CONSTRUCT VALIDATION OF THE SOCIAL-EMOTIONAL CHARACTER DEVELOPMENT SCALE IN BELIZE: MEASUREMENT INVARIANCE THROUGH EXPLORATORY STRUCTURAL MODELING

Krystal M. Hinerman, B.A., M.A.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2014

APPROVED:

Darrell M. Hull, Major Professor
Robin K. Henson, Committee Member
Rebecca J. Glover, Committee Member
Endia Lindo, Committee Member
Abbas Tashakkori, Chair of the Department of Educational Psychology
Jerry R. Thomas, Dean of the College of Education
Mark Wardell, Dean of the Toulouse Graduate School

Hinerman, Krystal M. Construct Validation of the Social-Emotional Character Development Scale in Belize: Measurement Invariance Through Exploratory Structural Equation Modeling. Doctor of Philosophy (Educational Research – Research, Measurement and Statistics), August 2014, 90 pp., 10 tables, chapter references.

Social-emotional learning (SEL) measures assessing social-emotional learning and character development across a broad array of constructs have been developed but lack construct validity. Determining the efficacy of educational interventions requires structurally valid measures which are generalizable across settings, gender, and time. Utilizing recent factor analytic methods, the present study extends validity literature for SEL measures by investigating the structural validity and generalizability of the Social-Emotional and Character Development Scale (SECDS) with a large sample of children from schools in Belize (n = 1877, ages 8 to13). The SECDS exhibited structural and generalizability evidence of construct validity when examined under exploratory structural equation modeling (ESEM). While a higher order confirmatory factor structure with six secondary factors provided acceptable fit, the ESEM six-factor structure provided both substantive and methodological advantages. The ESEM structural model situates the SECDS into the larger body of SEL literature while also exhibiting generalizability evidence over both gender and time.

ACKNOWLEDGEMENTS

First, I would like to thank my committee chair, Dr. Darrell Hull. The opportunities, support, and mentorship you have provided throughout my journey to become a research professional are invaluable. You have continually promoted me and had faith in my ability even when I did not. To Dr. Robin Henson, your expertise and guidance throughout my doctoral studies have provided constant enlightenment and encouragement. To my committee members, Dr. Becky Glover and Dr. Endia Lindo, thank you for providing additional guidance throughout the dissertation process and for sharing your valuable substantive content knowledge. To my advisor, Dr. Qi Chen, thank you for assisting in my initiation into the UNT department and for introducing me early on to higher level modeling. To Dr. Abbas Tashakkori, thank you for always having an open door and providing both teaching and funding opportunities throughout my doctoral studies. I would be remiss not to send a special thank you to Laura Musgrove and Alecia Adams who have continued to look out for me from my first days at UNT. Above all, I would like to express a special thanks to my husband, Kelvin, and my two sons, Bishop and Wesley. I could not have completed this endeavor without your continual love and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

Tables

CONSTRUCT VALIDATION OF THE SOCIAL-EMOTIONAL CHARACTER DEVELOPMENT SCALE IN BELIZE

Introduction

Social-emotional learning developed in response to school programs designed to target specific problem youth behaviors such as violence and substance abuse (CASEL, 2002). Instead of focusing on the resulting problem behavior, social emotional learning (SEL) provides a preventative framework for addressing underlying causes of negative youth behaviors while also supporting academic improvement (Damon, 2004; Greenberg et al., 2003; Weissberg & O'Brien, 2004). Although several frameworks exist in the literature, SEL generally addresses a set of five inter-related cognitive, affective, and behavioral competencies: self-awareness, social awareness, responsible decision making, self-management, and relationship management (Zins et al., 2004; Weissberg & O'Brien, 2004; CASEL, 2011). Table 1 includes core descriptors of the five SEL competencies.

Character development programs extend the five SEL compentencies to include direct teaching of core values such as respect, responsibility, honesty, fairness, compassion, courtesy, and courage (Park, 2004). As a result, character development focuses on developing good character, defined as principles valued by society to reflect decisions beneficial to the person but also to others and society as a whole (Park & Peterson, 2008). Character education in the SEL framework is associated with positive behaviors under a wide variety of social-emotional learning competencies such as prosocial behavior, school attachment, responsibility, respect, self-efficacy, self-control, social skills, and academic performance (Berkowitz & Bier, 2004).

Table 1

Social Emotional Learning Competencies

Self-Awareness
Identifying and recognizing emotions and thoughts
Accurate self-perception
Recognizing strengths, limitations and values
Self-efficacy
Well-grounded sense of confidence and optimism
Spirituality
Social Awareness
Perspective taking
Empathy
Appreciating diversity
Respect for others
Understanding social and ethical norms for behavior
Recognizing resources and supports
Responsible Decision Making
Problem identification and situational analysis
Making constructive and respectful choices
Problem solving
Evaluation and reflection
Personal, moral and ethical responsibility
Self-Management
Regulating emotions, thoughts and behaviors
Impulse control and stress management
Self-motivation and discipline
Goal setting and organizational skills
Relationship Management
Communication, social engagement and building relationships
Establishing and maintaining relationships with diverse individuals
Resisting inappropriate social pressure
Working cooperatively
Negotiation, refusal and conflict management
Help seeking and providing

Adapted from Zins, J.E., Weissberg, R.P., Wang, M.C., & Walberg, H.J. (2004, p. 7) and CASEL (2012, What is SEL?).

Measuring Social-Emotional Learning and Character Development:
Social Emotional and Character Development Scale

Although decades of empirical research surrounding the effects of social-emotional learning and character development have been published, issues regarding instruments to measure SECD skills remain unresolved. In a report issued by the Society for Prevention Research intended to standardize the criteria for identifying prevention programs which have been sufficiently empirically tested, a standard was set to include measures which were psychometrically sound, meaning the measures have been demonstrated to exhibit construct validity and reliability (Flay et al., 2005). Greenberg's (2004) suggestions for future research in prevention science called for the development of easily utilized, valid and reliable assessments of social, emotional, ethical and health outcomes. More specifically, Greenberg highlighted the need to develop meaningful and easily understood assessments of social and emotional competence. The meta-analysis by Durlak et al. (2011) concluded 24% of the examined empirical studies on SEL programs did not use reliable outcome measures and 50% did not use valid outcome measures. Likewise, Wigelsworth et al. (2010) called for examination of the psychometric properties and application of SEL measures across varying populations and ethnicities. In a systematic review of 187 currently used SEL instruments, Humphrey et al. (2011) concluded the majority of measures have been developed only with American populations and there is little analysis of the applicability of the measures across different groups (e.g. ethnicity, gender).

Ji, DuBois, and Flay (2013) developed and conducted initial validation of a social-emotional and character development scale under the SEL and character development framework. Meant to address the need for a multi-dimensional SEL instrument which captures both social and emotional skills, the Social Emotional and Character Development Scale

(SECDS) includes 29 four point Likert scale items designed to assess skills and behaviors with likely relevance to both social-emotional learning and character development programs.

Table 2

Comparison of Student and Teacher Versions of the SECDS

Construct	Item	Teacher*
Prosocial Behavior	I play nicely with others.	X
	I do things that are good for the group.	X
	I treat my friends the way I like to be treated.	
	I am nice to kids who are different from me.	X
	I try to cheer up other kids if they are feeling sad.	
	I am a good friend to others.	X
	I think about how others feel.	X
Honesty	I apologize when I have done something wrong.	X
	I tell the truth when I have done something wrong.	X
	I tell others the truth.	
	I keep promises I make to others.	
	I admit my mistakes.	X
Self-Development	I make myself a better person.	
	I keep trying at something until I succeed.	X
	I set goals for myself (make plans for the future).	X
	I try to be my best.	X
Self-Control	I wait my turn in line patiently.	
	I keep my temper when I have an argument with other kids.	X
	I follow the rules even when nobody is watching.	X
	I ignore other children when they tease me or call me bad names.	X
Respect at School	I speak politely to my teacher.	X
	I obey my teacher.	X
	I follow the directions of my teacher.	X
	I listen (without interrupting) to my teacher.	X
	I follow school rules.	X
Respect at Home	I speak politely to my parents.	
	I obey my parents.	
	I listen (without interrupting) to my parents.	
	I follow the rules at home.	

*X indicates corresponding item on the Teacher Rating of Student assessment.

The six SECDS constructs were intended to capture school-related aspects of the five larger social emotional learning constructs. Ji, Dubois, and Flay (2013) utilizing data from 459 Chicago students Grades 3 to 5 over five waves of data collection, demonstrated concurrent validity with

several related outcome measures in addition to high test-retest and internal reliability across gender and ethnic groups. Table 2 includes the SECDS items and associated constructs.

Methodological Overview

Few studies have used factor analysis to investigate the construct validity of instruments designed to measure social-emotional and character development skills (Humphrey et al., 2011). In order to establish construct validity in SECD measurement models, methodologies such as exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and exploratory structural equation modeling (ESEM) must be applied to current SECD measures. With few studies reporting evidence for factorial invariance across groups or time, it is important to continue the factorial investigations to include multi-group analysis.

Exploratory Structural Equation Modeling

An integration of EFA, CFA and SEM, exploratory structural equation modeling (ESEM) was developed to help aleviate commonly encountered CFA problems associated with goodness of fit, differentiation of factors, measurement invariance across time or groups and differential item functioning (Asparouhov & Muthèn, 2009; Marsh et al., 2009, 2010). As such, instead of associating each item with only one factor and constraining all other non-target loadings to zero as is typical in the highly restrictive independent clusters model (ICM), ESEM allows for less restrictive models in which all factor loadings are estimated and where items are free to cross-load on other factors within the same set of factors (Marsh et al., 2011; Asparouhov & Muthèn, 2009). Instead of calculating structure coefficients in a separate analysis as authors such as Thompson (1997) demonstrate, ESEM includes the structure coefficient parameter estimation

along with the standard errors for the structure coefficients. ESEM retains the capability of rotating factors and also comparing model fit through comparing model fit statistics.

ESEM Model Fit

In lieu of depending on arbitrary cutoff points to fit indices, ESEM usually involves testing a hypothesized model along with other alternative models. In a CFA independent clusters model (ICM-CFA), each item is regressed on only one factor and all other factor-to-item loadings are assumed to be zero (Marsh, Ludtke, Nagengast, & Morin, 2013). The ICM-CFA can be considered nested within the comparable ESEM model. Since model fitting depends on correctly specifying the model, the fit of hypothesized and alternate models are compared to see which better reproduces the variance in the assessed data (Byrne, 1998; Marsh et al., 2004). When comparing the fit of nested models imposing differing numbers of invariance constraints with an adequate sample size ($N > 300$), Chen (2007) suggests less than 0.01 decrease in incremental model fit indices (e.g., CFI) and a RMSEA increase of less than 0.015 supports retaining the more parsimonious model. However, Chen cautions against over-generalization of these suggestions since sample size and model complexity can affect the magnitude of changes in fit statistics. For the purposes of testing invariance models where the indicators are categorical, the Satorra-Bentler scale chi-square difference testing function (DIFFTEST in MPlus; Muthèn & Muthèn, 2010). The MPlus DIFFTEST analysis involves first running the least restrictive model (H_1), saving the derivitives of the model, then subsequently using the saved derivities to compare the fit of the more restrictive model (H_0). A statistically significant DIFFTEST result indicates the more parsimonious (more restrictive) model to be a worse fit for the data (H_0 is rejected).

Multi-Group Analysis

Multi-group factorial invariance and time invariance can be examined under the ESEM framework (Marsh et al., 2009, 2010; Morin, Marsh, & Nagengast, 2013; Guay, Morin, Litalien, Valois, Vallerand, 2014). The purpose of multi-group analysis is to determine if item responses from identified groups follow similar response patterns as related to the underlying theoretical structure of the instrument. Time invariance analysis determines the similarity of response patterns across different waves of data collection. Testing factorial and time invariance follows a sequential constraint imposition procedure comparing a set of partially nested models ranging from the least restrictive model with no parameters constrained to be invariant, to a model with complete factorial invariance where all parameters are constrained to be invariant (Marsh et al., 2011; Dimitrov, 2010). This forward approach to testing factorial invariance provides for examing configural, measurement and structural invariance. Table 3 provides the taxonomy of the multiple-group exploratory structural equation models (MGESEM). Again, the Satorra-Bentler scaled chi-square difference ($\Delta\chi^2$) and change in CFI (ΔCFI < -.01) can be used to compare models (Dimitrov, 2010).

Rationale for the Present Study

The Social Emotional and Character Development scale (SECDS) is a recently developed scale (Ji, DuBois, & Flay, 2013) in the initial stages of psychometric evaluation. Previous investigation of the SECDS psychometric properties involved a sample of U.S. students (Ji et al., 2013), and SEL measures are badly needed to assess the efficacy of interventions in developing countries, especially in the Caribbean which includes many English speaking countries such as Belize. The present study extends the validity literature for SEL measures by investigating the

structural validity and generalizability of the Social-Emotional and Character Development scale using both traditional and more recently utilized factor analytic tools (ESEM), all with a large sample from Belize. The research questions posed are as follows:

1. To what extent does the structure of the SECD scale, as demonstrated in a sample of Belizean students, replicate results published by Ji et al. (2013)?
2. Can the factor structure of the SECDS be better represented by extending the factor structure evidence to include exploratory structural equation models?
3. Does the SECDS scale exhibit factorial invariance across gender and time?

Method

Initial psychometric investigation of the SECDS demonstrated structural validity in a longitudinal sample of U.S. youth (Ji et al., 2013). In an effort to provide cross-cultural validity evidence for the SECDS, the present study utilized data from 24 primary schools in the Belize District, Belize. Situated in Central America and bordered by Mexico, Guatemala, and the Caribbean Sea, Belize has 8,800 square miles of land and a population of 334,060 (United Nations, 2013). With a GDP-per capita of $8,900 (2012 U.S. dollars), Belize has the second highest per capita income in Central America; however, 4 out of 10 people live in poverty (Mundi, 2013).

Sample

Data was collected from schools in the Belize District. In 2011, the Belize District contained 68 primary schools, ranging in enrollment from n = 12 to 1056 students per school (*Mdn* = 207) inclusive of eight grades, which in Belize are referred to as Infant 1 and 2 (generally aged 5 and 6 years) and 1st through 6th Standard (comprising ages 7 – 13 years). Data for the present study was collected from a sample of n = 24 schools which were randomly

selected from the Belize District. Only students in Standards 4 through 6 were administered the SECDS. A total of 1877 students provided SECD scale data for at least one of two waves of measurement. Of the represented upper elementary students, 36% were Standard 4, 33% were Standard 5, and 31% were Standard 6. The demographics of the students with completed demographic information (n =1781) were as follows: 51% male, 49% female; Creole 55%, Metizo 25%, Garifuna 6%, Maya 2%, and 6% other ethnicity.

Procedure

Students were administered the SECDS at the beginning of the school year in the Fall of 2011, and again at the end of the school year in July of 2011. An attempt was made to collect data from all students attending the schools at Time Two regardless of inclusion or omission at Time One. School staff administered the SECDS as part of several self-report measures to all students in their classrooms. In order to minimize socially desirable responses, schools were instructed to have school personnel other than the classroom teacher administer the SEL measures and students were informed their responses would be kept confidential. At time one, separate assessments were administered for the demographic questionairre and a 79-item positive youth development survey (including the SECDS). Teachers were provided with administration instructions complete with instructions on directing students to complete a 10-digit identification number. At Time 2 students were provided with a assessment booklet which included the previously administered positive youth development battery. If students had participated in Time 1 administration, a printed identification label was included on their booklet. Testing administrators were again provided with an administration manual. At both time one and Time 2 test administrators were instructed to read the instructions as well as each item to the students in

order to ensure reading ability did not prevent appropriate completion of the positive youth development battery.

Data Analysis

Preliminary Analysis. Following procedures outlined by Osborne (2013), data was cleaned for outliers. Missing data was evaluated across demographic variables in order to determine the missing data mechanism (Enders, 2010). For the purposes of multiple imputation, data was considered missing at random (MAR) and 20 item-level imputed datasets were generated at the time of each SEM analysis using MPlus Version 6.12 (Muthèn & Muthèn, 2010). For the purposes of comparing models where the chi-square DIFFTEST function (which does not allow for multiple imputation) was utilized, data was considered MAR and models were estimated using a four step estimation method which utilizes maximum likelihood estimation for the first two steps (Muthèn & Muthèn, 2010).

Phase I: Generalizability of Structural Validity

Since the purpose of the present study, in part, is to replicate the results of a previous psychometric investigation of the SECDS using a different sample of students from a different cultural context, Phase I of the data analysis followed the structural analysis as set forth in Ji et al., (2013). Replication of previous psychometric analysis addresses both structural evidence and generalizability evidence of the SECDS (Messick, 1995; Dimitrov, 2010). Confirmatory factor analysis (CFA) was used to evaluate the degree to which the SECDS responses were consistent with the theorized multidimensional, hierarchical conceptualization of social-emotional skills and character (Brown, 2006). In order to initially test this conceptualization, the

hypothesized higher order model and three comparative models was fit to the data. The hypothesized model mirrors the higher-order factor conceptualization of self-concept (Shavelson et al., 1976). In this model, all 28 items were assigned to their respective SECDS dimension, and all of the dimensions or sub-factors were nested within a higher-order SECD factor. The first order factors were not correlated. The first alternative model includes all 28 indicators assigned to a single SECD factor. The second alternative model associated all 28 items with the respective dimensions; however, in lieu of a higher order factor, all factors were specified to correlate. The third alternative model included all items as indicators for a single first order factor. Appendix D contains path models of the apriori CFA comparison models.

MPlus Version 6.12 was used to conduct all competing CFA models. Since responses to the SECDS included ordered categorical data from a 4-point Likert scale, CFAs employed weighted least squares estimation using a diagonal weight matrix with standard errors and mean and variance adjusted chi-square test statistic using a full weight matrix (WLSMV; Muthèn & Muthèn, 2010). Model fit was evaluated using indices which are adjusted for sample-size: root mean square error of approximation (RMSEA), comparative fit index (CFI), the Tucker-Lewis index (TLI) . Criteria for the various indices using categorical data were followed as recommended by Schreiber, Stage, King, Nora, and Barlow (2006): RMSEA .06-.08, CFI .90-.95, and TLI .90- .96. When comparing the fit of nested models, suggestions by Chen (2007) will be followed where a less than 0.01 decrease in incremental model fit indices (e.g. $\Delta CFI < -0.01$) and a RMSEA increase of less than 0.015 supports retaining the more parsimonious model ($\Delta RMSEA < 0.015$). In addition, the Satorra-Bentler scaled chi-square difference (DIFFTEST in MPlus) will be used to compare the fit of the hypothesized model to alternative models (Dimitrov, 2010; Muthèn & Muthèn, 2010).

Phase II: Structural Validity

Extending the Ji et al. (2013) psychometric investigation beyond commonly accepted CFA models where items are required to load on only one factor, Phase II of the present study examined the factor structure of the SECDS using exploratory structural equation modeling (ESEM). Appendix D includes the path diagram for the ESEM model. Following recommendations by Marsh et al. (2009) and because previous evaluation of the SECDS scale indicated some of the SECDS factors were correlated at .7 or more, the CFA factor structure was examined under an oblique geomin rotation with an epsilon value of .5 as well as an oblique target rotation where all non-target loadings were set to be influenced towards zero. In order to remain consistent with Phase I of the analysis, the RMSEA, CFI, TLI, and DIFFTEST in addition to examination of the parameter estimates were used to assess model fit (Muthèn & Muthèn, 2010; Dimitrov, 2010; Asparouhov & Muthen, 2009).

