Assessment Simplified

A Series of Briefs on
Key SEL Assessment Issues and Practices

SSISCoLab

Assessment and Intervention. Aligned. Simplified!

Attention Colleagues

All materials in this booklet may be shared freely with colleagues and parents interested in the assessment of children and youth's social emotional learning skills.

New *Assessment Simplified Briefs* are written quarterly. To access these and the corresponding *Intervention Simplified Briefs* series, visit www.ssiscolabs.com.

Stephen N. Elliott, PhD
Mickelson Foundation Professor
Arizona State University
and
President
SAIL Collaborative, LLC
snelliott25@gmail.com

Table of Contents

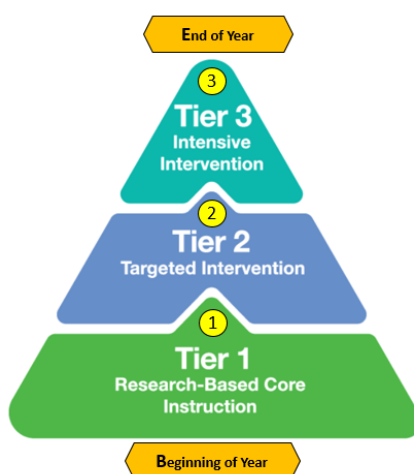# ASSESSMENT SIMPLIFIED

SSIS™ SEL
Social-Emotional Learning Edition

BRIEF #1:
ASSESSMENT DRIVEN MTSS
USING ITEM RESPONSE THEORY TO IMPROVE SEL ASSESSMENT

Prevention, positive behavior support, and response-to-intervention initiatives are in many schools today as central parts of a multi-tiered system of supports (MTSS) designed to improve the social and academic lives of all students and help create positive school culture. A widely accepted framework for conceptualizing MTSS involves three levels (or tiers) of support services. Tier 1 support represents research-based core instruction for all students. Tier 2 support represents intervention targeted to approximately 15% of students who need extra support (e.g., more time, practice, and feedback) to improve, and Tier 3 support represents a more intensive intervention (e.g., more individualized, longer duration) for an estimated 5% to 10% of students. Effective MTSS management is predicated on (a) systematically identifying students' needs, (b) using assessment data to guide decisions about the selection of effective strategies to improve student behavior and/or achievement, (c) implementing effective interventions, and (d) evaluating student progress and intervention outcomes (Glover, 2018). A number of challenges exist to implementing effective multi-tiered services focused on improving students' SEL skills. These include the existence of few reliable screening assessments aligned to interventions, a limited number of proven effective universal interventions, and few formal assessments for monitoring students' progress.

Brief Assessments for Making Evidence-Based Decisions

The need for psychometrically sound assessments for making decisions within multi-tiered student support (MTSS) systems spurred the development of the *SSIS SEL Brief Scales*. These rating forms measure a representative sample of social–emotional competencies related to *self-awareness, self-management, social awareness, relationship skills,* and *responsible decision making*. They can be completed in 5 minutes by a teacher or students (Grades 3-12) to screen an entire class, grade level, or school on 20 key SEL skills (Elliott, DiPerna, Anthony, Lei, & Gresham, 2020). These

End of Year

③
Tier 3
Intensive Intervention

②
Tier 2
Targeted Intervention

①
Tier 1
Research-Based Core Instruction

Beginning of Year

Decision Point E. Determine end of year competency levels to establish changes from baseline competency levels for all students.

Decision Point 3. For participants, (a) determine outcome of Tier 3 intervention, (b) note progress from baseline, and (c) decide on subsequent support needs.

Decision Point 2. For participants, (a) determine outcome of Tier 2 intervention, (b) note progress from baseline, and (c) decide on subsequent support needs.

Decision Point 1. For participants, (a) determine outcome of Tier 1 intervention, (b) note progress from baseline, and (c) decide on subsequent support needs.

