school are useful.	.60
6. Being a student is one of the most important parts of who I am.	.64

Note. MSS denotes the modified DAP model based on items from the Minnesota Student Survey.

As reported in Table 3, item 3 is below the established threshold of .30, and item 4 is barely above this threshold (De Vaus, 2013). This implies that item 3 has the lowest discrimination in distinguishing between students with the highest and lowest ability levels of the CtL construct. The results are further discussed in the next section.

DISCUSSION

This study contributes to the analysis of score reliability for the MSS scores (based on the DAF) in that 82% of the variation in the MSS items is attributable to a general SEL factor. The use of reliability analyses on scores from the MSS suggested that omega subscales are negligible and, hence, not reported in the estimation of reliability. Interpretations of score reliability are derived solely through the general factor by computing hierarchical omega for the bifactor model. In the context of this instrument which is used for low stakes purposes, the .82 reliability estimate of hierarchical omega is deemed adequate.

Since the original SI field test for the DAP scales only reported coefficient alpha in analyzing reliability, alpha is computed for the MSS measures (including three DAP measures) to compare results across administrations. The reliability estimates were similar between the SI and MSS scales with the exception of the CtL scale. According to the corrected item-total correlations in Table 3, two items have relatively lower values on the CtL scale: item 3 and item 4. Item 3 (i.e., "How often do you go to class unprepared?" (reversed)) is the only reverse-coded item on the six reported scales. The negative wording may alter the response patterns of students. The item with the second lowest corrected item-total correlation is item 4 (i.e., "If something interests me, I try to learn more about it."). Although somewhat low, the authors suggest retaining item 4 on the CtL scale because the item sufficiently captures the underlying notion of this construct.

Results from this study suggest that the only reverse-coded item from the CtL scale has a relatively lower corrected item-total correlation and is slightly below the threshold of .30. There is a chance, however, that removing this item may adversely impact validity evidence based on test content resulting in construct underrepresentation (Sireci & Faulkner-Bond, 2014). Further examination of content-oriented validity evidence is necessary before removing item 3 from the CtL scale. To improve the reliability of this scale, the authors suggest requesting that the MSS reword the item "How often do you go to class unprepared?" to "How often do you go to class *prepared*?" This proposed change modifies the item to no longer be reverse coded which may improve the extent to which this item correlates with aggregated scores of the remaining items on the CtL scale. Future research on the MSS measures can re-administer the CtL scale with the revised item and compare estimates of reliability to those reported in the present study. More specifically, we are interested in checking if the corrected item-total correlations for each item exceed .30 (De Vaus, 2013) and if the coefficient alpha confidence interval is closer in proximity to enclosing coefficient alpha of .83 as reported by the original DAP model from Search Institute's (2013) initial field test.

ACKNOWLEDGEMENT

Thank you to Dr. Michael Rodriguez for providing feedback on the manuscript and access to the data.

ENDNOTE

1. The measures Positive Identity, Empowerment, and Social Competence come from the Search Institute Developmental Asset Profile (https://www.search-institute.org/surveys/choosing-a-survey/dap/) and were licensed to the Minnesota Student Survey for statewide use. The remaining scales were constructed from existing Minnesota Student Survey items, to be consistent with the asset domains from Search Institute (i.e., not Search Institute items).

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238.

Brown, T.A. (2015). *Confirmatory Factor Analysis for Applied Research*. Guilford Publications.

Collaborative for Academic, Social, and Emotional Learning. (2012). *2013 CASEL guide: Effective social and emotional learning programs—Preschool and elementary school edition*.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

Curby, T.W., Brock, L.L., & Hamre, B.K. (2013). Teachers' emotional support consistency predicts children's achievement gains and social skills. *Early Education & Development*, *24*(3), 292-309.

Davenport, E.C., Davison, M.L., Liou, P.Y., & Love, Q.U. (2015). Reliability, dimensionality, and internal consistency as defined by Cronbach: Distinct albeit related concepts. *Educational Measurement: Issues and Practice*, *34*(4), 4-9.

Davis, A., Solberg, V.S., de Baca, C., & Gore, T.H. (2014). Use of social emotional learning skills to predict future academic success and progress toward graduation. *Journal of Education for Students Placed at Risk (JESPAR)*, *19*(3-4), 169-182.

De Vaus, D. (2013). *Surveys in social research*. Routledge.

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

Feldt, L.S., Woodruff, D.J., & Salih, F.A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, *11*(1), 93-103.

Garver, M.S., & Mentzer, J.T. (1999). Logistics research methods: Employing structural equation modeling to test for construct validity. *Journal of Business Logistics*, *20*(1), 33.