Phase III: Generalizability Across Groups and Time

Utilizing the final measurement model retained from Phase I and II, the multi-group factorial invariance and time invariance was assessed using SEM procedures outlined by Guay, Morin, Litalien, Valois, & Vallerand (2014), Byrne (2012), Dimitrov (2010), and Marsh et al. (2009, 2010). Prior research has reported gender differences in SEL intervention effects (e.g. Endrulat, Tom, Ravitch, Wesley, & Merrell, 2010; Taylor, Liang, Tracy, Williams, & Seigle, 2002; Wilson & Lipsey, 2007). In order to assess the SECDS factor structure and determine if comparing latent means across gender is appropriate, a set of multigroup models were compared.

Testing factorial invariance followed a sequential constraint imposition procedure comparing a set of partially nested models ranging from the least restrictive model with no

parameters constrained to be invariant to a model with complete factorial invariance with all parameters constrained to be invariant (Guay et al., 2014; Marsh et al., 2011; Dimitrov, 2010; Byrne, 2012). This forward approach to testing factorial invariance provides for examing configural, measurement and structural invariance. Table 3 provides the taxonomy of the multiple-group confirmatory factor analysis (MGCFA) models included in the factorial invariance analyses.

Table 3

Taxonomy of ESEM Factorial Invariance Models Using Categorical Indicators

Model	Parameters Constrained to be Invariant					Invariance Level
	Factor Elements			Indicator Elements		
	Loadings	Variance-Covariance	Means	Uniqueness	Thresholds	
1						Configural
2	X					Weak Factorial
3	X				X	Strong Factorial
4	X			X	X	Strict Factorial
5	X	X		X	X	Variance-Covariance
6	X	X	X	X	X	Latent Means/Complete

Adapted from Marsh et al., 2011 and Guay et al. (2014).

Similar to testing invariance across groups, the six invariance models can be adapted to evaluate test - re-test instrument performance (Marsh et al. 2011). One adaptation is the inclusion of correlated uniqueness (CU) for the same indicator between Time 1 and Time 2. Failure to include the correlated uniqueness between the same items in two different testing periods is likely to inflate test-retest correlations (Marsh et al. 2004, 2011); therefore in addition to the nested time invariance models, a comparison between models estimating CU and not estimating CU was conducted . The DIFFTEST, CFI (ΔCFI < -.01) and RMSEA (ΔRMSEA = .015) was used to compare all invariance models (Chen, 2007; Dimitrov, 2010).

Results

Phase I: Generalizability of Structural Validity

For the purposes of replicating construct validity procedures as demonstrated by Ji et al. (2013), CFAs comparing the hypothesized higher order model and three comparative models were fit to the first wave of data. Table 4 presents the model fit indices for the four compared models. While the hypothesized higher order factor model provides reasonably good fit, comparisons of model fit indicates Alternative 2: six-correlated factor model (ΔCFI = 0.008, ΔRMSEA = -0.008) to be a slightly better fit. The DIFFTEST comparing the hypothesized Higher Order CFA nested within the alternative 6 Correlated Factor CFA suggests the addition of a higher order factor provided decrement in model fit (H_0: Higher Order v. H_1: 6 Correlated Factors; MD$\Delta\chi 2$ = 180.862, df = 9, p < .001).

Table 5 includes the factor loadings, structure coefficients, and factor correlations for the six-correlated factors model. The target factor loadings for all factors are substantial (.511 - .745). However, the structure coefficients for all non-target loadings are above .3, indicating the factors are not distinct as is required for the independent cluster model CFA (ICM-CFA) where all non-target cross loadings are predetermined to be zero. As would be expected, the factor correlations are also high (.629 - .909) indicating the factors are highly related even though the higher-order factor model does not provide a substantially better fit

.

Table 4

Model Fit Comparing Hypothesized CFA and Three Alternatives

Model	χ^2	df	CFI	TLI	RMSEA	RMSEA CI	MD$\Delta\chi^2$	$df_{\Delta\chi 2}$	$p_{\Delta\chi 2}$	ΔCFI	ΔTLI	ΔRMSEA
Hypoth: Higher Order	2009.178	371	0.943	0.937	0.049	[.047, .051]	180.862[a]	9	<.001	-0.008	-0.008	0.003
Alt 1: Correlated Factors	1772.769	362	0.951	0.945	0.046	[.044, .048]	985.876	-	-	-	-	-
Alt 2: Uncorrelated Factors	23418.85	377	0.192	0.130	0.181	[.179, .183]	5392.856[b]	15	<.001	-0.759	-0.815	0.135
Alt 3: Single Factor	3162.856	377	0.902	0.895	0.063	[.069, .065]	791.051[c]	6	<.001	-0.041	-0.042	0.014

a. H0: Higher Order v. H1: 6 Correlated Factors; b. H0: 6 Uncorrelated Factors v. H1: 6 Correlated Factors; c. H0: Single Factor v. H1: Higher Order.

Note. All models estimated using WLSMV. Missing values <5% on all indicators.

Table 5

Pattern Coefficients for CFA and ESEM Models

	CFA and Structure Coefficients*						ESEM (geomin, ε = .5)**						Target (all non-CFA indicators ~0)**						
I#	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	Item
51	**0.59**	0.51	0.50	0.39	0.52	0.38	**0.32**	*0.20*	0.13	*0.05*	*0.19*	*-0.10*	0.24	0.16	0.26	*0.05*	0.19	-0.13	I wait my turn in line patiently.
56	**0.52**	0.45	0.44	0.35	0.46	0.34	**0.47**	0.09	*0.04*	0.07	0.09	-0.02	**0.39**	0.13	0.16	*0.07*	0.11	*-0.04*	I keep my temper when I have an argument with other kids.
63	**0.72**	0.63	0.61	0.48	0.64	0.47	0.14	0.24	0.24	0.02	0.27	*0.06*	*0.04*	0.24	0.29	*0.02*	0.24	*0.02*	I follow the rules even when nobody is watching.
73	**0.51**	0.44	0.43	0.34	0.45	0.33	**0.44**	0.09	0.10	0.03	0.05	0.00	**0.36**	0.20	0.16	*0.03*	*0.05*	*-0.02*	I ignore children when they tease me or call me bad names.
55	0.53	**0.62**	0.47	0.53	0.53	0.53	**0.32**	0.13	0.18	*0.06*	*0.22*	*-0.06*	0.23	0.23	0.17	*0.06*	0.21	-0.09	I play nicely with others.
60	0.57	**0.65**	0.50	0.43	0.59	0.52	0.07	0.14	*0.21*	0.04	*0.26*	0.19	*-0.02*	0.23	0.16	*0.04*	0.24	0.16	I do things that are good for the group.
62	0.50	**0.57**	0.44	0.38	0.52	0.45	0.15	*0.01*	*0.22*	*0.13*	*0.19*	*0.11*	0.08	0.25	*0.02*	0.13	0.16	0.07	I treat my friends the way I like to be treated.
65	0.56	**0.65**	0.49	0.43	0.59	0.51	*0.24*	0.01	***0.33***	0.05	0.10	*0.19*	0.14	**0.42**	*0.03*	*0.05*	*0.06*	0.14	I am nice to kids who are different from me.
67	0.49	**0.56**	0.43	0.37	0.51	0.44	*0.20*	*-0.16*	0.24	*-0.03*	*0.24*	*0.27*	0.10	**0.39**	-0.16	*-0.03*	0.23	0.22	I try to cheer up other kids if they are feeling sad.
68	0.51	**0.59**	0.45	0.39	0.54	0.47	0.23	-0.13	0.20	*-0.01*	*0.17*	***0.37***	0.14	**0.37**	-0.12	*-0.01*	0.16	**0.33**	I am a good friend to others.
72	0.55	**0.64**	0.48	0.42	0.58	0.50	0.14	*0.11*	***0.40***	0.02	0.04	0.20	*0.04*	**0.44**	0.14	*0.02*	*-0.03*	0.15	I think about how others feel.
53	0.50	0.45	**0.59**	0.44	0.43	0.41	0.30	***0.40***	*-0.27*	**0.32**	*-0.02*	0.15	0.29	**-0.31**	**0.46**	**0.32**	*0.02*	0.21	I speak politely to my teacher.
54	0.60	0.54	**0.71**	0.53	0.51	0.49	*-0.03*	***0.67***	0.01	0.20	0.06	0.07	*-0.06*	-0.17	**0.74**	0.20	*0.02*	0.09	I obey my teacher.
66	0.61	0.55	**0.72**	0.54	0.52	0.50	0.09	**0.55**	*0.10*	0.02	0.09	0.14	*0.01*	*0.03*	**0.63**	*0.02*	*0.05*	0.15	I follow the directions of my teacher.
74	0.55	0.49	**0.65**	0.49	0.47	0.45	0.18	***0.39***	0.18	0.10	-0.03	0.06	0.12	0.14	**0.46**	0.10	-0.08	*0.05*	I listen (without interrupting) to my teacher.
75	0.63	0.57	**0.75**	0.56	0.54	0.52	0.18	***0.42***	*0.24*	-0.05	0.18	0.02	*0.07*	0.22	**0.50**	*-0.05*	0.14	*-0.01*	I follow school rules.
58	0.40	0.40	0.45	**0.60**	0.41	0.38	0.06	-0.08	*-0.07*	***0.88***	0.00	*0.04*	0.16	-0.26	-0.16	**0.88**	*0.01*	*0.07*	I speak politely to my parents.
59	0.44	0.44	0.50	**0.66**	0.46	0.42	-0.29	0.23	*0.28*	**0.51**	0.10	-0.01	-0.26	*0.00*	0.17	**0.51**	*0.02*	*-0.02*	I obey my parents.
70	0.44	0.44	0.49	**0.66**	0.46	0.42	0.05	0.14	***0.31***	**0.35**	0.00	*-0.03*	*0.04*	0.18	0.13	**0.35**	*-0.07*	*-0.06*	I listen (without interrupting) to my parents.
71	0.48	0.48	0.54	**0.72**	0.50	0.45	-0.08	*0.07*	***0.41***	0.29	0.18	*0.00*	-0.11	0.28	*0.03*	0.29	0.11	*-0.05*	I follow the rules at home.
52	0.55	0.57	0.45	0.43	**0.62**	0.41	**0.41**	0.05	0.03	*0.16*	*0.19*	*0.03*	**0.34**	0.11	0.10	0.16	0.21	*0.01*	I apologize when I have done something wrong.
57	0.52	0.53	0.42	0.40	**0.58**	0.38	*0.06*	*0.08*	-0.12	0.14	***0.62***	-0.01	*0.02*	-0.13	0.07	0.14	**0.68**	*-0.03*	I tell the truth when I have done something wrong.
61	0.55	0.57	0.45	0.43	**0.63**	0.41	-0.03	*-0.01*	*0.09*	*0.02*	**0.71**	0.00	-0.11	0.10	*-0.03*	*0.02*	**0.75**	*-0.05*	I tell others the truth.
64	0.46	0.48	0.38	0.36	**0.53**	0.34	0.11	-0.06	*0.28*	*0.14*	*0.13*	0.12	*0.06*	**0.31**	-0.07	0.14	0.09	0.08	I keep promises I make to others.
69	0.52	0.53	0.42	0.41	**0.59**	0.38	0.22	-0.03	*0.03*	*0.09*	**0.32**	0.18	0.15	0.11	*-0.02*	0.09	**0.35**	0.15	I admit my mistakes.
76	0.48	0.59	0.51	0.47	0.49	**0.75**	0.05	0.11	*0.25*	0.07	*0.07*	**0.39**	*-0.02*	0.28	0.12	0.07	*0.02*	**0.37**	I make myself a better person.
77	0.42	0.52	0.45	0.41	0.43	**0.65**	*-0.02*	0.07	*0.02*	*0.16*	0.07	**0.56**	*-0.05*	*0.05*	*0.06*	0.16	*0.06*	**0.56**	I keep trying at something until I succeed.
78	0.34	0.42	0.37	0.33	0.35	**0.53**	*-0.08*	0.07	-0.03	*0.08*	0.00	***0.67***	-0.11	*0.02*	*0.07*	0.08	*-0.01*	**0.69**	I set goals for myself (make plans for the future).
79	0.47	0.57	0.50	0.45	0.47	**0.72**	*0.13*	0.18	*0.09*	0.04	-0.01	***0.51***	*0.06*	0.16	0.22	*0.04*	*-0.04*	**0.51**	I try to be my best.
Factor Correlations																			
F2	0.87						0.01						0.11						
F3	0.84	0.76					0.43	0.01					0.16	0.22					
F4	0.67	0.67	0.75				0.09	0.16	0.17				0.14	0.16	0.29				
F5	0.89	0.91	0.73	0.69			0.33	0.06	0.37	0.14			0.19	0.28	0.25	0.26			
F6	0.65	0.79	0.69	0.63	0.65		0.37	0.02	0.42	0.12	0.27		0.14	0.17	0.18	0.22	0.18		

* CFA coefficients include structure coefficients for the non-target loadings. NON-shaded indicates structure coefficient.

***Italics* indicates NON-statistically significant coefficient (p>0.05). **BOLD** indicates coefficient >0.3. Underline indicates highest loading for indicator.

Phase II: Structural Validity

As emphasized by Marsh et al. (2010, 2011) and Morin et al. (2013), the first step in conducting an ESEM analysis is to compare the a priori factor model with the hypothesis that the ESEM model provides a better fit over the more restrictive ICM-CFA model. Table 7 includes model fit indices for the CFA and ESEM models. As noted in Phase I, the six-factor model provided the most appropriate fit of the CFA-ICM models. However, comparison of model fit indices warrants retention of the less parsimonious ESEM model (ΔCFI = .035, ΔTLI = .032; ΔRMSEA = -.016, Chen, 2007). Additionally the DIFFTEST indicates the ESEM model fits the responses at least somewhat better (MD$\Delta\chi 2$ = 985.876, df = 115, p <.001).

When considering the ESEM solution with target rotation's factor pattern coefficients shown in Table 5, the Prosocial Behavior, Respect for Teacher, Respect for Parent and Self-Development factors show higher coefficients on target loadings (.883 to .229) with lower loadings on non-target factors. For the Self-Control factor, only two of the target items show the highest factor pattern on Self-Control: Item 2 – I keep my temper when I have an argument with other kids; Item 3 – I ignore other children when they tease me or call me bad names. These two items seem to focus on peer relations. The other two target indicators show higher factor patterns on the Respect for Teacher factor: Item 1 – I wait my turn in line patiently; Item 3 – I follow the rules even when nobody is watching. Both of these items could be associated with school related tasks. For the Honesty factor, only three of the target items show the highest factor pattern coefficient on the target factor: Item 2 – I tell the truth when I have done something wrong; Item 3 – I tell others the truth; Item 5 – I admit my mistakes. The other two Honesty target items load higher on other factors. Item 1 (I apologize when I have done something wrong) exhibits a higher association (p =.342) with the Self-Control factor, which as discussed previously seems to

be associated with peer relations. Item 4 (I keep promises I make to others) has a higher association (p = .305) with Prosocial Behavior. Overall, the ESEM non-target loadings are systematically smaller (.004 to .342, M = .111) than the target loadings (.043 to .883, M = .427). Table 6 reflects the SECDS constructs with indicators rearranged to include items with high cross-loadings.

When comparing target and non-target loadings of the CFA-ICM and the ESEM models, the profile similarity index (PSI = correlation between CFA-ICM loadings where non-target loadings are constrained to 0 and the ESEM loadings) indicates an overall similarity of .698 which illustrates the factor patterns are fairly similar. However, when just considering the more distinct Prosocial Behavior, Respect for Teacher, Respect for Parent and Self-Development factors, the PSI increases to .744 indicating higher similarity between loadings after removing the factors with the highest cross-loadings. Examination of the inter-factor correlations indicates a critical advantage of the ESEM model over the CFA-ICM. Although the patterns of loadings are moderately similar, the factor correlations in the ESEM model (-.024 to .433) are much lower than the CFA-ICM (.629 to .909). The decrease in factor correlations from the CFA-ICM to the ESEM is indicative of misspecifing all CFA-ICM non-target loadings to zero, a problem which is further illustrated by the high CFA-ICM structure coefficients.

Table 6

Comparison of SECDS Factors under ESEM Framework to SEL Components

SECDS Factors	SEL Competencies	Items
Self-Control	**Self-Control**	I keep my temper when I have an argument with other kids.
	Filter negative input	I ignore other children when they tease me or call me bad names.
	Impulse control	*I apologize when I have done something wrong.*
	Regulate emotions and behavior	*I play nicely with others.*
Pro-Social	**Peer Relationship Mgmt & Social Awareness**	I play nicely with others.
	Builds relationships	I do things that are good for the group.
	Relationships with diverse individuals	I treat my friends the way I like to be treated.
	Working cooperatively	I am nice to kids who are different from me.
	Respect for others	I try to cheer up other kids if they are feeling sad.
	Empathy and perspective taking	I am a good friend to others.
	Appreciating diversity	I think about how others feel.
		I keep promises I make to others.
Respect Teacher	**Responsible Decision Making**	I speak politely to my teacher.
	Respectful choices	I obey my teacher.
	Obey and follow rules	I follow the directions of my teacher.
		I listen (without interrupting) to my teacher.
		I follow school rules.
		I wait my turn in line patiently.
		I follow the rules even when nobody is watching.
Respect Parents	**Adult Relationship Management**	I speak politely to my parents.
	Respect for others	I obey my parents.
		I listen (without interrupting) to my parents.
		I follow the rules at home.
		I speak politely to my teacher.
Honesty	**Moral & Ethical Decision Making**	I apologize when I have done something wrong.
	Moral & ethical responsibility	I tell the truth when I have done something wrong.
	Evaluation & reflection	I tell others the truth.
		I admit my mistakes.
Self-Development	**Self-Management**	I make myself a better person.
	Goal setting	I keep trying at something until I succeed.
	Self-motivation	I set goals for myself (make plans for the future).
	Improving self	I try to be my best.

Note. Italics indicates item discovered to have high cross-loadings when examined under the ESEM framework. SECDS factors from Ji et al. (2014). SEL competencies from CASEL (2013).

Phase III: Generalizability Across Gender and Time

Gender Invariance. The gender invariance models follow those outlined by Guay et al. (2014) and Marsh et al. (2011). Since the 4-point likert scale model indicators were considered categorical, the theta parameterization was utilized in order to include uniqueness as a point of constraint among the two groups. In addition, in lieu of item intercepts, categorical indicators warrant the calculation of item thresholds which is the point at which an individual transitions from a response of 0 to a response of 1 on the categorical outcome. Model fit indices for the six models are shown in Table 7.

Weak Factorial/Measurement Invariance: Model 1 vs. Model

Weak factorial/measurement invariance determines if the factor loadings are similar across groups by comparing models where the pattern coefficients are estimated freely across groups versus a model where pattern coefficients are constrained to be equal across groups. Although the DIFFTEST results indicate the more restrictive model provides a decrease in fit, comparisons between fit indices for Model 1 and Model 2 provide support for weak factorial invariance since the change in RMSEA and CFI does not warrant rejection of the more constrained model (ΔCFI = .002, ΔRMSEA = -.005; Chen, 2007).

Strong Measurement Invariance

Model 2 vs. Model 3. Strong measurement invariance is determined by comparing models where, in addition to pattern coefficients, item thresholds are estimated freely (Model 2) versus models where the item thresholds are constrained to be equal across groups (Model 3). Comparisons between Model 2 and Model 3 support retention of the more parsimonious Model 3

(ΔCFI = -.001, ΔRMSEA = <.001). When considering the DIFFTEST and testing at an alpha of .01 as is appropriate when dealing with large sample sizes, the more constrained model would not be considered a decrease in model fit (MDΔχ2 = 77.233, df = 52, p = .013). Support of the more constrained Model 3 provides evidence for lack of differential item functioning or strong measurement invariance which justifies comparison of the latent means across gender.