Decision Point B. Determine beginning of year competency levels to establish baselines for all students and level of support needed.

assessments are functionally short-form versions of the comprehensive *SSIS SEL Edition Rating Forms-Teacher or Student* (Gresham & Elliott, 2017). The accompanying Figure illustrates key assessment decision points where these Brief Scales can be used within a 3-level MTSS system across a school year.

SEL Assessment Decisions

Educators involved in SEL programs differ with regard to how they use assessments to make the best decisions for their students and schools. The SSIS SEL Brief Scales-Teacher and Student versions help users make a number of decisions about individual and groups of students. Specifically, these decisions are about:
- Students' overall SEL competency status or proficiency level,
- Students' specific SEL skill strengths and areas in need of improvement,
- Changes in students' SEL competency level and skills over time, and
- Performance benchmarks for SEL competences across grades/developmental levels.

When the skills assessed are well aligned with skills taught in a program based on the CASEL competency framework, the *SSIS SEL Brief Scales* also can be used to make decisions about:
- SEL skill units to be taught and
- Effects of intervention programs on students' SEL skills improvement.

As highlighted in the K-12 MTSS model figure, there are five basic assessment decision points across the three levels of support. Specifically, in coordination with the implementation and completion of each specific level of support, are summative assessment decision points or gates. In addition to the intervention level assessment decision points, the comprehensive framework encourages beginning of year and end of year assessments. We consider this 5-decision point assessment approach to implementing an SEL focused MTSS model as a ***Basic Intervention Assessment (BIA) model***. This BIA model can be implemented effectively using SSIS assessments with both student and teacher informants. For the majority of students, three assessments over the course of a school year are likely: Baseline, Point 1, and End. For students receiving more support than Tier 1, four or five assessments may be needed: Baseline, Point 1, Point 2 and/or Point 3, and End. Clearly, with any of these assessment plans, it is important that the assessments be time-efficient, well aligned with the skills being provided in the supportive intervention, and yield reliable and valid scores for making moderate stakes decisions.

Alternative models that build upon or significantly modify the BIA model are possible. An assessment model that features more frequent progress monitoring within tiers of support could supplement or supplant Decision Points 1, 2, or 3. Progress monitoring (PM) could be completed via direct observations by a program leader during intervention sessions, self-monitoring charts, or other assessment tools like the *SSIS SEL Screening/Progress Monitoring Scales* (Elliott & Gresham, 2017), Thus, this approach could be called ***BIA+PM model*** when progress monitoring assessments occur prior to Decision Points 1, 2, or 3.

Another model could be conceptualized within a typical fall-winter-spring Benchmarking approach prevalent with many academic assessment programs. The Benchmark Points would be calendar based with the fall and spring assessment points replacing the Baseline and Exit points in the BIA, and the winter assessment replacing the Decision Points 1, 2, and 3. A **Benchmarking Model (BM)** for SEL that is coordinated with academic benchmark assessments may be attractive to many educators and likely provides opportunities for integrating students' social and academic data.

Conclusion

The *SSIS SEL Brief Scales* are flexible, efficient, and validated measures for making fundamental decisions about students and programs to improve students' SEL skills. An MTSS services framework provides an effective structure for prioritizing and integrating SEL assessment and intervention actions for individual students and groups of students. In general, however, the *SSIS SEL Brief Scales* can be a primary assessment for use in a variety of MTSS models because they are time-efficient, cost effective, and aligned to the CASEL competency model and most states' SEL standards.

References

Elliott, S.N., DiPerna, J.C., Anthony, C.J., Lei, P., & Gresham, F.M. (2020). *SSIS SEL Brief Scales.* Scottsdale, AZ: SAIL CoLab.

Elliott, S. N., & Gresham, F. M. (2017b). *SSIS SEL Edition Screening/Progress Monitoring Scales* [Measurement instrument]. Bloomington, MN: NCS Pearson.