Green, S.B., & Yang, Y. (2015). Evaluation of dimensionality in the assessment of internal consistency reliability: Coefficient alpha and omega coefficients. *Educational Measurement: Issues and Practice*, *34*(4), 14-20.

Leman, P.J., Smith, E.P., & Petersen, A.C. (2017). Introduction to the special section of Child Development on positive youth development in diverse and global contexts. *Child Development*, *88*(4), 1039-1044.

Lerner, R.M., Lerner, J.V., Almerigi, J.B., Theokas, C., Phelps, E., Gestsdottir, S., . . . von Eye, A. (2005). Positive youth development, participation in community youth development programs, and community contributions of fifth-grade adolescents: Findings from the first wave of the 4-H study of positive youth development. *The Journal of Early Adolescence*, *25*(1), 17-71.

McDonald, R.P. (1999). *Test Theory: A Unified Treatment*. Hillsdale, NJ: Erlbaum.

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412.

Minnesota Department of Education. (2016). *Data Reports and Analytics*. MN Department of Education. Retrieved June 28, 2020, from public.education.mn.gov/MDEAnalytics/DataTopic.jsp?TOPICID=2.

Moore, K.A. (2017). Commentary: Positive youth development goes mainstream. *Child Development*, *88*(4), 1175-1177.

Moore, K.A., Lippman, L., & Brown, B. (2004). Indicators of child well-being: The promise for positive youth development. *The Annals of the American Academy of Political and Social Science*, *591*(1), 125-145.

Oshri, A., Topple, T.A., & Carlson, M.W. (2017). Positive youth development and resilience: Growth patterns of social skills among youth investigated for maltreatment. *Child Development*, *88*(4), 1087-1099.

Pan, W., & Bai, H. (2009). A multivariate approach to a meta-analytic review of the effectiveness of the DARE program. *International Journal of Environmental Research and Public Health*, *6*(1), 267-277.

Petrosino, A., Turpin-Petrosino, C., & Buehler, J. (2003). Scared Straight and other juvenile awareness programs for preventing juvenile delinquency: A systematic review of the randomized experimental evidence. *The Annals of the American Academy of Political and Social Science*, *589*(1), 41-62.

Raykov, T. (2007). Reliability if deleted, not 'alpha if deleted': Evaluation of scale reliability following component deletion. *British Journal of Mathematical and Statistical Psychology*, *60*(2), 201-216.

Raykov, T., & Marcoulides, G.A. (2019). Thanks coefficient alpha, we still need you! *Educational and Psychological Measurement*, *79*(1), 200-210.

Revelle, W. (2020). Psych: Procedures for personality and psychological research [Computer software manual]. http://cran.r-project.org/web/packages/psych/. (R package version 2.0.12)

Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema*, *26*(1), 108-116.

Rodriguez, M.C. (2017). *2013-2016 MSS Technical Report on Measures of Developmental Skills, Supports, & Challenges*. Retrieved from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/195197.

Rodriguez, M.C., Dosedel, M., & Kang, Y. (2019). *Interpretation and Use Validation of Social and Emotional Learning Measures in Inequitable Settings*. Retrieved from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/202391.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Scales, P.C., Benson, P.L., & Roehlkepartain, E.C. (2011). Adolescent thriving: The role of sparks, relationships, and empowerment. *Journal of Youth and Adolescence*, *40*(3), 263-277.

Scales, P.C., Benson, P.L., Roehlkepartain, E.C., Sesma, A., Jr., & van Dulmen, M. (2006). The role of developmental assets in predicting academic achievement: A longitudinal study. *Journal of Adolescence*, *29*(5), 691-708.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*(4), 350.

Search Institute. (2013, May 15). *Developmental Assets Profile: Technical Summary*. Retrieved January 3, 2020, from https://www.search-institute.org/wp-content/uploads/2017/11/DAP-Psychometric-Information.pdf

Sebastian, C., Burnett, S., & Blakemore, S.J. (2008). Development of the self-concept during adolescence. *Trends in Cognitive Sciences*, *12*(11), 441-446.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107.

Sijtsma, K. (2015). Delimiting coefficient α from internal consistency and unidimensionality. *Educational Measurement: Issues and Practice*, *34*(4), 10-13.

Sireci, S., & Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema*, *26*(1), 100-107.

Smith, E.P., Witherspoon, D.P., & Wayne Osgood, D. (2017). Positive youth development among diverse racial–ethnic children: Quality afterschool contexts as developmental assets. *Child Development*, *88*(4), 1063-1078.

Smith, M., Kang, Y., Vue, K., Miranda, A., & Rodriguez, M.C. (2019). *Investigating Internal Structure of Social Emotional Learning Measures: A Bifactor Approach*. Retrieved from the University of Minnesota Digital Conservancy. Retrieved from http://hdl.handle.net/11299/202394

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