Strict Measurement Invariance: Model 3 vs. Model 4

Strict measurement invariance is determined by comparing Model 3 where the indicator uniqueness is freely estimated across groups versus Model 4 where uniqueness is constrained to be equal. Comparisons between Model 3 and Model 4 support retention of the more restrictive Model 4 (ΔCFI = <.001, ΔRMSEA = -.001). Likewise, the DIFFTEST supports retention of the more constrained Model 4 (MDΔχ2 = 48.685, df = 29, p = .013). Support of strict measurement invariance indicates measurement error is similar across groups and therefore manifest scores could be reasonably compared.

Factor Variance-Covariance Invariance: Model 4 vs. Model 5

Factor variance-covariance (FVCV) invariance is determined by comparing Model 4 where the FVCV is freely estimated across groups to Model 5 where the FVCV is constrained to be equal. Comparisons between Model 4 and Model 5 provide evidence for retaining the more parsimonious constrained Model 4 (ΔCFI = .008, ΔRMSEA = -.008). The DIFFTEST also provides evidence for adopting the more constrained Model 5 (MDΔχ2 = 24.585, df = 21, p = .266). Determining FVCV invariance across groups is important to being able to compare correlations between the SECDS and other concurrent measures.

Table 7

Model Fit Indices for GENDER Multigroup ESEM Models (Guay, 2014)

Model	Invariant Parameters	χ^2	df	χ^2_{GIRL}	χ^2_{BOY}	CFI	TLI	RMSEA	RMSEA CI	MD$\Delta\chi^2$	$df_{\Delta\chi2}$	$p_{\Delta\chi2}$	ΔCFI	ΔTLI	ΔRMSEA
6 Correlated Factors CFA	-	1773	362	-	-	0.951	0.945	0.046	[.044, .048]	985.876	115	<.001	0.035	0.032	-0.016
ESEM	-	651.8	247	-	-	0.986	0.977	0.030	[.027, .032]	-	-	-	-	-	-
TI-1 Configural Invariance	NONE	908.6	494	458.659	449.891	0.983	0.973	0.031	[.028, .034]	-	-	-	-	-	-
TI-2 Weak Invariance	FL	1014	632	497.915	516.44	0.985	0.980	0.026	[.023, .029]	200.582	138	<.001	0.002	0.007	-0.005
TI-3 Strong Invariance	FL, THOLD	1080	684	515.293	564.324	0.984	0.981	0.026	[.023, .028]	77.233	52	0.013	-0.001	0.001	<0.001
TI-4 Strict Invariance	FL, THOLD, UNIQ	1110	713	548.509	561.191	0.984	0.982	0.025	[.022, .028]	48.685	29	0.013	<.001	0.001	-0.001
TI-5 Variance-Covar Invariance	FL, THOLD, UNIQ, FVCV	924.4	734	469.314	455.044	0.992	0.992	0.017	[.013, .020]	24.585	21	0.266	0.008	0.010	-0.008
TI-6 Latent Means Invariance	FL, THOLD, UNIQ, FVCV, FMN	1498	740	777.672	720.428	0.970	0.967	0.034	[.031, .036]	215.193	6	<.001	-0.022	-0.025	0.017

Where FL = factor loading; THOLD = thresholds; UNIQ = indicator uniqueness/residual; FVCV = factor variance/covariance; FMN = factor means.

Table 8

Difference in Latent Means for BOYS with GIRLS as Referent Group

Factor	M	SE	*p*
Self-Control	-0.270	0.074	<.001
Pro-Social	-0.319	0.073	<.001
Respect for Teacher	-0.297	0.058	<.001
Respect for Parent	-0.108	0.059	0.069
Honesty	-0.437	0.060	<.001
Self-Development	-0.522	0.065	<.001

Based on the evidence of FVCV invariance, comparison of correlations between SECDS manifest variables and other concurrent measures is warranted.

Latent Factor Mean Comparison Across Gender*:* Model 5 vs. Model 6

Invariance across latent means can be determined by comparing Model 5 where the FVCV, thresholds, uniqueness, and pattern coefficients are constrained but the latent factor means are freely estimated to Model 6 where all elements are constrained to be equal across groups. Comparison of the model fit indices supports retention of the less parsimonious Model 5 (ΔCFI = -.022, ΔRMSEA = .017). In other words, constraining the latent means to be equal across groups resulted in decreased model fit. Retention of Model 5 where latent factor means are freely estimated provides evidence for gender differences between the latent means. Since previous multi-group model comparisons provided evidence for strong measurement invariance, the differences indicate latent means vary systematically between boys and girls. Table 8 includes latent means for boys as expressed in SD units from girls' means. When compared to the girls' means which are set at 0 for identification purposes, the boys' means are statistically significantly lower on all factors with the exception of Respect for Parent. The greatest difference in means between girls and boys occurs on the Self-Development factor where boys' mean is 0.522 standard deviations lower than girls' mean (M = -.522, SE = .065, $p < .001$). The Respect for Parent factor showed the lowest gender-based differences (M = -.108, SE = .06, $p = .069$).

Time Invariance. In order to evaluate the potential impact of omitting correlated uniqueness between time periods, two configual models were compared. Model 1 included estimating the correlated uniqueness while Model 1a did not. Comparisons of model fit indices

shown in Table 9 indicate while although the model fit does not decrease substantially (Chen, 2007), the RMSEA confidence intervals do not overlap which suggests there are indeed at least some identifiable differences between the two models. Table 10 compares factor correlations in Model 1a and 1. Although there appears to be no systematic decrease in factor correlations across all factors, the mean of all correlations does decrease slightly (M = .330, SD =.287 versus M = .266, SD = .213), and the factor correlations differ greatly in some comparisons. For example, under Model 1a the test-retest correlation for Respect Teacher is .590 while under Model 1 the test-retest correlation is only .121. Because of the potential impact on future test-retest analysis, the a' priori correlated uniquenesses were included in all further time invariance models – even though inclusions of these additional parameters increase model complexity.

Similar to the protocol for testing multigroup invariance, time invariance models evaluate the stability of components over waves of data instead of groups. Model fit indices for the time invariance models are shown in Table 9. Weak factorial invariance is evidenced by comparison of fit indices for Model 1 and Model 2. Comparison of Model 2 and Model 3 provides evidence of strong measurement invariance which inturn justifies comparison of latent means over time. Strict measurement invariance where uniqueness is held constant is demonstrated by Model 3 and 4 comparisons. Invariance of the factor variance-covariance matrix is supported by Model 4 and 5 comparisons. Comparison of Model 5 where latent means are freely estimated versus Model 6 where latent means are constrained to be equal indicates the more parsimonious constrained model provides an equivalent fit to the data. This can be further interpreted to indicate factor means do not differ systematically over time. It is interesting to note the DIFFTEST probability values indicated differences between all models comparisons except when comparing Model 2 and Model 3 ($MD\Delta\chi 2 = 76.772$, df = 52, p = .014).

Table 9

Model Fit Indices for TIME Invariance ESEM Models (Guay, 2014)

Model	Invariant Parameters	χ^2	df	CFI	TLI	RMSEA	RMSEA CI	MD$\Delta\chi^2$	$df_{\Delta\chi2}$	$p_{\Delta\chi2}$	ΔCFI	ΔTLI	ΔRMSEA
CFA	-	1772.769	362	0.951	0.945	0.046	[.044, .048]	985.876	115	<.001	0.035	0.032	-0.016
ESEM	-	651.841	247	0.986	0.977	0.03	[.027, .032]	-	-	-	-	-	-
TI-1 Configural Invariance	NONE	2049.625	1270	0.987	0.983	0.018	[.017, .020]	-	-	-	-	-	-
TI-1a Configural Invariance	(no correlated uniqueness)	2509.916	1299	0.980	0.975	0.022	[.021, .024]	707.445	29	<.001	-0.007	-0.008	0.004
TI-2 Weak Invariance	FL	2137.892	1408	0.988	0.986	0.017	[.015, .018]	205.548	138	<.001	-0.001	-0.003	0.001
TI-3 Strong Invariance	FL, THOLD	2213.239	1460	0.988	0.986	0.017	[.015, .018]	76.772	52	0.014	0.000	0.000	0.000
TI-4 Strict Invariance	FL, THOLD, UNIQ	2346.482	1489	0.986	0.984	0.018	[.016, .019]	108.896	29	<.001	0.002	0.002	-0.001
TI-5 Variance-Covariance Invariance	FL, THOLD, UNIQ, FVCV	2588.666	1510	0.982	0.980	0.020	[.018, .021]	109.033	21	<.001	0.004	0.004	-0.002
TI-6 Latent Means Invariance	FL, THOLD, UNIQ, FVCV, FMN	2672.439	1516	0.981	0.979	0.020	[.019, .021]	54.563	6	<.001	0.001	0.001	0.000

Where FL = factor loading; THOLD = thresholds; UNIQ = indicator uniqueness/residual; FVCV = factor variance/covariance; FMN = factor means.

Table 10

Test-Retest Correlations Between SECDS Factors With and Without CU Estimation

	Model 1a - No CU	Model 1 - CU estimated
F1: Self-Control	.133	.143
F2: Prosocial	.075	.155
F3: Respect Teacher	.590	.121
F4: Respect Parent	.782	.516
F5: Honesty	.188	.098
F6: Self-Develop	.209	.563
Mean (SD)	.330 (.287)	.266 (.213)

Note. All correlations are statistically significant at p=.05.

However, evaluation of the RMSEA CIs between models show clear overlap - and in the instance of Model 2 and 3, complete overlap. In lieu of any published simulation studies investigating the sensitivity of DIFFTEST, it is assumed the discrepancy between interpretation based on model fit indices and interpretation of DIFFTEST significance could be attributed to the large sample size.

Discussion

In the present study, the validity of the SECDS was examined through a three phase investigation. Phase I examined the generalizability and structural aspects of validity under the methodological framework demonstrated in a recently published article which examined the SECDS construct validity utilizing a sample of U.S. students (Ji et al., 2013). Phase II extended the structural evidence of construct validity by examining the SECDS measurement model under the ESEM framework. Phase III sought to extend the generalizability evidence of the SECDS construct validity through multi-group and time invariance ESEM models.

In Phase I, the replication of the structural model as demonstrated by Ji et al., (2013) seemed to fit the Belize sample data. Although the hypothesized higher-order factor model met acceptable fit standards where model fit indices are concerned, the Belize data was slightly better fitted to the six-correlated factor model. Since recent SEL and character development reviews call for instruments which measure multiple distinguishable facets of the SEL constructs, retention and further examination of the six-factor model was substantively warranted (Humphrey, et al., 2011; Wigelsworth et al., 2010). Similar to Ji et al.'s (2013) findings, examination of the ICM-CFA six factor structure revealed high factor correlations as well as high structure coefficients. As Asparouhov and Muthen (2009), Marsh et al., (2011), Marsh et

al., (2010), Morin et al. (2013), and others point out, misspecification of non-target zero loadings in ICM-CFA models can lead to over inflation of factor correlations which in turn can lead to biased estimates in further examined SEM models. In addition, high factor correlations are indicative of low discriminant validity, rendering the SECDS factors virtually indistinguishable as separate constructs. The ICM-CFA high factor correlations and high structure coefficients provide substantive cause for further investigation of the SECDS under the ESEM framework.

In Phrase II the structural evidence of construct validity was extended through evaluation of the SECDS under the ESEM framework. Consistent with demonstrations in recently published ESEM literature, the ESEM six-factor structure of the SECDS provided a slightly better fit and suggests that the magnitude of inter-factor correlations is lower (Guay, 2014; Marsh et al. 2011). Substanatively speaking, the reduction in factor correlations greatly improves the viability of the SECDS by helping distinguish between factors associated with different SEL programing components. While in many instances factor loadings show similar patterning to the ICM-CFA loadings, the ESEM model allowed for expression of some very notable cross-loadings

In addition to methdological advantages of the ESEM model, inclusion of non-target loadings indicates the need for a substantive change in how the SECDS factors are being defined. Table 6 shows the alignment of the SECDS six factor structure with the generalized SEL competencies as defined by CASEL (2013). As noted, the items in italics include those with high cross-loadings as discovered through the ESEM model.

The SECDS Self-Development factor aligns well with the SEL Self-Management factor to include goal setting, motivation and improvement of self. No additional indicators loaded heavily on the Self-Development construct which would seem to indicate a certain degree of

discriminate validity. Instead of retaining only a single SEL Decision Making compentency, evaluation of the items loading on Respect for Teacher and Honesty seem to key in on both responsible decision making and moral/ethical decision making elements. The SECDS Honesty factor aligns with the SEL Decision Making compentency but more specifically with the moral and ethical decision making facet. Items which loaded on the original SECDS factor congregate around the theme of following rules and making respectful choices – or rather Responsible Decision Making. Similarly, instead of a single SEL Relationship Management compentency, the crossloadings on the SECDS Pro-Social and Respect Parents factors provide for interpretation of separate peer and adult Relationship Management elements. The high cross-loadings of Teacher Respect indicators on the Parent Respect items point specifically to Adult Relationship Management compentency. While the highly loaded items on the SECDS Pro-Social factor are specific to Peer Relationship Management. The remaining SECDS Self-Control factor is relatively analogous to the SEL Self-Control compentency in that the high-loading indicators involve regulating emotions, filtering negative input, and impulse control.Considering the re-conceptualization of the SECDS factor structure under the ESEM framework, the six factor structure can be considered to fit more generally into the larger conceptualization of the SEL compentencies while also retaining is applicability to the specific Positive Action program components (CASEL, 2012; Positive Action, 2013; Zins et al., 2004). Retaining the original six factors, yet re-defining the factors under the findings of the ESEM model increases the utility of the SECDS and helps meet a noted need in the SEL literature for instruments designed to measure unified concepts across multiple programs (Humphrey et al., 2011).

Phase III extended the generalizability evidence of the SECDS over time and gender. The series of models examining the invariance of components across gender indicates the

SECDS held up to strict measurement invariance as well as factor variance-covariance invariance. As a result, the latent mean differences discovered in the final model comparison can be interpreted as systematic differences in the latent mean scores of boys and girls. Similar results, where males exhibit lower SEL and character development manifest means scores have been noted by other authors (e.g. Endrulat et al., 2010; Taylor et al., 2002).

The occurance of varied gender-based latent mean differences on the six factors provides additional evidence of discriminate validity provided by examination of the SECDS under the ESEM framework. In opposition, under the ICM-CFA model with high correlations between factors variations of the latent mean differences for the different SECDS factor would likely not be noticed since the high correlations render the factors essentially identical mathematically. Being able to detect the variation in gender-based latent mean differences across constructs is an additional benefit of examining the SECDS under the ESEM framework. Following a similar protocol to evaluating group differences, the time invariance models demonstrate the SECDS to exhibit strict invariance across time in addition to indicating there are no systematic latent mean differences between Time 1 and Time 2.

Conclusion

The SECDS exhibits structural and generalizability evidence of construct validity when examined under the ESEM framework. While the initial higher order SECD factor with six secondary factors provided acceptable fit to the Belize sample data, the ESEM six factor structure provided both substantive and methodological advantages. The ESEM six-factor structure decreased the high factor correlations as seen under the ICM-CFA model and allowed for the expression of high cross-factor loadings. The lower factor correlations provide at least

some level of discriminate validity, which renders the six factors usable in larger SEM models designed to compare the SEL facets to other purported concurrent constructs. Interpretation of the SECDS factors under the ESEM framework allows for fitting of the SECDS into the larger body of SEL literature. In addition, the ESEM SECDS six-factor structure exhibits generalizability evidence over both gender and time.

While evaluation of the SECDS under the ESEM framework poses significant substantive advantages and exhibits structural and generalizability evidence of construct validity, this initial investigation utilizing a Belizean sample does not warrant cessation of further examination of the SECDS under the ICM-CFA framework. Instead the current findings demonstrate the need to expand the construct validation of the SECDS and other similar SEL instruments to include evaluation under both ICM-CFA and ESEM frameworks. As shown with the SECDS, examination under the more flexible ESEM framework could allow previously developed SEL instruments to be redefined or expanded to include the more generally accepted SEL competency constructs.

Limitations and Future Work

The present investigation examined the structure of the SECDS under the ESEM framework using only data gathered from a sample of Belizean children ages 9 to 13; therefore the results cannot be generalized to other populations. The currently assessed self-reported SECDS version could also be impacted by students engaging in socially desirable response patterns. A multigroup analysis evaluating model fit over both Belizian and U.S. samples should be conducted under the ESEM framework. In addition, further investigation surrounding the SECDS's discriminant validity is needed. For example, an ESEM-MTMM as outlined by Morin

et al. (2013) would further elucidate the differences between SECDS factors and other related constructs as called for by Wigelsworth et al. (2010). Since the SECDS also includes a yet unexamined teacher report version, efforts should be made to establish the SECDS as a multiple-reporter cross-validated instrument, another need noted in Wigelsworth et al.'s (2010) review of current SEL measures. Although the SECDS has been subjected to brief evaluation of reliability under classical test theory applications, no published literature has included an examination of SECDS indicators' performance under IRT applications. Since SEL instruments seek to measure levels of SEL construct competencies over all levels (as opposed to establishing a cutoff score), it is important to add IRT indicator performance into consideration when establishing reliabilities instead of interpreting solely the omnibus alpha coefficient.

Being a more recently utilized method in the construct validity literature, the methodological limitations surrounding the use of ESEM are numerous. One of the more obvious areas for future work in the area of comparing ESEM models includes further investigation of best practice concerning comparing models. For example, while previous studies have established general guidelines for comparison of model fit indices for nested models which included continuous indicators, no published literature establishes guidelines for use of the model fit comparisons in models with categorical indicators. In addition, no model fit indices have been developed for comparision over multiple imputed datasets. Another limitation includes the current limitation of MPlus to evauate ESEM measurement models under multilevel design or to include the ESEM measurement model in higher order factor models.

References

Asparouhov, T., & Muthen, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 397-438. doi:10.1080/10705510903008204

Berkowitz, M. W., & Bier, M. C. (2004). Research-based character education. *American Academy of Political and Social Science*, 72-85. Retrieved from http://www.jstor.org/stable/4127636

Byrne, B. (2012). *Structural equation modeling with Mplus.* New York, NY: Routledge.

Byrne, B. M. (1998). *Structural equation modeling with Lisrel, Prelis, and Simplis: Basic concepts, applications, and programming.* Mahwah, NJ: Lawrence Earlbaum.

CASEL. (2002). *Guidelines for social emotional learning.* Retrieved September 20, 2013, from www.casel.org: http://casel.org/wp-content/uploads/GuidelinesAug02.pdf

CASEL. (2011). *What is SEL?* Retrieved September 18, 2014, from www.casel.org: http://casel.org/why-it-matters/what-is-sel/

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

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 464-504. doi:10.1080/10705510701301834

Damon, W. (2004). What is positive youth development? *American Academy of Political and Social Science*, 13-24. doi:10.1177/0002716203260092

Dimitrov, D. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 121-149. doi:10.1177/0748175610373459

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, 294-309.

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*, 405-432. doi:10.1111/j.1467-8624.2010.01564.x

Endrulat, N. R., Tom, k., Ravitch, K., Wesley, K., & Merrell, K. W. (2010). Gender differences in positive social emotional functioning. *Annual Meeting of the National Association of School Psychologists.* Lecture conducted from Chicago, IL.