Glover, T. A. (2018). Response-to-intervention models and access to services for all students. In S. N. Elliott, R. J. Kettler, P. A. Beddow, & A. Kurz (Eds.), *Handbook of Accessible Instruction and Testing Practices: Issues, Innovations, and Applications* (2nd ed., pp. 157–166). New York, NY: Springer.

Learn More …

Intervention Simplified Briefs #1 through #10 are available at
https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Behavior/SSIS-Social-Emotional-Learning-Edition/p/100001940.html

Assessment Simplified Briefs #1 and beyond are available at https://ssiscolab.com/

Founder & Senior Scientist-Practitioner:
Stephen N. Elliott, PhD, Mickelson Foundation Professor
Arizona State University

Contact information: snelliott25@gmail.com.

BRIEF #2:
MAXIMIZING MEASUREMENT EFFICIENCY
*USING ITEM RESPONSE THEORY TO IMPROVE SEL ASSESSMENT*

Services targeting students' social and emotional learning (SEL) skills is on the rise in today's schools and much of this activity occurs within the context of multi-tiered systems of support (MTSS). MTSS as a system has many advantages for the efficient allocation of resources to best meet the SEL needs of all students. One emphasis point of well-functioning MTSS models is assessment. In contrast to other service delivery approaches, MTSS requires multiple and varied assessments that are tailored to their specific function within different tiers of service delivery. For example, traditional rating scales are likely too long to function well as universal measures or progress monitoring tools for evaluating the SEL strengths and growth areas for all students within a school. Despite this need, most currently available SEL assessments are not well-tailored to applications within MTSS models. As such, there is a need for the development of a range of efficient and technically sound assessment that can be applied to make different types of decisions within MTSS models (see Assessment Simplified Brief # 1 for more information on Assessment Driven MTSS models).

Item Response Theory and the Development of Brief Forms

This need was spurred the development of the *SSIS SEL Brief Scales*, a set of time-efficient, multi-informant, rating scales designed for use with SEL-focused MTSS. To develop the *SSIS SEL Brief Scales* we utilized established items from the comprehensive *SSIS SEL Edition Rating Forms* (Gresham & Elliott, 2017) with advanced psychometric approaches that are well suited for the task for maximizing the efficiency of applied SEL assessments. Specifically, we used methods grounded in *Item Response Theory.*

Item Response Theory (IRT) refers to a set of analytic techniques that enable fine-grained evaluation of the functioning of individual items on assessments. Although most psychometricians consider IRT to be "modern measurement theory," it is not frequently used for assessment in socioemotional/behavior or SEL assessment. Briefly, IRT entails models that estimate the probability of an item response (e.g., being assigned a score of 2 on a SSIS SEL
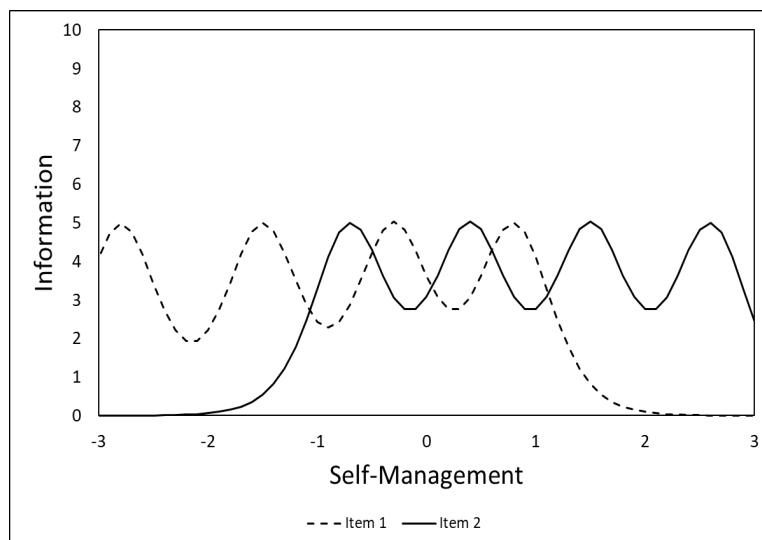
Brief Scale item) as a mathematical function of (a) characteristics of items such as their difficulty and (b) the target of assessment such as a student's SEL skills. Because analysis proceeds at the item level rather than from the vantage point of entire scales, IRT allows test-developers to isolate and retain the best functioning items when creating or refining measures.