Flay, B. R., Biglan, A., Boruch, R. F., Gonzàlez Castro, F., Gottfredson, D., Kellam, S., . . . Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 151-175. doi:10.1007/s11121-005-5553-y

Greenberg, M. T. (2004). Current and future challenges in school-based prevention: The research perspective. *Prevention Science*, 5-13. doi:10.1023/B:PREV.0000013976.84939.55

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologists*, 466-474. doi:10.1037/0003-066X.58.6-7.466

Guay, F., Morin, A. J., Litalien, D., Valois, P., & Vallerand, R. J. (2014). Application of exploratory structural equation modeling to evaluate the academic motivation scale. *Journal of Experimental Education*, 1-32. doi:10.1080/00220973.2013.876231

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement*, 617-637. doi:10.1177/0013164410382896

Ji, P., DuBois, D. L., & Flay, B. R. (2013). Social-emotional and character development scale: Development and initial validation with urban and elementary school students. *Journal of Research on Character Education*, 121-147.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, 320-341. doi:10.1207/s15328007sem1103_2

Marsh, H. W., Ludtke, O., Muthen, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big-five factor structure through exploratory structural equation modeling. *Psychological Assessment*, 471-491. doi:10.1037/a001922

Marsh, H. W., Ludtke, O., Nagengast, B., & Morin, A. J. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right - camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, 257-284. doi:10.1037/a0032773

Marsh, H. W., Muthen, B., Asparouhov, T., Ludtke, O., Robitzsch, A., Morin, A., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Applications to students' evaluations of university teaching. *Structural Equation Modeling*, 439-476. doi:10.1080/10705510903008220

Marsh, H. W., Nagengast, B., Morin, A. J., Parada, R. H., Graven, R. G., Craven, R. G., & Hamilton , L. R. (2011). Construct validity of the multidimensional structure of bullying and victimization: An application of exploratory structural equation modeling. *Journal of Educational Psychology*, 701-732. doi:10.1037/a0024122

Messick, S. (1995). Validity of psychological assessment: Validaton of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 741-749. doi:10.1037/0003-066X.50.9.741

Morin, A. J., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 395-436). Charlotte, NC: Information Age Publishing, Inc.

Mundi. (2013, Feb 21). *Belize economy profile 2013*. Retrieved Feb 2, 2014, from www.indexmundi.com: http://www.indexmundi.com/belize/economy_profile.html

Muthen, L. K., & Muthen, B. O. (2010). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthen & Muthen.

Park, N. (2004). Character strengths and positive youth development. *The ANNALS of the American Academy of Political and Social Science*, 40-54. doi:10.1177/0002716203260079

Park, N., & Peterson, C. (2008). Positive psychology and character strengths: Application to strengths-based school counseling. *Professional School Counseling*, 85-92. doi:10.5330/PSC.n.2010-12.85

Shavelson, R. J., Hubner, J. J., & Stanton, J. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 407-441. doi:10.3102/00346543046003407

Taylor, C. A., Liang, B., Tracy, A. J., Williams, L. M., & Seigle, P. (2002). Gender differences in middle school adjustment, physical fighting, and social skills: Evaluation of a social competency program. *The Journal of Primary Prevention*, 259-272. doi:doi:10.1023/A:1019976617776

Thompson, B. (1997). The importance of structure coefficients in structural equation modeling confirmatory factor analysis. *Educational and Psychological Measurement*, 5-19. doi:10.1177/0013164497057001001

Weissberg, R. P., & O'Brien, M. U. (2004). What works in school-based social and emotional learning programs for postive youth development. *Annals of the American Academy of Political and Social Science*, 86-97. doi:10.1177/0002716203260093

Wigelsworth, M., Humphrey, N., Kalambouka, A., & Lendrum, A. (2010). A review of key issues in the measurement of children's social and emotional skills. *Educational Psychology in Practice*, 173-186. doi:10.1080/02667361003768526

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior. *American Journal of Preventative Medicine*, 130-143. doi:10.1016/j.amepre.2007.04.011

Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (2004). *Building academic success on social and emotional learning: What does the research say?* New York, NY: Teachers College Press.

APPENDIX A

EXTENDED LITERATURE REVIEW

Strategies to improve school achievement have shifted the balance beyond academic core curriculum to focus on social-emotional learning and character development as a means to increase academic achievement. In the United States, 18 states legislatively mandate character education, 18 encourage character education, and seven support character education without direct legislation (Character Education Partnership, 2013). Societies such as the International Academy of Education (IAE) and International Bureau of Education (IBE) have stressed the value of social-emotional learning to every society as a means to achieve universally expressed criteria concerning what parents want young people to know and be able to do (Elias, 2003).

In response to educational policy changes schools have begun molding dual-core curriculum emphasizing both social-emotional and character development (SECD) and academic learning (Elias, 2009). National organizations such as the Collaborative for Academic, Social, and Emotional Learning (CASEL) and the Character Education Partnership (CEP) recommend schools implement social-emotional interventions through adoption of empirically researched and theory-based social-emotional learning and character education curriculum. The purpose of SECD programs is to educate the whole child, going beyond academic literacy to address the students' social-emotional and character development competencies in order to prevent negative behavior and promote positive outcomes such as social competence and academic achievement. Meta-analyses of empirically evaluated social-emotional learning (SEL) and SECD programs indicate the programs lead to improvements in social emotional skills, behavior, and academic performance. However several studies point to the need for better developed measures and accountability systems for evaluating the programs (Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Greenberg et al., 2003; Weare & Nind, 2011; Weissberg & O'Brien, 2004; Zins, Weissberg, Wang, & Walberg,

2004). Although many program-specific instruments have been designed to measure SECD, few have ventured beyond single program use to establish a valid and reliable instrument to be used across different programs to evaluate program efficiency. Even fewer have been subjected to psychometric investigation to establish generalizability evidence of construct validity across time and different groups (Catalano et al., 2004; Coryn, Spybrook, Evergreen, & Blinkiewicz, 2009; Wigelsworth, Humphrey, Kalambouka, & Lendrum, 2010).

Social Emotional Learning and Character Development

Schooling is inherently a social activity. Children learn within the context of interactions with teachers, peers, families, and the larger community. Because of the closely intertwined relationship between education and socialization, social-emotional skills not only impact students' societal success but also broadly impact academic performance. An integrated approach to social-emotional learning along with academic learning leads to greater school success (Elias et al., 1997).

Grounded largely in Bandura's (1986) social cognitive theory, social emotional learning (SEL) involves the processes of developing social and emotional competencies (CASEL, 2011). Viewed as a broadly defined category, SEL provides an over-arching framework for several subsets of developmental programs such as character development, positive youth development, and emotional intelligence. SEL programs strive to holistically address students' social, emotional, ethical, and academic development within a safe supportive environment (Zins et al., 2004; Elias, 2009; Weissberg & O'Brien, 2004).

Although social responsibility and moral character has long been included in educating children, historically, programs with social-emotional components have appeared in school

routines as fragmented responses to target a single facet of non-academic development. As a result, schools often resort to adopting several social-emotion programs at one time such as one program for health promotion, another program for violence prevention and another to reduce teen pregnancy. In a 2001 survey of 848 public and private schools, the median number of problem behavior prevention programs was 14, with an upper range of 66 programs (Gottfredson & Gottfredson, 2001). The end result of multi-program adoption meant social-emotional programs were ultimately unsuccessful due to uncoordinated efforts across several domains and programs being sporadically implemented and eventually dropped from use. In addition, social-emotional skills were taught in isolation without providing implicit teaching of the underlying social-emotional developmental attributes which impact the root of many risk behaviors (Zins et al., 2004).

Social-emotional learning developed in response to school programs designed to target specific problem youth behaviors such as violence and substance abuse (CASEL, 2002). Instead of focusing on the resulting problem behavior, SEL provides a preventative framework for addressing underlying causes of negative youth behaviors while also supporting academic improvement (Greenberg et al., 2003; Weissberg & O'Brien, 2004). Although several frameworks exist in the literature, SEL generally addresses 5 sets of inter-related cognitive, affective, and behavioral competencies: self-awareness, social awareness, responsible decision making, self-management, and relationship management (Zins et al., 2004; Weissberg & O'Brien, 2004; CASEL, 2011).

Self-Awareness

Self-awareness encompasses the ability to identify and recognize one's own emotions

and the ability to reflect on how those emotions relate to particular behaviors. In addition, self-awareness involves accurately perceiving one's self by recognizing personal strengths, needs, and values. Self-awareness acts as a precursor to the development of self-efficacy which is described by Bandura (1994) as the belief in one's ability to self-regulate and manage behavior. Self-efficacy associated with self-awareness is indicated by maintaining a well-adjusted amount of self-confidence and optimism (Zins et al., 2004; Weissberg & O'Brien, 2004; CASEL, 2011).

Social Awareness

Extending self-awareness beyond self into interactions with the environment, social-awareness includes the ability to grasp other's perceptions and empathize with people from a variety of different backgrounds and cultures, while recognizing social norms. Respecting others' opinions, thoughts, and actions in addition to appreciating diversity are considered social-awareness. In addition, social-awareness involves seeing how one fits into a larger network of support systems involving family, friends, and community (Zins et al., 2004; Weissberg & O'Brien, 2004; CASEL, 2011).

Self-Management

Similar to Bandura's (1987) idea of self-regulation, self-management extends self-awareness into regulating emotions, thoughts, and behaviors to provide appropriate responses according to the contextual environment and situation. Self-management skills include managing stress and controlling impulses in addition to self-motivating and exercising self-discipline in order to set and meet personal and academic goals (Zins et al., 2004; Weissberg & O'Brien, 2004; CASEL, 2011).

Responsible Decision-Making

Responsible decision-making involves making responsible decisions based on a global consideration of how those decisions interact with personal moral and ethical beliefs, in addition to how those decisions impact others and fit into societal norms. Responsible decision-making also includes evaluating and reflecting upon decisions in order to effectively problem solve (Zins et al., 2004; Weissberg & O'Brien, 2004; CASEL, 2011).

Relationship Management

Relationship management includes the ability to communicate and socially engage to establish and maintain healthy beneficial relationships with diverse individuals and groups. Relationship management skills also include providing help to others, working cooperatively, negotiating, and conflict management, while also maintaining the ability to resist maladaptive social pressures and seek help from others (Zins et al., 2004; Weissberg & O'Brien, 2004; CASEL, 2011).

Social-Emotional Learning Program Characteristics

The purpose of social-emotional learning is to develop social and emotional compentencies in children. The goal of social-emotional learning programing is to create an environment conducive to social-emotional learning by developing challenging, engaging and meaningful relationships. In following Bandura's (1987) idea of reciprocal causation where person, and environment mutually influence behavior, social-emotional learning provides a framework for addressing social, emotional, and academic needs which includes a dual focus on both person-centered and environment-centered objectives (Hawkins, Smith, & Catalano, 2004).

Person-centered objectives promote social-emotional growth for students to self-evaluate and regulate, engage in appropriate positive social interactions, develop healthy relationships, and learn to provide and lean on others for support. Person-centered objectives help develop goal-driven actions and behaviors in addition to encouraging collaborative problem-solving while considering the impact on others, personal ethics, and societal norms. Academically, a person-centered focus leads students to be (a) able to set personal academic goals, (b) engage in productive communication with peers and teachers, (c) organize themselves towards reaching their goals, and, ultimately, (d) foster commitment to school and academics (Zins et al., 2004).

The achievement of person-centered SEL objectives depends on creating a learning environment with positive social-emotional environmental factors. As such, success of SEL programming hinges on creating a safe, caring, and supportive environment. Under the SEL framework, open communication, high expectations, established classroom structure and rules, supportive district policies, and active involvement of parents and community are important considerations (Jennings & Greenberg, 2009).

Reaching beyond the classroom, the environmental focus promotes an extendedsupport system while also giving students the opportunity to develop and practice skills across a variety of settings. Providing students with opportunities to see behaviors modeled and reinforced across environments eventually facilitates transferring behaviors to real world situations (Bandura, 1987). Allowing students to develop SEL skills under a larger context enables students to eventually become mobilized, productive, caring citizens (Greenberg et al., 2003; Schaps, Battistich, & Solomon, 2004).

In an effort to better understand how schools can successfully impact student social-emotional and character development through SEL programs, CASEL conducted a series of

visits to schools strong both academically and at building social, emotional, and character skills. From the successful schools, five main characteristics of SEL implementation were noted: a school climate articulating specific themes, character elements, and values; explicit instruction in SEL skills, health promotion, and problem-prevention skills; systems to enhance coping skills and social support for transitions, crisis, and conflict resolution; widespread systematic opportunities for contributory service; and strong parent education and involvement components (Devaney, O'Brien, Resnik, Keister, & Weissberg, 2006; Elias, 2009; Elias et al., 1997; Payton et al., 2008). Adding to the list of school characteristics, a growing body of research indicates successful SEL programs generally incorporate the following attributes: grounded in theory and research; teach children to apply SEL skills in daily life; build connections between school and communitity; provide developmentally and culturally-appropriate instruction; unify often fragmented programs; address the affective and social dimensions of academic learning; create family and community partnerships; establish organizational supports and policies to foster success; provide high quality staff development and support; and incorporates contiuous evaluation and improvement (CASEL, 2013; Cohen, 2003; Greenberg, et al., 2003).

Several school-based programs have been developed under the SEL unifying framework (Catalano et al., 2004; Durlak et al., 2011; Greenberg et al., 2003; Weissberg & O'Brien, 2004; Zins et al., 2004). These programs operate on the understanding of the underlying cause to a variety of youth problems are based on the same risk factors. As such, SEL programs incorporate social-emotional learning through prevention-focused, long-term programs of effective classroom instruction coupled with emphasis on developing supportive, collaborative relationships among peers and community (Weissberg & O'Brien, 2004).

Character Development

Social-emotional and character development subsumes many of the SEL skills yet extends to specifically address character development (Elias, 2009). Berkowitz (2004) explains character as a multi-faceted psychological construct consisting of moral action, values, emotions, reasoning, identity, and foundational characteristics. Grounded in the moral development theory of notable psychologists such as Kohlberg (1969), who believed moral thinking could be promoted educationally through social interaction and participation in a positive moral environment, character education is a form of moral education which includes teaching students respect, compassion, responsibility, self-control and loyalty (Park, 2004). Since character is considered a multi-dimensional psychological construct, recent research calls for approaching character development programs from a multidimensional perspective, focusing on character as a whole instead of in individual components (Park, 2004).

Character development programs extend the five SEL compentencies to include direct teaching of core values such as respect, responsibility, honesty, fairness, compassion, courtesy, and courage (Park, 2004). As a result, character development focuses on developing good character, defined as principles valued by society to reflect decisions beneficial to the person but also to others and society as a whole (Park & Peterson, 2008). Character education in the SEL framework is associated with positive behaviors under a wide variety of social-emotional learning competencies such as prosocial behavior, school attachment, responsibility, respect, self-efficacy, self-control, social skills, and academic performance (Berkowitz & Bier, 2004).

Evidence for Social-Emotional Learning and Character Development Program Effectiveness

Empirical evidence have shown Social-Emotional Learning programs to demonstrate both direct and indirect effects on academic performance and greater success in school in life

(Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Zins, Weissberg, Wang, & Walberg, 2004). Following the dual-core purpose as discussed above, SEL programs work to (a) enhance the learning environment by creating safe, supportive, well-managed classrooms and schools; and (b) provide developmentally-appropriate direct instruction centered around the five main SEL competencies (CASEL, 2013). Extending beyond correlational relationships, numerous experimental and quasi-experimental studies have consistently demonstrated both the direct and indirect effects the SEL learning environment have had on academic and life success. In an effort to summarize the empirical evidence surrounding the effectiveness of SEL programming on increased academic performance and societal success, several meta-analyses and reviews of program effectiveness have been recently published (Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Durlak, Weissberg, & Pachan, 2010; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Greenberg et al., 2003; Social and Character Development Research Consortium, 2010; Zins, Weissberg, Wang, & Walberg, 2004;).

Three separate meta-analysis encompasing 317 studies of SEL interventions with over 324,303 students in kindergarten through 8th grade examined effectiveness of SEL programs on increasing positive feelings and behaviors, decreasing problem behaviors, and elevating academic performance (Durlak et al., 2010; Durlak et al., 2011; Payton et al., 2008). The meta-anlyses included (a) universal interventions across entire general student bodies without any identified behavior or emotional difficulties (Durlak et al., 2011); (b) indicated interventions which focus on students identified as having displayed early signs of emotional or behavioral problems (Payton et al., 2008); and (c) after-school interventions involving primarily students not identified as having behavioral problems (Durlak et al., 2010). Main findings of the metaanalyses concluded SEL programs to be effective in increasing social-emotional skill

development, decreasing conduct problems, and increasing academic performance. These effects were demonstrated to be consistent across setting (in-school and after-school) and subjects (general population and identified problem population) (Payton, et al., 2008). Appendix B summarizes the results of these meta-analysis.

In another meta-anlysis examining SEL from a positive youth development perspective, Catalano et al. (2002) initially identified 161 programs to be potentially included. Out of the 161 programs only 25 were implemented across general populations, included strong evaluative quasi-experimental or experimental designs, presented an acceptable standard of statistical proof with adequate methodological detail and indicated a positive PYD program effect on behavioral and academic outcomes. Nineteen of the PYD effective programs showed significant improvement in interpersonal skills, peer and adult relationships, self-efficacy, self-control, problem-solving, commitment to schooling, cognitive competencies, and academic achievement. Twenty-four of the implemented PYD programs indicated a decrease in maladaptive behaviors such as violence, high-risk sexual behavior, truance, and drug and alcohol use.

Wang, Haertel and Walberg (1997) analyzed 179 handbook chapters and reviews, 91 meta-analyses, and surveyed 61 educational researchers to identify the most influential factors on academic learning. Twenty-eight categories were considered, and of the top 11 categories, 8 included social-emotional competencies: school culture, peer group, classroom climate, classroom management, parental support, student-teacher social interactions, and motivation. Based on these findings, Wang et al. concluded social-emotional interventions to be the most promising towards achieving greater academic performance.

Programs such as *Positive Action*, which focuses on both social emotional and character development through a series of units designed to incorporate learning across school, home, and

community environments, have been shown to increase SEL skills and subsequently decrease negative behaviors (Lewis et al., 2012; Beets et al., 2008; Flay, 2010; Washburn et al., 2011; Flay et al., 2001). Specific to the *Positive Action* program, Washburn et al. (2011) summarized the results of three longitudinal randomized trials including four years of data from 20 Hawaii schools, three years of data from 14 Chicago schools, and three years of data from eight schools in a southeastern state. In all three studies, students assigned to the *PA* intervention experienced a lesser decrease in positive behaviors. These results indicate targeted social-emotional and character development, as presented in the *PA* program, mitegated the decrease in positive behaviors often experienced with children aged 6 to 11. Flay et al. (2001) conducted a matched-control comparison study of a *PA* intervention which indicated a 16% improvement in academic achievement in one district and 52% in the other, while also decreasing disciplinary referrals by 78% and 85%, respectively. In another 2013 study by Lewis et al., results of a cluster-randomized trial involving 14 Chicago public schools over a 6-year period with grades 3 to 8 indicated increased positive affect (ES = .17), life satisfaction (ES = .13) and lower depression (ES = -.14) and anxiety (ES = -.26) for students assigned to the treatment group.

Measuring Social-Emotional Learning and Character Development

Although decades of empirical research surrounding the effects of social-emotional learning and character development have been published, issues regarding instruments to measure SECD skills remain unresolved. In a report issued by the Society for Prevention Research intended to standardize the criteria for identifying prevention programs which have been sufficiently empirically tested, a standard was set to include measures which were psychometrically sound, meaning the measures have been demonstrated to exhibit construct

validity and reliability (Flay et al., 2005). Greenberg's (2004) suggestions for future research in prevention science called for the development of easily utilized, valid and reliable assessments of social, emotional, ethical and health outcomes. More specifically, Greenberg highlighted the need to develop meaningful and easily understood assessments of social and emotional competence. Likewise, Cohen (2003) emphasized the need to reframe the goals of education to extend beyond examining academic outcomes to include outcomes utilizing psychometrically sound measures of social-emotional and ethical learning. These SEL measures could then be used for planning and educating the public on school needs beyond academic performance to include a focus on protective factors, problem behaviors and school climate (Greenberg, 2004).