As with many aspects of IRT, visual explication is helpful. Based on IRT results, test developers can evaluate items by the amount of *information* they provide. Information in IRT is a technical term that is akin to reliability and indicates the level of precision provided by single items or sets of items. When plotted, each item has an information function that shows the level of precision across the spectrum of the targeted skill or behavior domain. For example, consider the displayed figure on the first page, showing information functions for two hypothetical items targeting students' self-management skills. As can be seen in the figure on page 1, Item 2 (solid line) provides more precision at higher levels of self-management than Item 1 (dashed line), which provides more precision at lower levels of self-management. This level of detail on item functioning provides test developers guidance on streamlining assessments to be as efficient as possible, while still retaining as much precision as possible. For more in-depth information about IRT, please see Embretson and Reise (2013), Thomas (2011; 2019) or DeMars (2010).

By utilizing IRT, researchers (e.g., Anthony, DiPerna, & Lei, 2016; Anthony & DiPerna, 2017; 2018; Moulton, von der Embse, Kilgus, & Drymond, 2019) have been able to maximize the efficiency of several widely used measures. Yet, there are very few options available for brief, technically sound SEL assessment. This limitation motivated us to utilize IRT methodology to develop the *SSIS SEL Brief Scales*. This goal was highly successful. For example, the displayed table shows traditional reliability estimates for the *SSIS SEL Brief Scales – Teacher K-12 Form*. Note the high level of reliability

Scale	Cronbach's α	Test-Retest
Self-Awareness	.83	.83
Self-Management	.83	.83
Responsible Decision Making	.87	.78
Relationship Skills	.79	.75
Social Awareness	.83	.75
SEL Composite	.93	.84

even though each SSIS SEL Brief Scale is only 4 items and the entire scale only includes 20 items. Such reliability estimates support the use of the *SSIS SEL Brief Scales* for making universal screening decisions involving competency areas in need of improvement and for making general progress assessments across MTSS tiers.

Conclusion

The *SSIS SEL Brief Scales* (Teacher, Student, & Parent versions) were all designed for efficient usage within MTSS models. These student support models require assessments that are streamlined for specific applications with large groups of students. Unfortunately, few such measures exist for students' SEL skills. Thus, the *SSIS SEL Brief Scales* are unique in their efficiency, content coverage, and psychometric sophistication and should support expansion of services for all children's SEL skill development.

References

Anthony, C. J., & DiPerna, J. C. (2017). Identifying sets of maximally efficient items from the Academic Competence Evaluation Scales—Teacher Form. *School Psychology Quarterly, 32(4),* 552. Doi: 10.1037/spq0000205

Anthony, C. J., & DiPerna, J. C. (2018). Piloting a short form of the academic competence evaluation scales. *School Mental Health, 10,* 1–8. doi: 10.1007/s12310-018-9254-7.

Anthony, C. J., DiPerna, J. C., & Lei, P.-W. (2016). Maximizing measurement efficiency of behavior rating scales using item response theory: An example with the Social Skills Improvement System–Teacher Rating Scale. *Journal of School Psychology, 55,* 57-69. doi:10.1016/j.jsp.2015.12.005

DeMars, C. (2010). *Item response theory*. New York: Oxford.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory.* Mahwah, NJ: Psychology Press.

Moulton, S., von der Embse, N., Kilgus, S., & Drymond, M. (2019). Building a better behavior progress monitoring tool using maximally efficient items. *School Psychology, 34*(6), 695–705. doi: 10.1037/spq0000334

Thomas, M. L. (2010). The value of Item Response Theory in clinical assessment: A review. *Assessment, 18(3),* 291–307. doi: 10.1177/1073191110374797

Thomas, M. L. (2019). Advances in applications of item response theory to clinical assessment. *Psychological Assessment, 31*, 1442–1455. doi: 10.1037/pas0000597

Learn More …

Other *Assessment Simplified Briefs* are available at https://ssiscolab.com/

Research Scientist:
Christopher J. Anthony, PhD
University of Florida
Contact information: canthony@coe.ufl.edu

ASSESSMENT SIMPLIFIED

BRIEF #3:
Students' Self-Assessment of Social Emotional Competencies
Tools and Technical Considerations

Self-assessment via rating scales is a very common, if not the most widely used, approach to documenting students' social emotional learning (SEL) competencies. Indeed, some SEL scholars have noted that self-report measures are "irreplaceable sources of information about children's views of themselves" (McKown, 2015, page 323). An examination of the *CASEL Assessment Guide* (https://measuringsel.casel.org/assessment-guide/) indicates that 16 of the 26 SEL assessment reviewed are student self-rating scales. Most of these scales are for students in grades 6 to 12 and broadly focus on a portion of key social-emotional competencies. Specifically, the CASEL framework involves five competency domains as defined in the figure.

Because of the prominence of this framework in state and school SEL content standards (Eklund, Kilpatrick, Kilgus, & Haider, 2018), it is important that content alignment between standards and assessment is high, especially for users who want to build interventions based on assessment results.

Challenges in Using Self-Assessments. Regardless of the uses, however, challenges persist for self-report assessments. These

Domain	Definition
Self-Awareness	The ability to accurately recognize one's emotions and thoughts, and their influence on behavior. This includes accurately assessing one's strengths and limitations and possessing a well-grounded sense of confidence and optimism. *Example items:* I ask for help when I need it. I pay attention when others present their ideas.
Self-Management	The ability to regulate one's emotions, thoughts, and behaviors effectively in different situations. This includes managing stress, controlling impulses, motivating oneself, and setting and working toward achieving personal and academic goals. *Example items:* I stay calm when dealing with problems. I try to find a good way to end a disagreement.
Social Awareness	The ability to take the perspective of and empathize with others from diverse backgrounds and cultures, to understand social and ethical norms for behavior, and to recognize family, school, and community resources and supports. *Example items:* I help my friends when they are having a problem. I stand up for others when they are not treated well.
Relationship Skills	The ability to establish and maintain healthy and rewarding relationships with diverse individuals and groups. This includes communicating clearly, listening actively, cooperating, resisting inappropriate social pressure, negotiating conflict constructively, and seeking and offering help when needed. *Example items:* I work well with my classmates. I try to forgive others when they say "sorry."
Responsible Decision Making	The ability to make constructive and respectful choices about personal behavior and social interactions based on consideration of ethical standards, safety concerns, social norms, the realistic evaluation of consequences of various actions, and the well-being of self and others. *Example items:* I follow school rules. I am careful when I use things that aren't mine.

Note. All definitions from CASEL (2013, p. 9).

include readability, response bias, and social desirability, and lack of accessibility for children typically younger than 8 years of age. Another very important issue to consider is administration time, which is influenced directly by the length of an assessment. Considering

the facts that SEL competencies are increasingly targeted in universal programs and student "voice" is a valued premise of social emotional development (Elias, et al. 2015), brief, student-friendly, and reliable assessments are important to the viability of evidence-driven universal SEL programs.