Despite the call for the utilization of psychometrically sound instrumentation, the meta-analysis by Durlak et al. (2011) concluded 24% of the examined empirical studies on SEL programs did not use reliable outcome measures and 50% did not use valid outcome measures. In response to the call for further empirical evaluation of social-emotional and character development programs, organizations such as the National Center for Education Evaluation and Regional Assistance and the Forum for Youth Investment produced reports outlining several available measures of youth program outcomes for social-emotional and character education programming (Wilson-Ahlstrom, Yohalem, DuBois, & Ji, 2011; Person, Moiduddin, Hague-Angus, & Malone, 2009). All of these reports indicate a need for further psychometric investigation of the available instruments.

A review by Wigelsworth et al. (2010) summarizes the continued need surrounding the measurement of children's social and emotional skills to include (a) establishment of a concensus regarding the definitions of social-emotional skills and compentencies; (b) implementation of measures for multi-dimensional constructs instead of uni-dimensional

constructs; (c) utilization of typical (self-report) and maximal (task oriented to provide evidence of an underlying construct) measures; (d) evidence to distiguish SEL measures from measures of personality and cognitive ability; (e) utilization of multiple perspective reporting (student, parent, teacher, peers); (f) establishment of the purpose and setting for which the measures might be used (time, cost, monitoring, identifying problem students, targeted population); and (g) examination of psychometric properties of SEL measures. More important to the present study, Wigelsworth et al. (2010) points out there have been little psychometric analysis of the application of SEL measures across varying populations and ethnicities. In addition, more advanced analysis such as item response theory evaluation is lacking.

In a systematic review of measures designed to assess social-emotional skills Humphrey et al. (2011) began with 187 measures of social and emotional skills and reduced the initial list by only including measures which targeted children within a wide age range, sought to measure a broad range of social and/or emotional skills, were available in English, were intended to be completed by the child, and had been validated. The resulting 52 potential measures were further reduced by only including measures which had been used in four or more articles in peer-reviewed academic journals. After conducting an indepth review of the final 12 retained measures, Humphrey et al. (2011) concluded (a) most measures of social-emotional skills have a relatively short "shelf-life," appearing very infrequently in the literature; (b) there is a great imbalance among the scope and type of measures identified, with more measures designed to capture social skills as opposed to emotional skills or both; (c) few have developed versions to offer a range of possible respondents (e.g. child version, parent version and teacher version); (d) the majority of measures have been developed only with American populations and few have

developed any norms; and (e) there is little analysis of the applicability of the measures across different groups (e.g. ethnicity, gender).

Social Emotional and Character Development Scale

Specific to *Positive Action*, Ji, DuBois and Flay (in press) developed and conducted initial validation of a social-emotional and character development scale under the SEL framework. Meant to address the need for a multi-dimensional SEL instrument which captures both social and emotional skills, the Social Emotional and Character Development Scale (SECDS) was designed to assess skills and behaviors with likely relevance to both social-emotional learning and character development programs. Spanning across six person-centered and environment-centered compentencies, the SECDS factor structure was found to be consistent with current theories regarding the multi-dimensional and heirarchical nature of SECD. The six SECDS constructs capture the school-related aspects of the five larger social emotional learning constructs which have been shown to be beneficial to increasing academic performance. The study, which utilized data gathered from 459 Chicago students grades 3 to 5 over five waves of data collection, also indicated the SECDS to show concurrent validity with several related outcome measures in addition to high test-retest and internal reliability across gender and ethnic groups. Table 2 includes the SECDS items and the associated constructs.

References

Bandura, A. (1986). *Social foundations of thoughts and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (pp. 71-81). San Diego, CA: Academic Press.

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K.-K., & Allred, C. (2008). School climate and teachers' beliefs and attitudes associated with implementation of the positive action program: A diffusion of innovations model. *Prevention Science*, 264-275. doi:10.1007/s11121-008-0100-2

Berkowitz, M. W., & Bier, M. C. (2004). Research-based character education. *American Academy of Political and Social Science*, 72-85. Retrieved from http://www.jstor.org/stable/4127636

CASEL. (2002). *Guidelines for social emotional learning.* Retrieved September 20, 2013, from www.casel.org: http://casel.org/wp-content/uploads/GuidelinesAug02.pdf

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

Catalano, R. F., Berglund, L., Ryan, J. A., Lonczak, H. S., & Hawkins, J. D. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *The ANNALS of the American Academy of Political and Socal Science*, 98-124. doi:10.1037/1522-3736.5.1.515a

Character Education Partnership. (2013). *Character Education Legislation*. Retrieved September 18, 2013, from www.character.org: http://www.character.org/more-resources/character-education-legislation/

Cohen, J. (2003). Social, emotional, ethical, and academic education: Creating a climate for learning, participation in a democracy, and well-being. *Harvard Educational Review*, 201-285.

Coryn, C. L., Spybrook, J. K., Evergreen, S. D., & Blinkiewicz, M. (2009). Development and evaluation of the social-emotional learning scale. *Journal of Psychoeducational Assessment*, 283-295. doi:10.1177/0734282908328619

Devaney, E., O'Brien, M. U., Resnik, H., Keister, S., & Weissberg, R. P. (2006). *Sustainable schoolwide social and emotional learning: Implementation guide and toolkit.* Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, 294-309.

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*, 405-432. doi:10.1111/j.1467-8624.2010.01564.x

Elias, M. J. (2003). *Academic and social-emotional learning.* Brussels, Belgium & Geneva, Switzerland: International Academy of Education & International Bureau of Education.

Elias, M. J. (2009). Social-emotional and character development and academics as a dual focus of educational policy. *Educational Policy*, 831-846. doi:10.1177/0895904808330167

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., . . . Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators.* Alexandria, VA: Association for Supervision and Curriculumn Development.

Enders, C. (2010). *Applied missing data analysis.* New York, NY: Guilford Press.

Flay, B. R. (2010). *The Positive Action family program: A pilot randomized trial and replication.* Retrieved September 29, 2013, from www.positiveaction.net: https://www.positiveaction.net/research/downloads/family-article-2010.pdf

Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the Positive Action program on achievement and discipline: Two matched-control comparisons. *Prevention Science*, 71-89. doi:10.1023/A:1011591613728

Flay, B. R., Biglan, A., Boruch, R. F., Gonzàlez Castro, F., Gottfredson, D., Kellam, S., . . . Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 151-175. doi:10.1007/s11121-005-5553-y

Gottfredson, D. C., & Gottfredson, D. C. (2001). What schools do to prevent problem behaviors and promote safe environments. *Journal of Educational and Psychological Consultation*, 313-344. doi:10.1207/S1532768XJEPC1204_02

Greenberg, M. T. (2004). Current and future challenges in school-based prevention: The research perspective. *Prevention Science*, 5-13. doi:10.1023/B:PREV.0000013976.84939.55

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologists*, 466-474. doi:10.1037/0003-066X.58.6-7.466

Hawkins, J. D., Smith, B. H., & Catalano, R. F. (2004). Social development and social and emotional learning. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 135-150). New York, NY: Teachers College Press.

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement*, 617-637. doi:10.1177/0013164410382896

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, 491-525.

Ji, P., DuBois, D. L., & Flay, B. R. (in press). Social-emotional and character development scale: Development and initial validation with urban and elementary school students. *Journal of Research on Character Education*.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Golsin (Ed.), *Handbook of socialization theory and research* (pp. 347-480). Chicago, IL: Rand McNally.

Lewis, K. M., Bavarian, N., Snyder, F. J., Acock, A., Day, J., Dubois, D. L., . . . Flay, B. R. (2012). Direct and mediated effects of a social-emotional and character development program on adolescent substance use. *The International Journal of Emotional Education*, 56-78. Retrieved from http://www.academia.edu/2833668/Direct_and_Mediated_Effects_of_a_Social-Emotional_and_Character_Development_Program_on_Adolescent_Substance_Use

Lewis, K. M., DuBois, D. L., Bavarian, N., Acock, A., Silverthorn, N., Day, J., . . . Flay, B. R. (2013). Effects of Positive Action on the emotional health of urban youth: A cluster-randomized trial. *Journal of Adolescent Health*, 1-6. doi:10.1016/j.jadohealth.2013.06.012

Muthèn, L. K., & Muthèn, B. O. (2010). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthen & Muthen.

Osborne, J. W. (2013). *Best practices in data cleaning.* Thousand Oaks, CA: Sage.

Park, N. (2004). Character strengths and positive youth development. *The ANNALS of the American Academy of Political and Social Science*, 40-54. doi:10.1177/0002716203260079

Park, N., & Peterson, C. (2008). Positive psychology and character strengths: Application to strengths-based school counseling. *Professional School Counseling*, 85-92. doi:10.5330/PSC.n.2010-12.85

Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). *The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews.* Chicago, IL: Collaborative for Academic, Social, and Emotional Learning (CASEL). Retrieved from http://casel.org/publications/positive-impact-of-social-and-emotional-learning-for-kindergarten-to-eighth-grade-students-findings-from-three-scientific-reviews/

Person, A. E., Moiduddin, E., Hague-Angus, M., & Malone, L. M. (2009). *Survey of outcomes measurement in research on character education programs (NCEE 2009-006).* Institute of Education Sciences, U.S. Department of Education. Washington, DC: National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/pubs/2009006/

Schaps, E., Battistich, V., & Solomon, D. (2004). Community in school as key to student growth: Findings from the Child Development Project. In J. E. Zins, R. P. Weissberg, M. C.

Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 189-205). New York, NY: Teachers College Press.

Social and Character Development Research Consortium. (2010). *Efficacy of schoolwide programs to promote social and character development and reduce problem behavior in elementary school children: Report from the social and character development research program (NCER 2011-2001).* Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1997). Fostering educational resilience in inner-city schools. In M. C. Wang, H. Walberg, & H. J. Walberg (Eds.), *Children and youth* (pp. 119-140). Newbury Park, CA.

Washburn, I. J., Acock, A., Vuchinich, S., Snyder, F., Li, K.-K., Ji, P., . . . Flay, B. R. (2011). Effects of a social-emotional and character development program on the trajectory of behaviors associated with social-emotional and character development: Findings from three randomized trials. *Prevention Science*, 314-323. doi:10.1007/s11121-011-0230-9

Weissberg, R. P., & O'Brien, M. U. (2004). What works in school-based social and emotional learning programs for postive youth development. *Annals of the American Academy of Political and Social Science*, 86-97. doi:10.1177/0002716203260093

Wigelsworth, M., Humphrey, N., Kalambouka, A., & Lendrum, A. (2010). A review of key issues in the measurement of children's social and emotional skills. *Educational Psychology in Practice*, 173-186. doi:10.1080/02667361003768526

Wilson-Ahlstrom, A., Yohalem, N., DuBois, D., & Ji, P. (2011). *From soft skills to hard data: Measuring youth program outcomes.* The Forum for Youth Investment. Washington, DC: The Forum for Youth Investment. Retrieved from http://forumfyi.org/content/soft-skills-hard-data-

Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (2004). *Building academic success on social and emotional learning: What does the research say?* New York, NY: Teachers College Press.

APPENDIX B

EXTENDED METHODOLOGY OVERVIEW

Adopting Messick's unified construct-based model of validity, the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) refers to validity as the "degree to which evidence and theory support the interpretation of test scores entailed by proposed uses of tests" (p. 9). Messick (1995) outlines six aspects of construct validity which test developers should address: (a) *content*, which includes evidence of relevance, representativeness and technical quality; (b) *substantative*, which adresses the theoretical rationales underlying the responses; (c) *structural*, which examines the relationship between items and the targeted constructs; (d) *generalizablity*, which determines how well the instrument performs across varied populations, groups, settings and tasks; (e) *external*, to establish discriminant and convergent as well as criterian evidence; and (f) *consequential*, to appraise the implications of score interpretation as a basis for action. Establishing content, substantive, and consequential aspect evidence relies more on theory building with the support of substantive experts. Establishing structural, generalizability, and external aspects of validity depends on conducting a series of psychometric investigations in order to determine how well the item responses correspond to the underlying theoretical constructs (Dimitrov, 2010).

Few studies have used factor analysis to investigate the construct validity of instruments designed to measure social-emotional and character development skills (Humphrey et al., 2011). In order to establish construct validity in SECD measurement models, methodologies such as exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and exploratory structural equation modeling (ESEM) must be applied to current SECD measures. With few studies reporting evidence for factorial invariance across groups or time, it is important to continue the factorial investigations to include multi-group analysis. Furthermore, social-emotional character

development skills surveys offer limited evidence of concurrent criterian, convergent or discriminant validility (Person et al., 2009; Wilson-Ahlstrom et al., 2011; Durlak et al., 2011). In order to address these validity deficiencies, established measurement models of SECD scales must be extended through structural equation modeling (SEM) to include (a) canonical correlation analysis (CCA) relating the SECD construct to other related constructs, and (b) multi-trait multi-method analysis (MTMM) examining student self-reports against outside reports (e.g. teachers) across multiple traits.

Exploratory Factor Analysis

The purpose of exploratory factor analysis (EFA) is to discover the underlying structure of an instrument by determining the number of latent factors which most adequately capture the correlations among factor indicators. Factor indicators are observed item responses which are assumed to be representative of underlying continuous latent variable constructs. Unlike principal components analysis (PCA) which analyzes all the variance in a set of items, EFA only analyzes the covariance or reliable variance between items (Bryant & Yarnold, 1995). In EFA, no hypothesized structure (factor to item relationships) is designated, but instead factors are mathematicially derived. As such, an EFA is seen as a technique for theory development (Henson & Roberts, 2006). In order to conduct EFA, several analytical decisions such as how many factors to retain and which rotation techinique is most appropriate must be made. Ultimately, EFA depends on being able to adequately describe and explain the resulting retained factors.

Factor Retention

The goal of identifying the number of factors to retain involves adequately representing and explaining the underlying constructs, but avoiding retention of extraneous factors where trivial variance is being elevated in importance and assigned meaning (Hayton, Allen, & Scarpello, 2004). Retaining too few or too many factors has been shown to result in uninterpretable and non-replicable factors (Zwick & Velicer, 1986). Methods for determining the number of factors to retain include Bartlett's (1950, 1951) chi-square test, Kaiser's (1960) eigenvalue greater than 1 (K1) rule, and Cattell's (1966) visual inspection of scree plots to include more acceptable applications such as Horn's (1965) parallel analysis (PA) and Velicer's (1976) minimum average partial method (MAP). Zwick and Velicer's (1986) review of these factor retention methods determined the PA and MAP methods performed best across varying conditions, while the often used K1 method overestimates the number of factors to retain and the sample-size-sensitive Bartlett's test is more variable and less accurate than visual inspection of scree plots.

Factor Rotation

The purpose of factor rotation is to rotate the factor axis in an attempt to simplify the factor solution where factor patterns are more closely clustered and, therefore, easier to interpret. The goal is to achieve "cleaner" factors with high factor pattern coefficients on a single factor and lower, near-zero loadings on other factors, thus distinguishing distinct latent constructs. Numerous rotation strategies exist but can be grouped into two main categories: (a) orthogonal rotations in which factors are not allowed to correlate, and (b) oblique rotations where factors are allowed to correlate in representation of indicators being related to more than one factor (Crocker

& Algina, 2008). Decisions on which type of rotation strategy to use partially depends on whether the underlying structure is theorized to be simple (uncorrelated factors) or complex (correlated factors).

When deciding which items will be attributed to which factors, Schmitt and Sass (2011) and Cudeck and O'Dell (1994) emphasize the importance of estimating standard errors of pattern coefficients since the choice of rotation criteria stands to create bias in either the interfactor correlations or the factor pattern coefficients. For example, in the case of orthogonal rotations with uncorrelated factors, item cross-loadings could be inflated to account for the variance explained. In a case of oblique rotations with correlated factors, the cross-loadings could be biased downward since the correlation between factors could be arbitrarily assigned to explain the variance. In other words, different rotation criteria can produce different factor pattern matrices which, in turn, result in different data-derived interpretations of the underlying construct structure. As such, in the case of complex factor structures with oblique rotation, interpretation of factor structure coefficients (the correlation between indicators and non-assigned factors) in addition to the factor pattern coefficients should be performed (Henson & Roberts, 2006).

Confirmatory Factor Analysis

The purpose of confirmatory factor analysis (CFA) is to determine if data obtained from a measure fit an a'priori determined factor structure. In CFA, a theoretical structure is superimposed on the data, then a series of fit indices are consulted to determine how well the predetermined factor structure reproduces the data's variance-covariance matrix. In contrast to exploratory factor analysis, CFA acts as a tool for theory testing – or confirmation (Bryant & Yarnold, 1995; Tabachnick & Fidell, 2007).

Model Fit

How well the data fits a specified model is determined by examining the difference between the theorized inter-item relationships and actual inter-item relationships from the data. The resulting differences between the predicted and observed inter-item relationships are referred to as fitted residuals. Standardardized residuals are obtained by dividing the fitted residuals by associated standard errors (Bryant & Yarnold, 1995). Goodness of model fit is determined by evaluating the size of the fitted residuals where smaller residuals results in better fit. Most common software programs (e.g., MPlus, Lisrel) provide a chi-square statistic along with several fit indices. Fit indices calculated to summarize the goodness of fit are organized into three categories: (a) absolute fit indices where a fit of zero indicates the best fit, (b) comparative (incremental) fit indices for comparing models where zero indicates worst fit and one indicates best fit, and (c) parsimony fit indices which correct for the number of varibles included in the model (Hooper, Coughlan, & Mullen, 2008; Kenny, 2012).

An overall maximum likelihood chi-square (χ^2, Likelihood Ratio Test) statistic of zero indicates the model is a perfect fit with no difference between the sample covariance matrix and the reproduced implied matrix (Hu & Bentler, 1999). The χ^2 associated *p*-value indicates whether the null hypothesis of the residual matrix being equal to zero has been confirmed (Bryant & Yarnold, 1995). Failure to reject the null hypothesis indicates a good fitting model where the CFA model reproduces the observed relationships (Mulaik et al., 1989). Due to sensitivity to sample size and strength of correlations, many researchers suggest limited use of chi-square null hypothesis testing for assessing a model's exact fit but still advocate evaluating the change in chi square when comparing alternative models (Bentler & Bonett, 1980; MacCallum, Browne, & Cai, 2006; Anderson & Gerbing, 1988).

Root mean square wrror of approximation (RMSEA; Steiger, 1990) assesses how well a model with unknown but optimal parameters would fit the population covariance matrix if it were available (Byrne, 1998). One benefit of RMSEA is confidence intervals can be calculated to understand the role of sampling error. Hu and Bentler (1999) proposed a RMSEA cutoff value of .06 as indication of good model fit between proposed and observed model fit; however, they caution against use of RMSEA with small sample sizes ($N \leq 250$). Simulation studies have illustrated the RMSEA cutoff point to be highly dependent on model specification, degrees of freedom, and sample size. Because of these dependencies RMSEA should only be used in conjunction with the chi square and other fit statistics (Chen, Curran, Bollen, Kirby, & Paxton, 2008; Kenny & McCoach, 2009; Fan & Sivo, 2007).

Two popular comparative fit indices are the Tucker Lewis Index (TLI) and Comparative Fit Index (CFI). Both TLI and CFI provide the ratio of change in chi-square between the null model (a model where all variables are allowed to vary but are uncorrelated) and a proposed theoretical model where 0.90 – 0.95 is considered marginal fit and above 0.95 is considered good fit. TLI compares the ratio of chi square and degrees of freedom while CFI subtracts the degrees of freedom from the chi-square. Both calculations award parsimony, with greater numbers of estimated parameters resulting in a lowered index. Since CFI and TLI are closely correlated, Kenny (2012) suggests only one be reported.

Several issues surrounding assessing model fit through indices and chi-square difference testing remain somewhat controversial. Barrett (2007) suggests interpreting only the chi-square when comparing model fit among alternative models, especially in light of the tendency to make non-theoretical model adjustment in order to achieve fit statistics which fall within the acceptable range. However, considering chi square is highly sensitive to sample size, Barrett

also concludes any model cannot be scientifically distinguised from competing models if no theory-relevant-criteria can be dertermined to support a substantative advantage of one model over another. Kenny and McCoach (2009) illustrate the effect of the number of variables on model fit and argue against using fit indices in models with small degrees of freedom. Arguing model fit indices allow misspecification among latent constructs to be masked, O'Boyle and Williams (2011) have proposed a root mean square error of approximation of the path component (RMSEA-P). Because of issues such as these and over-generalization of model fit indices' cutoff criteria across all model and sample types, current studies suggest application of fit indices are better suited at determining misspecified models when used to compare alternative models as opposed to a single model in isolation (Marsh et al., 2004; Chen, 2007; Williams & O'Boyle, 2011).

Alternative Models

In lieu of depending on arbitrary cutoff points to fit indices, CFA usually involves testing a hypothesized model along with other alternative models. In a CFA independent clusters model (CFA-ICM), each item is regressed on only one factor and all other factor-to-item loadings are assumed to be zero (Marsh, Ludtke, Nagengast, & Morin, 2013). One alternative to restrictive ICM allows the factors to correlate. Depending on the theoretical underpinnings, first order factors (factors with direct relationships with observed variables) may be viewed as mediating factors after the addition of a higher second order factor. Since fit of the CFA depends on correctly specifying the model, the fit of hypothesized and alternate models are compared to see which better reproduces the variance in the assessed data (Byrne, 1998; Marsh et al., 2004). When comparing the fit of nested models imposing differing numbers of invariance constraints

with an adequate sample size ($N > 300$), Chen (2007) suggests less than 0.01 change in incremental model fit indices (e.g., CFI) and a REMSEA increase of less than 0.015 supports retaining the more parsimonious model. However, Chen cautions against overgeneralization of these suggestions since sample size and model complexity can affect the magnitude of changes in fit statistics.

Exploratory Structural Equation Modeling

An integration of EFA, CFA and SEM, exploratory structural equation modeling (ESEM) was developed to help aleviate commonly encountered CFA problems associated with goodness of fit, differentiation of factors, measurement invariance across time or groups and differential item functioning (Asparouhov & Muthèn, 2009; Marsh et al., 2009, 2010). As such, instead of associating each item with only one factor and constraining all other non-target loadings to zero as is typical in the highly restrictive independent clusters model (ICM), ESEM allows for less restrictive models in which all factor loadings are estimated and where items are free to cross-load on other factors within the same set of factors (Marsh et al., 2011; Asparouhov & Muthèn, 2009). Instead of calculating structure coefficients in a separate analysis as Thompson (1997) demonstrates, ESEM includes the structure coefficient parameter estimation along with the standard errors for the structure coefficients. ESEM retains the capability of rotating factors and also comparing model fit through comparing model fit statistics. Asparouhov and Muthen (2009) demonstrate fitting ESEM models using change in RMSEA (ΔRMSEA) and change in CFI and TFI (ΔCFI; ΔTFI), in addition to the Satorra-Bentler change in chi-square.

Multi-Group Analysis

Multi-group factorial invariance and time invariance can be examined under SEM and ESEM frameworks as outlined by Dimitrov (2010) and Marsh et al. (2009, 2010). The purpose of multi-group analysis is to determine if item responses from identified groups follow similar response patterns as related to the underlying theoretical structure of the instrument. Time invariance analysis determines the similarity of response patterns across different waves of data collection. Testing factorial and time invariance follows a sequential constraint imposition procedure comparing a set of partially nested models ranging from the least restrictive model with no parameters constrained to be invariant, to a model with complete factorial invariance where all parameters are constrained to be invariant (Marsh et al., 2011; Dimitrov, 2010). This forward approach to testing factorial invariance provides for examing configural, measurement and structural invariance. Table 4 provides the taxonomy of the multiple-group exploratory structural equation models (MGESEM). Again, the Satorra-Bentler scaled chi-square difference ($\Delta\chi^2$) and change in CFI (ΔCFI < -.01) can be used to compare models (Dimitrov, 2010).

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing.* Washington, DC: American Educational Research Association.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 411-423. doi:10.1037/0033-2909.103.3.411

Asparouhov, T., & Muthen, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 397-438. doi:10.1080/10705510903008204

Barrett, P. (2007). Structural equation modeling: adjudging model fit. *Personality and Individual Differences*, 815-824. doi:10.1016/j.paid.2006.09.018

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology*, 77-85. doi:10.1111/j.2044-8317.1950.tb00285.x

Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. *British Journal of Psychology*, 1-2. doi:10.1111/j.2044-8317.1951.tb00299.x

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 588-606. doi:10.1037/0033-2909.88.3.588

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm, & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). Washington, DC: American Psychological Association.

Byrne, B. M. (1998). *Structural equation modeling with Lisrel, Prelis, and Simplis: Basic concepts, applications, and programming.* Mahwah, NJ: Lawrence Earlbaum.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 245-276. doi:10.1207/s15327906mbr0102_10

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 464-504. doi:10.1080/10705510701301834

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods and Research*, 462-494. doi:10.1177/0049124108314720

Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory.* Mason, OH: Cengage Learning.

Cudeck, R., & O'Dell, L. L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for factor loadings and correlations. *Psychological Bulletin*, 475-487. doi:10.1037/0033-2909.115.3.475

Dimitrov, D. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 121-149. doi:10.1177/0748175610373459

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American journal of community psychology*, 294-309.

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, 509-529. doi:10.1080/00273170701382864

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 191-205. doi:10.1177/1094428104263675

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, 393-416. doi:10.1177/0013164405282485

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 53-60. Retrieved from http://arrow.dit.ie/cgi/viewcontent.cgi?article=1001&context=buschmanart&sei-redir=1&referer=http%3A%2F%2Fscholar.google.com%2Fscholar_url%3Fhl%3Den%26q%3Dhttp%3A%2F%2Farrow.dit.ie%2Fcgi%2Fviewcontent.cgi%253Farticle%253D1001%2526context%253Dbuschmanart%2

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 179-185. doi:10.1007/BF02289447

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 1-55. doi:10.1080/10705519909540118

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 1-55.

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement*, 617-637. doi:10.1177/0013164410382896

Kenny, D. A. (2012, July 5). *Measuring model fit*. Retrieved October 20, 2013, from www.davidakenny.net: http://davidakenny.net/cm/fit.htm

Kenny, D. A., & McCoach, B. (2009). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 333-351. doi:10.1207/S15328007SEM1003_1

MacCallum, R. C., Browne, M. W., & Cai, L. (2006). Testing differences between nested covariance structure models: Power analysis and null hypothesis. *Psychological Methods*, 19-35. doi:10.1037/1082-989X.11.1.19

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, 320-341. doi:10.1207/s15328007sem1103_2

Marsh, H. W., Ludtke, O., Muthen, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big-five factor structure through exploratory structural equation modeling. *Psychological Assessment*, 471-491. doi:10.1037/a001922

Marsh, H. W., Ludtke, O., Nagengast, B., & Morin, A. J. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right - camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, 257-284. doi:10.1037/a0032773

Marsh, H. W., Muthen, B., Asparouhov, T., Ludtke, O., Robitzsch, A., Morin, A., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Applications to students' evaluations of university teaching. *Structural Equation Modeling*, 439-476. doi:10.1080/10705510903008220

Marsh, H. W., Nagengast, B., Morin, A. J., Parada, R. H., Graven, R. G., Craven, R. G., & Hamilton , L. R. (2011). Construct validity of the multidimensional structure of bullying and victimization: An application of exploratory structural equation modeling. *Journal of Educational Psychology*, 701-732. doi:10.1037/a0024122

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 741-749.

Mulaik, S. A., James, L. R., VanAlstine, J., Bennet, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 430-445. doi:10.1037/0033-2909.105.3.430

O'Boyle, E. H., & Williams, L. J. (2011). Decomposing model fit: measurement vs. theory in organizational research using latent variables. *Journal of Applied Psychology*, 1-12. doi:10.1037/a0020539

Person, A. E., Moiduddin, E., Hague-Angus, M., & Malone, L. M. (2009). *Survey of outcomes measurement in research on character education programs (NCEE 2009-006).* Institute of Education Sciences, U.S. Department of Education. Washington, DC: National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/pubs/2009006/

Schmitt, T. A., & Sass, D. A. (2011). Rotation criteria and hypothesis testing for exploratory factor analysis: Implications for factor pattern loadings and interfactor correlations. *Educational and Psychological Measurement*, 95-113. doi:10.1177/0013164410387348

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). New York, NY: Allyn and Bacon.

Thompson, B. (1997). The importance of structure coefficients in structural equation modeling confirmatory factor analysis. *Educational and Psychological Measurement*, 5-19. doi:10.1177/0013164497057001001

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 321-327. doi:10.1007/BF02293557

Wilson-Ahlstrom, A., Yohalem, N., DuBois, D., & Ji, P. (2011). *From soft skills to hard data: Measuring youth program outcomes.* The Forum for Youth Investment. Washington, DC: The Forum for Youth Investment. Retrieved from http://forumfyi.org/content/soft-skills-hard-data-

Zwick, W. R., & Velicer, W. F. (1986). Comparisons of five rules for determining the number of components to retain. *Psychological Bulletin*, 432-442. doi:10.1037/0033-2909.99.3.432

APPENDIX C

EXTENDED RESULTS

Extended Results from Phase I

After establishing the best fitting SECDS factor structure, internal reliability of the scales for the six SECD scales were evaluated for both waves and separately for gender and ethnic groups. Test-retest reliability was examined by correlating the overall SECD as well as the six composites at Time One and Time Two. Concurrent validity was be assessed by determining the extent to which the SECDS correlated with the six external measures previously described and hypothesized to relate to SECD.

Reliability estimates shown in Table 1 indicate acceptable internal consistencies for the total group on all six factors for Time 1 and Time 2 (α = .643 to .818). When comparing reliabilities across gender and ethnic groups the Maya, Garifuna, and Undesignated groups have lower reliabilities for Self-Control ($\alpha < .6$), and the Maya group has a lower reliability for the Respect for Parent factor (α = .598). Beyond these exceptions, internal consistencies are within acceptable range (α = .6 - .9) for all groups across all six factors and generally increase from Time 1 to Time 2.

Table C.1
SECDS six factor internal consistencies for pretest and posttest with subsets for total group, gender, and ethnicity.

Scale	Self-Cont (4 items)	Prosocial (7 items)	Respect Teacher (5 items)	Respect Parent (4 items)	Honesty (5 items)	Self-Develop (4 items)
Time 1						
All	0.643	0.765	0.757	0.662	0.662	0.681
Male	0.630	0.752	0.738	0.659	0.667	0.683
Female	0.624	0.749	0.754	0.662	0.632	0.657
Undesignated*	0.665	0.811	0.807	0.655	0.689	0.636
Creole	0.627	0.747	0.749	0.636	0.649	0.674
Garifuna	0.522	0.769	0.649	0.660	0.606	0.744
Maya	0.582	0.795	0.779	0.598	0.607	0.641
Metizo	0.655	0.786	0.793	0.735	0.713	0.703
Other	0.695	0.772	0.678	0.562	0.627	0.581
Undesignated*	0.638	0.804	0.808	0.652	0.661	0.646
Time 2						
All	0.653	0.784	0.818	0.735	0.724	0.719
Male	0.622	0.772	0.801	0.732	0.699	0.715

Female	0.675	0.775	0.825	0.745	0.727	0.686
Undesignated*	0.598	0.804	0.757	0.680	0.761	0.788
Creole	0.629	0.764	0.817	0.721	0.706	0.687
Garifuna	0.587	0.796	0.760	0.745	0.737	0.782
Maya	0.740	0.783	0.842	0.771	0.740	0.782
Metizo	0.780	0.823	0.842	0.778	0.756	0.756
Other	0.693	0.760	0.821	0.706	0.689	0.634
Undesignated*	0.580	0.778	0.758	0.679	0.756	0.775

*Those missing the gender or ethnicity indicators; *n* indicates sample size using listwise deletion.

Test-retest reliabilities for the six SECDS factors ranged from .302 to .591 with the highest average correlations on the Respect Teacher factor (ravg= .470) and lowest on the Respect Parent factor (ravg=.394) (Table 2). Those missing demographic information and therefore considered Undesignated (n<100) exhibited the lowest average reliabilities across the six factors (Gender ravg=..398; Ethnicity ravg=.377). No other patterns in test-retest reliabilities across groups were noted.

Table C.2
SECDS six factor summed score test-retest reliabilies for Time 1 and Time 2 with subsets for total group, gender, and ethnicity.

	Scale	Self-Cont	Prosocial	Respect Teacher	Respect Parent	Honesty	Self-Develop	Average
	All	0.440	0.446	0.491	0.390	0.421	0.440	**0.424**
Gender	Male	0.439	0.411	0.456	0.337	0.387	0.414	**0.407**
Gender	Female	0.431	0.430	0.505	0.445	0.423	0.427	**0.444**
Gender	Undesignated*	0.365	0.423	0.409	0.357	0.400	0.434	**0.398**
Ethnicities	Creole	0.421	0.411	0.470	0.337	0.413	0.415	**0.411**
Ethnicities	Garifuna	0.442	0.518	0.302	0.474	0.459	0.399	**0.432**
Ethnicities	Maya	0.436	0.491	0.506	0.426	0.326	0.277	**0.410**
Ethnicities	Metizo	0.506	0.497	0.591	0.504	0.443	0.517	**0.510**
Ethnicities	Other	0.383	0.487	0.569	0.305	0.433	0.489	**0.444**
Ethnicities	Undesignated*	0.339	0.387	0.403	0.360	0.374	0.400	**0.377**
	Average	**0.420**	**0.450**	**0.470**	**0.394**	**0.408**	**0.421**	

*Those missing the gender or ethnicity indicators; sample using listwise deletion for missing.

Correlations between the SECDS factors and concurrent measures are shown in Table 3. For both Time 1 and Time 2, the SECDS factors statistically significantly correlated with all concurrent measures with the exception of Anxiety (r = -.044 to .075). For all six SECDS

factors, the highest concurrent correlation was with Moral Center (r = .291 to .437). As expected, the SECDS factors were all negatively correlated with Risk Behaviors (r = -.198 to -293). Correlations in Table 4 indicate a similar pattern of concurrent correlations for both male and females.

TableC. 3
Correlations between SECDS factors and concurrent measures.

	Self-Control	Prosocial	Respect Teacher	Respect Parent	Honesty	Self-Develop	SECDS Total
Time 1							
Reward for Prosocial Behavior	.234	.279	.206	.177	.228	.217	.300
Anxiety	*-.042*	*.010*	-.063	-.062	*-.023*	*.033*	*-.019*
Participation in Neighborhood	.168	.192	.083	.138	.144	.105	.197
Peer Group Affiliation	.293	.294	.270	.169	.261	.222	.342
Risk Behaviors	-.243	-.222	-.285	-.254	-.215	-.198	-.289
Moral Center	.345	.327	.365	.305	.328	.291	.436
Time 2							
Reward for Prosocial Behavior	.212	.270	.244	.228	.277	.210	.303
Anxiety	*-.029*	.075	*-.009*	*-.044*	*.011*	.057	*.016*
Participation in Neighborhood	.191	.216	.127	.155	.176	.106	.209
Peer Group Affiliation	.269	.311	.287	.209	.296	.181	.334
Risk Behaviors	-.292	-.206	-.293	-.194	-.200	-.138	-.281
Moral Center	.414	.395	.437	.321	.391	.333	.485

Note. Correlations shown in italics are NOT statistically significant. For all other correlations p <.01.

Table C.4
Correlations between SECDS factors and concurrent measures across GENDER.

	Self-Control	Prosocial	Respect Teacher	Respect Parent	Honesty	Self-Develop	SECDS Total
MALE							
Reward for Prosocial Behavior	.194	.232	.245	.186	.228	.198	.273
Anxiety	*-.015*	*.095*	-.044	-.021	*.030*	*.048*	*.023*
Participation in Neighborhood	.145	.211	.144	.159	.149	.122	.206
Peer Group Affiliation	.262	.316	.285	.210	.288	.192	.336
Risk Behaviors	-.259	-.301	-.164	-.271	-.178	-.173	-.135
Moral Center	.390	.345	.418	.357	.374	.301	.465
FEMALE							
Reward for Prosocial Behavior	.225	.260	.180	.152	.222	.208	.273
Anxiety	*-.084*	-.075	*-.067*	*-.079*	*-.066*	*-.031*	-.092
Participation in Neighborhood	.183	.200	.105	.152	.148	.117	.201
Peer Group Affiliation	.282	.304	.234	.126	.244	.200	.312
Risk Behaviors	-.255	-.246	-.297	-.257	-.250	-.153	-.324
Moral Center	.323	.318	.334	.265	.330	.275	.400

Correlations shown in italics are NOT statistically significant. For all other correlations p <.01.

Population Invariance

With few studies reporting evidence for factorial invariance across populations, it is important to continue the factorial investigations to include cross-sample multi-group analysis. Since item level data (which is currently required for conducting ESEM in MPlus) for the comparative Ji et al. (2013) U.S. sample was not available, the preliminary multigroup analysis was conducted under the ICM-CFA framework where no items were allowed to crossload on non-target factors. This preliminary multigroup analysis comparing the Belize sample and the U.S. sample were conducted using Lisrel 9.1 where correlations matrices are allowed as the point of data entry.

Table C.4 includes model fit indices for multigroup comparisons across populations. Evaluation of model fit indices on initial CFA models ran separately on the Belize and U.S. samples separately indicate the data to be a moderately good fit for data from both groups (RMSEA: U.S. = .053 and Belize = .061; CFI: U.S. = .970 and Belize = .967). In addition, the unconstrained model where both groups are included in one model but no parameters are constrained to be equal across groups also indicates acceptable fit (RMSEA = .059; CFI = .968). However, the constrained model where both groups are included and all parameters are constrained equal across groups indicates a slightly lesser model fit (RMSEA = .064; CFI = .959). While the constrained model fit indices still fall within acceptable range, it is interesting the RMSEA confidence intervals do not overlap which would seem to indicate at least some degree of worsening of model fit in the constrained model.

Table C.5 includes the pattern coefficients, structural coefficients and factor correlations for both the Belize and U.S. samples. The pattern coefficients seem to follow a similar pattern across groups where all indicators show high coefficients on target factors but also high structure

coefficients for non-target factors. However in several instances the Belize target coefficients are elevated (PSIUS*Belize = -.298). For example on the Respect Teacher factor, the pattern coefficients for the U.S. sample range from .323 to .508 where for the Belize sample the range is between .609 and .735. Similar elevations can be seen on all factors. In opposition, the factor correlations for the U.S. sample are slightly elevated in comparison to the Belize population. For example, there is a .285 difference between the factor correlation for Self-Control and Self-Development between the U.S. (r = .946) and Belize (r = .661). As would be expected considering the high structure coefficients (rs = 171 to .603), factor correlations for both groups are elevated ($r_{U.S.}$ = .717 to .951; r_{Belize} = .651 to .902). All of the noted differences indicate the need for futher invariance evaluation under the ESEM framework.

The preliminary investigation of the U.S. versus Belize sample ICM-CFA multigroup testing indicates there is at least some variance in the parameters of the data for the two samples. Further multi-group evaluation under the ESEM framework would be able to determine if these differences are due in part to misspecification of the factor structure under the ICM-CFA framework. . Initial multigroup evaluation of the ICM-CFA SECDS measurement model comparing the Belize sample and the U.S. sample indicate a need for further evaluation.

Table C.4
Multigroup CFA fit indices for Ji et al. data versus Belize data.

Model	$\chi 2$	df	RMSEA	RMSEA CI	CFI
JI & FLAY	964.09	362	0.053	(.049; .057)	0.97
BELIZE	2886.1	362	0.061	(.059; .063)	0.967
UNCONSTRAINED	3848.586	724	0.059	(.057;.061)	0.968
CONSTRAINED	4766.65	797	0.064	(.062 ; .065)	0.959
DIFFERENCE	918.064	73			

Note. Preliminary mulitgroup analysis conducted in Lisrel 9.10

Table C.5

CFA factor loadings and structure coefficients using Ji et al. data vs. Belize data.

	Ji et al. Data						Belize Data					
ITEM	SELFCON	PROSOC	RTEACH	RPARENT	HONESTY	SELFDEV	SELFCON	PROSOC	RTEACH	RPARENT	HONESTY	SELFDEV
PRER51	0.512	0.471	0.477	0.378	0.487	0.484	0.598	0.449	0.477	0.378	0.487	0.484
PRER56	0.567	0.521	0.528	0.419	0.539	0.536	0.516	0.497	0.528	0.419	0.539	0.536
PRER63	0.522	0.480	0.487	0.386	0.496	0.494	0.694	0.457	0.487	0.386	0.496	0.494
PRER73	0.465	0.427	0.433	0.344	0.442	0.440	0.515	0.407	0.433	0.344	0.442	0.440
PRER55	0.362	0.394	0.314	0.362	0.362	0.362	0.541	0.589	0.314	0.362	0.362	0.362
PRER60	0.480	0.522	0.417	0.374	0.472	0.447	0.585	0.637	0.390	0.338	0.476	0.413
PRER62	0.438	0.477	0.381	0.342	0.432	0.408	0.524	0.570	0.357	0.309	0.435	0.377
PRER65	0.457	0.497	0.397	0.356	0.450	0.425	0.603	0.656	0.372	0.322	0.453	0.393
PRER67	0.474	0.516	0.412	0.370	0.467	0.442	0.538	0.585	0.386	0.334	0.470	0.408
PRER68	0.470	0.511	0.408	0.366	0.462	0.437	0.562	0.612	0.382	0.331	0.466	0.404
PRER72	0.527	0.573	0.457	0.411	0.519	0.490	0.578	0.629	0.429	0.371	0.522	0.453
PRER53	0.301	0.258	0.323	0.254	0.276	0.280	0.273	0.242	0.609	0.237	0.233	0.225
PRER54	0.392	0.336	0.421	0.331	0.360	0.365	0.355	0.315	0.731	0.309	0.304	0.293
PRER66	0.399	0.342	0.428	0.337	0.366	0.371	0.361	0.320	0.735	0.314	0.309	0.298
PRER74	0.473	0.405	0.508	0.400	0.435	0.440	0.429	0.380	0.645	0.372	0.366	0.354
PRER75	0.425	0.364	0.456	0.359	0.390	0.395	0.385	0.341	0.698	0.334	0.329	0.317
PRER58	0.202	0.196	0.215	0.273	0.224	0.241	0.178	0.177	0.200	0.640	0.187	0.171
PRER59	0.250	0.242	0.266	0.338	0.277	0.298	0.221	0.219	0.248	0.708	0.232	0.212
PRER70	0.316	0.307	0.337	0.428	0.351	0.378	0.279	0.277	0.314	0.622	0.293	0.268
PRER71	0.341	0.331	0.364	0.462	0.378	0.408	0.302	0.299	0.339	0.671	0.316	0.289
PRER52	0.399	0.380	0.360	0.344	0.420	0.353	0.379	0.383	0.303	0.288	0.604	0.273
PRER57	0.443	0.422	0.399	0.382	0.466	0.392	0.420	0.425	0.336	0.319	0.583	0.303
PRER61	0.541	0.515	0.487	0.466	0.569	0.479	0.513	0.518	0.410	0.390	0.639	0.370
PRER64	0.512	0.487	0.461	0.441	0.538	0.452	0.485	0.490	0.388	0.369	0.508	0.350
PRER69	0.444	0.423	0.400	0.382	0.467	0.393	0.421	0.425	0.337	0.320	0.578	0.304
PRER76	0.375	0.339	0.343	0.350	0.333	0.396	0.262	0.313	0.276	0.248	0.258	0.695
PRER77	0.547	0.495	0.501	0.510	0.486	0.578	0.382	0.457	0.402	0.362	0.376	0.665
PRER78	0.420	0.380	0.385	0.392	0.373	0.444	0.293	0.351	0.309	0.278	0.289	0.588
PRER79	0.362	0.328	0.332	0.338	0.322	0.383	0.253	0.303	0.267	0.240	0.249	0.715
Factor Correlations												
SELFCON	1						1					
PROSOC	0.919	1					0.876	1				
RTEACH	0.932	0.798	1				0.844	0.748	1			
RPARENT	0.739	0.717	0.787	1			0.653	0.648	0.733	1		
HONESTY	0.951	0.905	0.856	0.819	1		0.902	0.911	0.721	0.685	1	
SELFDEV	0.946	0.856	0.867	0.883	0.841	1	0.661	0.791	0.696	0.626	0.651	1

Convergent Validity

Social-emotional character development skills surveys offer limited evidence of concurrent criterian, convergent or discriminant validility (Person et al., 2009; Wilson-Ahlstrom et al., 2011; Durlak et al., 2011). In order to address these validity deficiencies, established measurement models of SECD scales must be extended to include structural equation modeling (SEM) models relating the SECDS constructs to other related constructs. In order to extend the external validity of SECDS, criteria evidence was evaluated in an expansion of the final measurement model. This new model relates the SECDS factors to the six independent measures proposed to be associated with social-emotional skills and character compentencies (Messick, 1995). This model included 12 latent factors: six SECD facets and a latent factor for each of the six external instrument variables. Appendix D includes a path diagram of the proposed concurrent validity SEM model.

In order to further evaluate the ESEM SECDS model and how it compares to associated external measures, a SEM model was conducted including the six SECDS latent factors and six related measures which were thought to also be useful in evaluating interventions. The purpose of evaluating the larger SEM model is to determine how similar or different the SECDS factors are to other social emotional and character development related constructs. The six external comparative factors include Reward for Prosocial Behavior, Anxiety, Participation in Neighborhood Context, Peer Group Affiliation, Risk Behaviors, and Moral Center (Arthur et al., 2000; CDC, 2004; Chipuer et al., 1999; Elliot et al., 1996; Reynolds & Kamphaus, 2002). Since the comparative measures included only single constructs taken from larger instruments, all constructs were included as an ICM-CFA measurement model and no cross-factor loadings were considered. Preliminary evaluation of the pattern coefficients for all the related constructs

indicated a relatively well defined measurement model with target loadings ranging from .284 to .846.

As an additional point of illustration of the utility of ESEM measurement models over CFA measurement models, the larger external relations SEM model was conducted with the inclusion of the ICM-CFA SECDS measurement model and the ESEM SECDS measurement model. Table C.6 compared the model fit of the two competing SEM models with the ICM-CFA included model being nested with in the ESEM included model. Although the ESEM included model indicates a very slightly improved fit, evaluation of the CFI and TLI in both instances indicates a less than optimal fit. Further investigation into a better fitted model fit goes beyond the scope of this paper; however, preliminary comparisons between correlations resulting from the ICM-CFA and ESEM included models were examined. As the correlations in in Table C.7 show, correlations between the SECDS factors and the six external measures under the ICM-CFA and ESEM frameworks are similar in pattern (PSI = .928). In addition, both the ICM-CFA and ESEM based models appear to reflect the expected substantive relationships. For example, Risk Behaviors and Anxiety are for the most part negatively associated with the SECDS factors (the exception being the Self-Development factor under the ESEM-based model). However, evaluation of the relationship magnitude between the associated measures shows somewhat larger coefficients for the model utilizing the ICM-CFA measurement model. The larger ICM-CFA interfactor correlations could be a interpreted as spurious inflation due to a function of misspecifying the SECDS measurement model through omitting non-target cross loadings. Further investigation examining the relationship between the ICM-CFA and ESEM included models is warranted.

Initial attempt to establish concurrent validity evidence provides a basis for further inquiry concerning the comparison of the SECDS factors to other related measures. In this study, a comparison of SEM models utilizing the ICM-CFA SECDS measurement model versus the ESEM SECDS measurement model highlighted the potential for upwardly biased correlations between the SECDS factors and other related factors based on misspecification of the measurement model. Although the pattern of correlations between factors was consistent across the ICM-CFA and ESEM utilized models, the magnitude for the correlations generated in the ICM-CFA are inflated.

Table C.6

Model fit comparing ESEM and CFA SEM models correlating SECDS factors with concurrent measures.

Model	χ^2	df	CFI	TLI	RMSEA	RMSEA CI	MD$\Delta\chi^2$	$df_{\Delta\chi 2}$	$p_{\Delta\chi 2}$	ΔCFI	ΔTLI	ΔRMSEA
ESEM FACTORS	9772.182	2821	0.871	0.859	0.036	[.036, .037]	-	-	-	-	-	-
CFA FACTORS	10244.002	2936	0.864	0.857	0.036	[.036, .037]	860.211	115	<.001	-0.007	-0.002	<.001

Table C.7

Correlations between SECDS factors and concurrent measures: ICM-CFA and ESEM.

	Reward Prosocial Behavior	Anxiety	Participation in Neighborhood	Peer Group Affiliation	Risk Behaviors	Moral Center
ESEM FACTORS						
Self-Control	.281	-.*053*	.249	.373	-.422	.493
Prosocial	.171	-.151	-.*038*	.236	-.347	.411
Respect Teacher	.257	-.*037*	.235	.505	-.343	.252
Respect Parent	.180	-.100	.191	.148	-.370	.323
Honesty	.322	-.*001*	.180	.247	-.146	.280
Self-Develop	.256	.133	.087	.168	-.186	.331
Mean Coefficient	.245	-.035	.151	.280	-.302	.348
CFA FACTORS						
Self-Control	.350	-.085	.251	.459	-.450	.542
Prosocial	.377	.*013*	.280	.437	-.366	.462
Respect Teacher	.294	-.107	.116	.380	-.463	.488
Respect Parent	.261	-.098	.211	.247	-.424	.400
Honesty	.340	-.*071*	.213	.384	-.380	.456
Self-Develop	.322	.*037*	.153	.338	-.359	.423
Mean Coefficient	.324	-.052	.204	.374	-.407	.462

Note. STDXY standardized correlations from Mplus output.

Concurrent Validity Measures

Reward for Pro-Social Behavior

In addition to the pro-social behavior component of the SECDS measure and the moral beliefs items, a separate scale targeted student perceptions of rewards for pro-social behavior (Arthur, Hawkins, Catalano, & Pollard, 2000). Items ask students to indicate frequency of parents and teachers noticing and rewarding student pro-social behaviors. The question stem was, "How much of the time do your PARENTS and TEACHERS notice when you…." Items were rated on a 4-point scale (NEVER, SOME TIMES, USUALLY, ALWAYS) with four indicating higher perception of social-emotional learning skills.

Anxiety

Anxiety was measured using items from the Behavior Assessment for Children (BASC) scale (Reynolds & Kamphaus, 2002). The BASC includes multiple methods for assessing behavior in children, but for the present study only the self-report items related to anxiety were included. The included items discuss common symptoms of anxiety for school aged children to determine if a student shows evidence of anxiety behavior (Reynold & Kamphaus, 2002; Ji et al., *2013*). The item stem was, "Tell us how you have felt during the last month." Statements were presented regarding student feelings about life in general and the participant was asked to mark YES, SOME TIMES, or NO for each statement. Example items include, "I often worry about something bad happening to me" and "I worry but I don't know why." Items were rated on a 3-point scale, with 3 indicating higher social-emotional skills and character.

Participation in Neighborhood Context

Student perception of their neighborhood context was assessed since the *Positive Action* program included a community component. Items from the *Neighborhood Youth Inventory* (Chipuer et al., 1999) target student experiences in their neighborhood, which is defined as "the streets and places around your home where you see people you know and do everyday things like visit and play with your friends." The instrument was developed as a measure of extra-individual characteristics of the environment to assess perceptions of the community as a whole and not individual people (Chipuer et al., 1999). The items were created from statements in interviews with youth related to their perspectives and opinions about activity in their neighborhood (Chipeur et al., 1999). Items ask students how much of the time people help others or drug dealing takes place in their neighborhood setting. Items were rated on a 4-point scale (NONE of the time, SOME of the time, MOST of the time, ALL of the time). Responses were coded where a higher score indicated more positive perceptions about the neighborhood context.

Peer Group Affiliation

Peer affiliation items were included based on the effect of emergent neighborhoods on adolescent development (Elliot et al., 1996). The measure was created based on a theoretical relationship between neighborhood culture and organization on adolescent behaviors, particularly in disadvantaged neighborhoods (Elliott et al.). The goal is to assess the peer group for each student, and the outcome provides two scores representing friends with positive and negative behaviors respectively (Elliott et al.; Ji et al., 2013). This measure targeted student perspectives of their peers, with questions regarding how many of their friends are bullies, are interested in school, and so on. The question stem was, "How many of your FRIENDS do these

things?" Items were rated on a 4-point scale (NONE, SOME, MOST, ALL), with 4 indicating positive behaviors.

Risk Behaviors

To assess substance use and serious violent tendencies, a measure was adapted from the Centers for Disease Control (2004). This measure asks students to indicate if they have ever taken controlled substances, been violent towards another person, as well as requests information about gang experiences. The outcome for this measure calculates the number of times students indicate they have participated in a risky behavior (CDC, 2004; Ji et al., *2013*). The question stem was, "The next questions ask if you have ever done some things. If you have never done them, you can circle 'No'." Items were rated on a 4-point scale (No, Yes, Yes 2-5 times, Yes more than 5 times), with an increased level of risky behavior indicating a reduction in social-emotional skills and character.

Moral Center. Negative and positive belief in moral center was measured with items from the Item-Construct Dictionary for the Student Survey of Risk and Protective Factors (Arthur et al., 2000). The measure targeted student preference and allowance of desirable and undesirable behaviors in a social setting (Arthur et al., 2000; Ji et al., 2013). These items are designed to assess student agreement or disagreement with moral issues such as cheating, violence, stealing, honesty, and self-sacrificing behaviors. The question stem was, "For each statement below, please circle one choice about how much you agree or disagree." Items were scored on a 5-point scale (*disagree a lot, disagree a little, do not agree, agree a little, agree a lot*). Outcomes from this measure consist of two scores: one related to positive behavior items

and one for the negative behavior items. For the purposes of the present analysis, all items were coded with four indicating higher social-emotional skills and character.

APPENDIX D

PATH DIAGRAMS FOR SEM MODELS

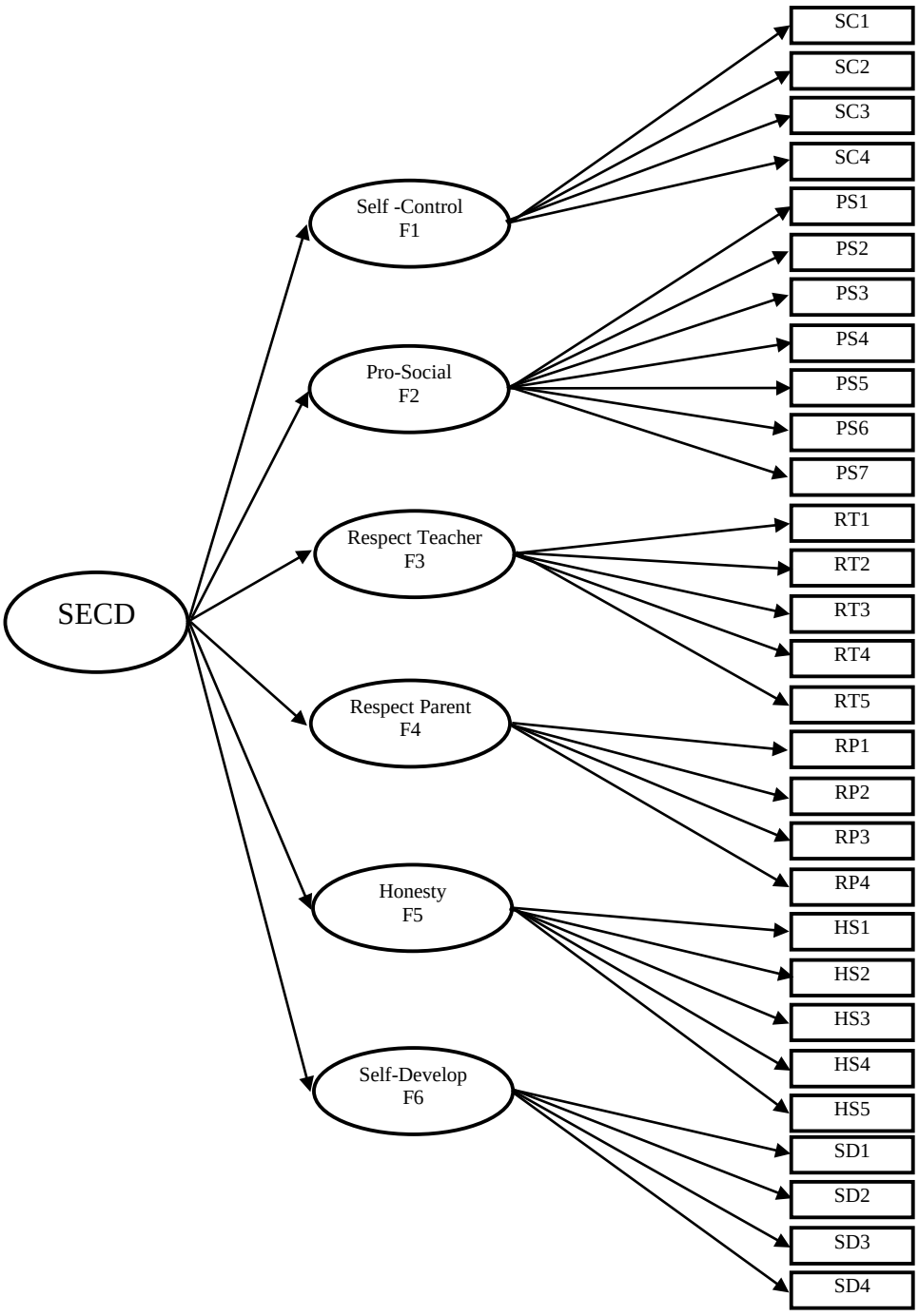

Figure D.1. Hypothesized baseline model of the SECDS structure featuring one higher-order social-emotional and character development trait (SECD) and six first order factors: self-control, pro-social, respect teacher, respect parent, honesty, and self-development. Associated error terms not shown.

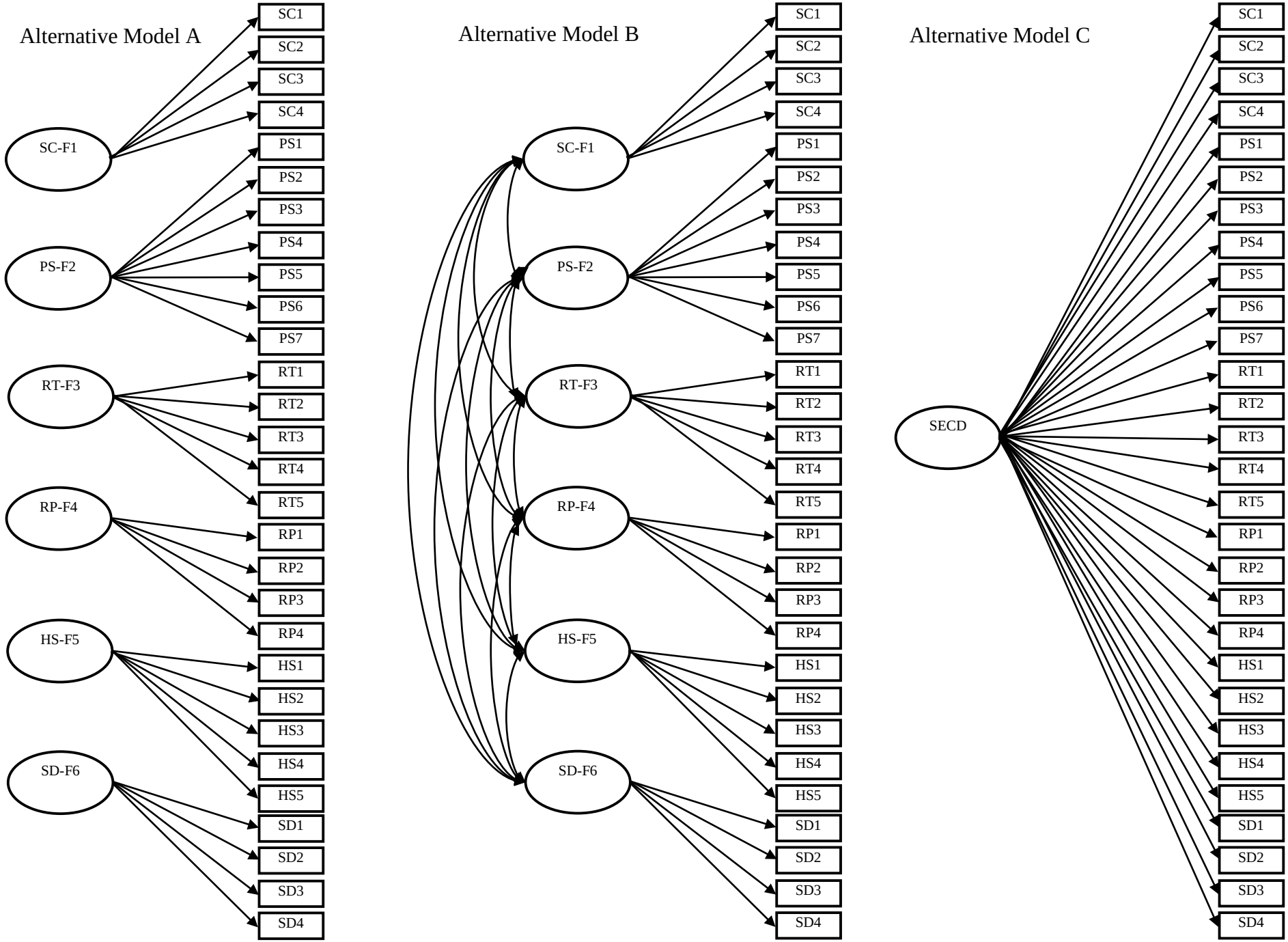

Figure D.2. Alternative models of the SECDS structure: Alternate A - Six uncorrelated factors; Alternate B - Six correlated factors; and Alternate C - One factor. Associated error terms not shown.

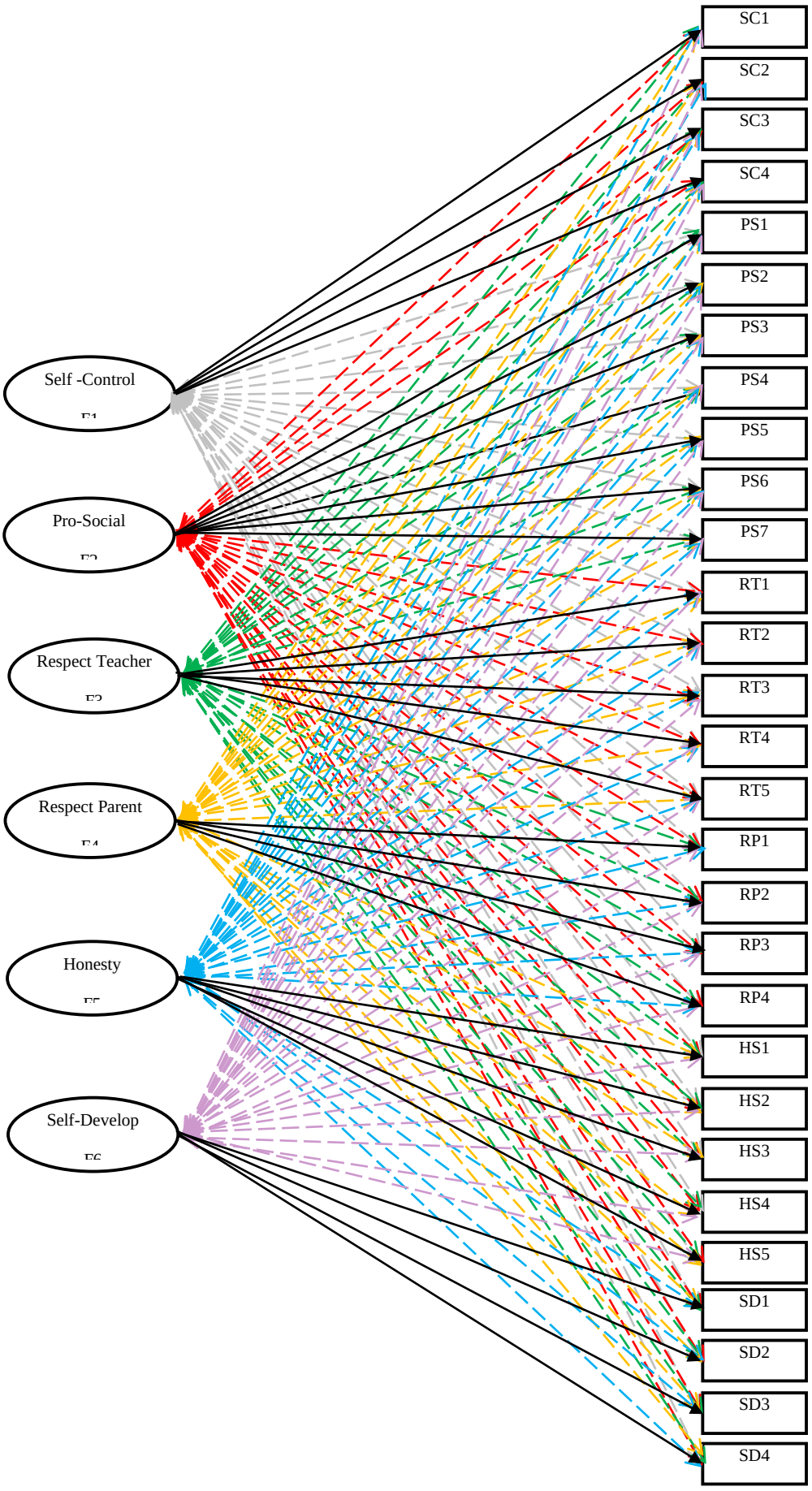

Figure D.3. Proposed ESEM model of the SECDS structure with six first order factors. Factor indicators are highlighted in solid black directional arrows. Structure coefficients are indicated in colored dashed lines. Associated error terms not shown.

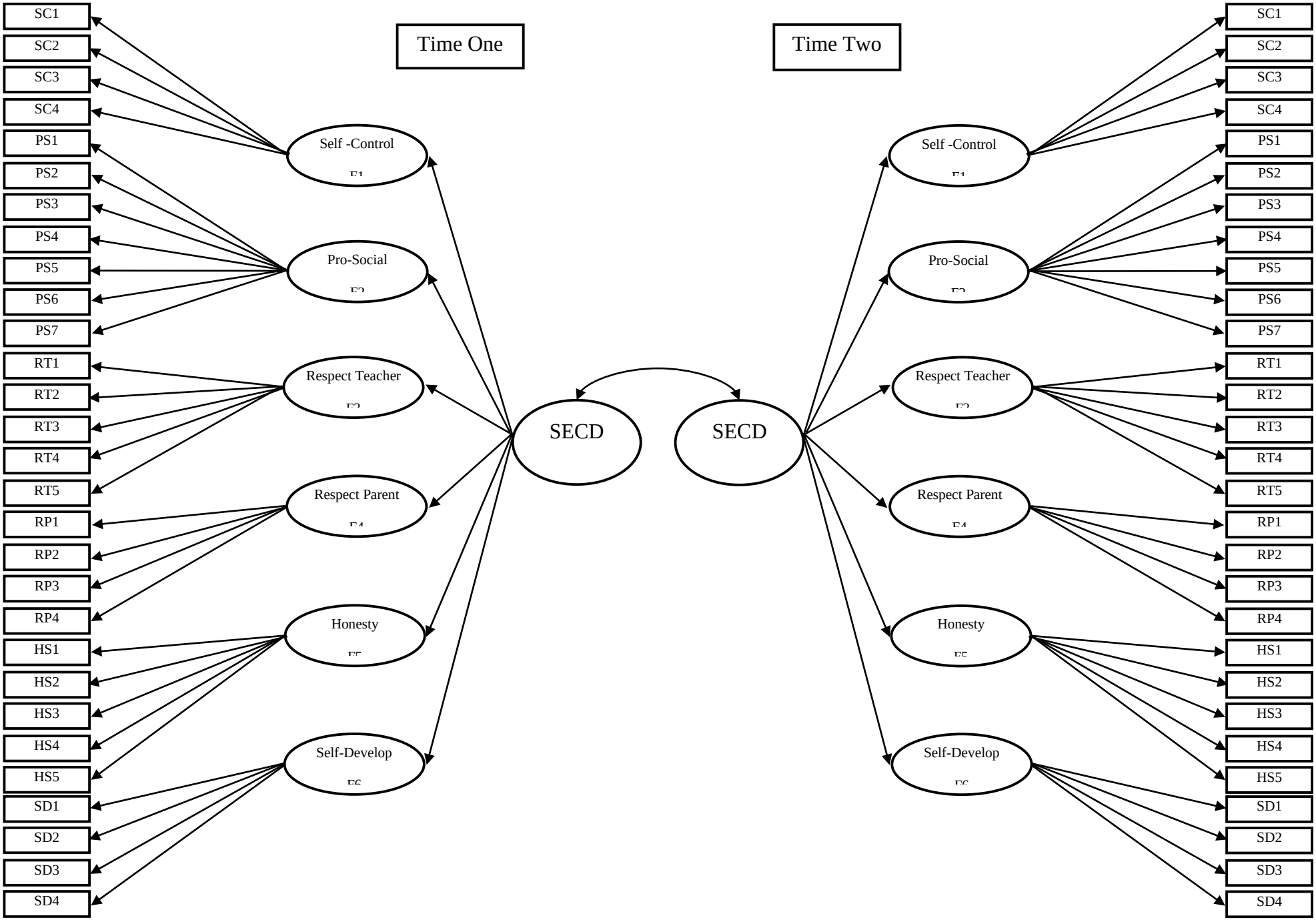

Figure D.4. CCA test-retest model correlating Time 1 and Time 2. Pending model fit, final SECDS structural model may include correlations as shown in ESEM model. Associated error terms not shown.

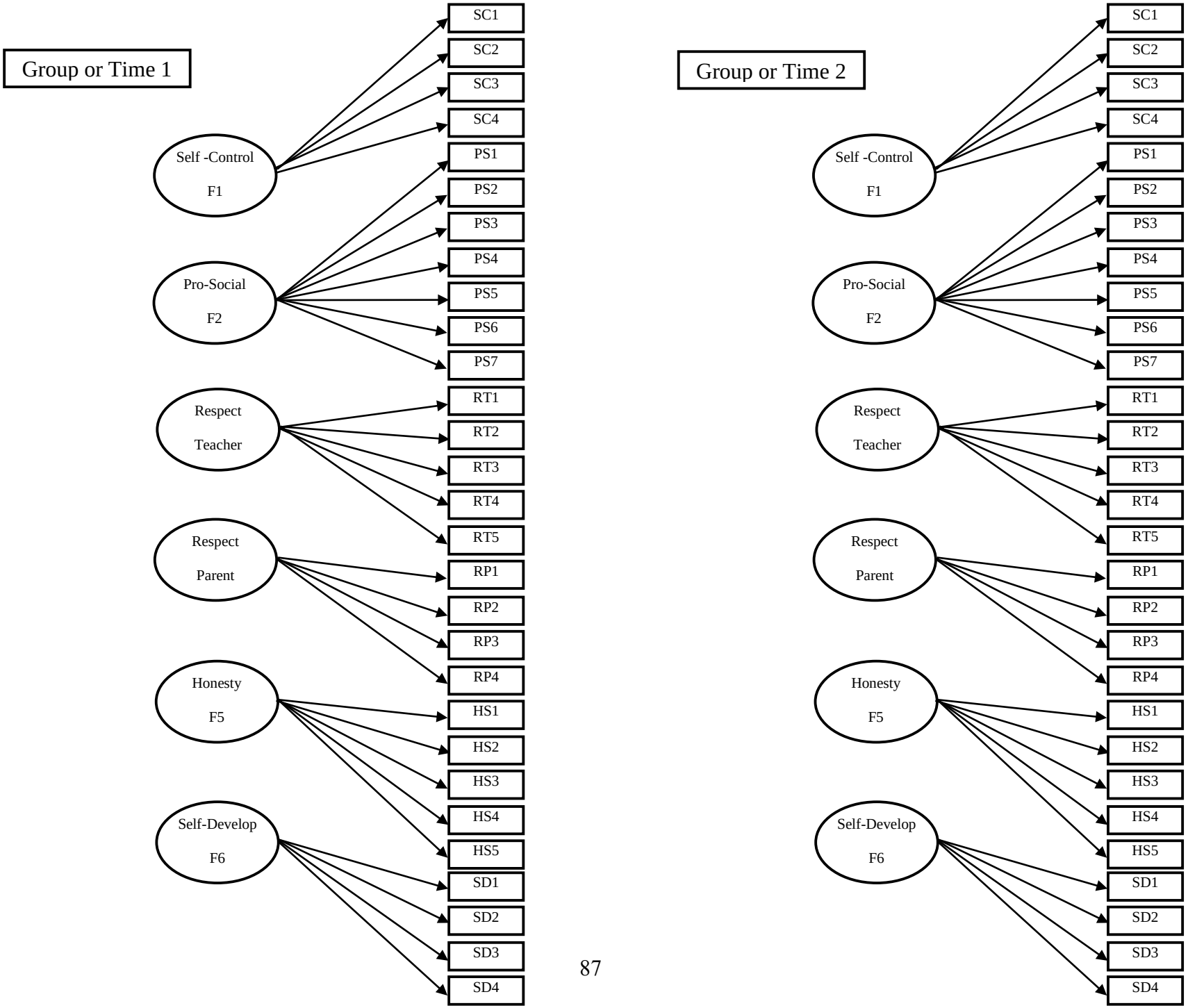

Figure D.5. Multigroup baseline configural model. Pending model fit, final SECDS structural model may include correlations as shown in ESEM.

APPENDIX E

MPLUS SYNTAX

```
 DATA:
   FILE IS F:\Belize\DISSERTATION\DATA\MPLUSDATA\4TO6SECDwcomposite.csv;

 VARIABLE:
 !NAMES OF VARIABLES IN ORDER OF FILE
 NAMES ARE
ID STD SCHGRP SCODE TCODE GENDER ETHNIC
!PYD PRETEST
!REWARDS FOR PROSOCIAL BEHAVIOR
PRER1 PRER2 PRER3 PRER4 PRER5 PRER6
!DEPRESSION AND ANXIETY
PRER7 PRER8 PRER9 PRER10 PRER11 PRER12
!PARTICIPATION IN NEIGHBORHOOD CONTEXT
PRER13 PRER14 PRER15 PRER16 PRER17 PRER18 PRER19 PRER20 PRER21
!PEER AFFILIATION
PRER22 PRER23 PRER24 PRER25 PRER26 PRER27 PRER28
!SUBSTANCE ABUSE AND VIOLENCE
PRER29 PRER30 PRER31 PRER32 PRER33 PRER34 PRER35 PRER36 PRER37 PRER38 PRER39 PRER40
!MORAL CENTER
PRER41 PRER42 PRER43 PRER44 PRER45 PRER46 PRER47 PRER48 PRER49 PRER50
!SOCIAL-EMOTIONAL CHARACTER DEVELOPMENT
PRER51 PRER56 PRER63 PRER73 PRER55 PRER60 PRER62 PRER65 PRER67 PRER68
PRER72 PRER53 PRER54 PRER66 PRER74 PRER75 PRER58 PRER59 PRER70 PRER71
PRER52 PRER57 PRER61 PRER64 PRER69 PRER76 PRER77 PRER78 PRER79
!PYD POSTTEST
!!REWARDS FOR PROSOCIAL BEHAVIOR
POSTR1 POSTR2 POSTR3 POSTR4 POSTR5 POSTR6
!DEPRESSION AND ANXIETY
POSTR7 POSTR8 POSTR9 POSTR10 POSTR11 POSTR12
!PARTICIPATION IN NEIGHBORHOOD CONTEXT
POSTR13 POSTR14 POSTR15 POSTR16 POSTR17 POSTR18 POSTR19 POSTR20 POSTR21
!PEER AFFILIATION
POSTR22 POSTR23 POSTR24 POSTR25 POSTR26 POSTR27 POSTR28
!SUBSTANCE ABUSE AND VIOLENCE
POSTR29 POSTR30 POSTR31 POSTR32 POSTR33 POSTR34 POSTR35 POSTR36 POSTR37 POSTR38
POSTR39 POSTR40
!MORAL CENTER
POSTR41 POSTR42 POSTR43 POSTR44 POSTR45 POSTR46 POSTR47 POSTR48 POSTR49 POSTR50
!!SOCIAL-EMOTIONAL CHARACTER DEVELOPMENT
POSTR51 POSTR56 POSTR63 POSTR73 POSTR55 POSTR60 POSTR62 POSTR65 POSTR67 POSTR68
POSTR72 POSTR53 POSTR54 POSTR66 POSTR74 POSTR75 POSTR58 POSTR59 POSTR70 POSTR71
POSTR52 POSTR57 POSTR61 POSTR64 POSTR69 POSTR76 POSTR77 POSTR78 POSTR79
!COMPOSITE VARIABLES PRIOR TO IMPUTATION
RESIL1re RESIL2re REWARD1 REWARD2 ANXIETY1 ANXIETY2
NEIGHB1 NEIGHB2 PEERS1 PEERS2  VIOLEN1 VIOLEN2 MORAL1 MORAL2
CHARACT1 CHARACT2 SELFCON1 SELFCON2 PROSOC1 PROSOC2
RESPTCH1 RESPTCH2 RESPPRT1 RESPPRT2 HONESTY1 HONESTY2
SELFDEL1 SELFDEL2
;

 !VARIABLES USED IN MODEL;
 USEVARIABLES =
 PRER51 PRER56 PRER63 PRER73 PRER55 PRER60 PRER62 PRER65 PRER67 PRER68
 PRER72 PRER53 PRER54 PRER66 PRER74 PRER75 PRER58 PRER59 PRER70 PRER71
 PRER52 PRER57 PRER61 PRER64 PRER69 PRER76 PRER77 PRER78 PRER79
;

 !VARIABLES AS CATEGORICAL;
 CATEGORICAL ARE
 PRER51 PRER56 PRER63 PRER73 PRER55 PRER60 PRER62 PRER65 PRER67 PRER68
 PRER72 PRER53 PRER54 PRER66 PRER74 PRER75 PRER58 PRER59 PRER70 PRER71
 PRER52 PRER57 PRER61 PRER64 PRER69 PRER76 PRER77 PRER78 PRER79
```

```
;
  !INDICATE CHARACTER FOR MISSING VALUES;
  MISSING ARE .
;
  ANALYSIS:
  ESTIMATOR = WLSMV;
  ROTATION=GEOMIN (OBLIQUE, .5);
  PARAMETERIZATION = THETA;
  DIFFTEST IS DERIVCFA.DAT;

  MODEL:
F1-F6 BY PRER51 PRER56 PRER63 PRER73 PRER55 PRER60 PRER62 PRER65 PRER67 PRER68
PRER72 PRER53 PRER54 PRER66 PRER74 PRER75 PRER58 PRER59 PRER70 PRER71
PRER52 PRER57 PRER61 PRER64 PRER69 PRER76 PRER77 PRER78 PRER79 (*1);

  OUTPUT:
  STANDARDIZED RESIDUAL TECH2 TECH4;

  SAVEDATA:
  DIFFTEST IS DERIVESEM.DAT;
```