There are, however, only two current self-assessments for students in elementary grades through high school that focus on the SEL competency framework advanced by CASEL. These assessments are the *Washoe County School District Student Social Emotional Competency Assessment* (Crowder, Gordon, Brown, Davidson, & Domitrovich, 2019) and the *SSIS – SEL Edition Rating Form - Student*; SSIS SEL-S; Gresham & Elliott, 2017). Furthermore, of these two measures, only the SSIS SEL-S has norms from a nationally representative sample of students and is aligned with an evidence-based intervention program (the *SSIS SEL Classwide Intervention Program*; Elliott & Gresham, 2017). Despite these advantages, the SSIS SEL-S is still relatively long (46 items taking about 12 minutes to complete), which results in administration time that hinders use for MTSS Tier 1 assessments where all students are screened and their progress is monitored via periodic repeated assessments.

Development and Initial Validation of the SSIS SEL*b*-S

Given these considerations of administration time, large number of students to assess, repeated measurement, and content alignment, a shortened version of the SSIS SEL-S was designed using IRT methods to create a maximal efficiency measure (see **Assessment Simplified Brief #2** for a more detailed account of our IRT procedures). The result is a strength-focused rating scale for 3rd to 12th graders called the *SSIS SEL Brief Scales-Student Form* (SSIS SEL*b*-S; Elliott, DiPerna, Anthony, Lei, & Gresham, 2020a). Specifically, we identified 20 items of the 46 items on the *SSIS SEL Rating Form-Student* version by selecting four highly informative items for each of the five SEL competency domains. Representative items for each competency domain are provided in the figure on the previous page and other items with their response options are illustrated below. Collectively, students' completion of these 20 items results in a Composite

	Social Behaviors	Not TRUE	A Little TRUE	A Lot TRUE	Very TRUE
1	I pay attention when others present their ideas.	Not	A Little	A Lot	Very
2	I stay calm when dealing with problems.	Not	A Little	A Lot	Very
3	I help my friends when they are having a problem.	Not	A Little	A Lot	Very
4	I work well with my classmates.	Not	A Little	A Lot	Very
5	I do the right thing without being told.	Not	A Little	A Lot	Very

SEL Score and five Competence Domain scores. All raw scores are transformed to standard scores and percentile ranks. In addition, scores +1 and -1 standard deviations from the mean score are used to create cut points for three proficiency levels – *Developing, Competent,* and *Advanced* – to describe students' social-emotional functioning. Research indicates the SSIS SEL*b*-S is fully aligned with the CASEL framework, has a readability of grade 2.5, and can be completed in 5 minutes by virtually all students. There is also a Teacher version of the SSIS SEL*b* (Elliott, DiPerna, Anthony, Lei, & Gresham, 2020b) to facilitate multi-informant assessments of the same five SEL competencies.

Reliability. As documented in the *SSIS SEL Brief Scales User Guide & Technical Manual* (2020; online at SSIScolab.com), traditional reliability evidence for the SSIS SEL*b*- S was strong.

Specifically, (a) Cronbach's α levels were .91 for the SSIS SEL*b*-S SEL Composite and ranged from .67 to .72 (median = .69) across SSIS SEL*b*-S scales and (b) test-retest reliability coefficients were .87 for the SSIS SEL*b*-S SEL Composite and ranged from .62 to .88 (median = .77) across SSIS SEL*b*-S scales. Using IRT-based Test Information Functions as an indicator of reliability all SSIS SEL*b*-S scales maintain a .70 level across broad levels of each SEL construct.

Validity. The evidence was largely supportive of the construct validity of SSIS SEL*b*-S scores. Considering its brevity, it is likely that the current evidence would be acceptable to identify children with potential SEL needs (using the SEL Composite score) and to identify preliminary strengths and weaknesses that might be used to inform intervention planning.

Summary

Self-report assessments, when well-constructed and used appropriately, can be useful for making group screening and intervention decisions, and individual diagnostic decisions with students age 8 and older (Denham, 2015; Elliott, Frey, & Davies, 2015). The SSIS SEL*b*-S holds promise to meet the needs of educators who value students' perspectives on their own SEL competence development. There are a number of SEL self-report assessments, but few are well-aligned with the prominent CASEL framework and fewer still that involve most elementary students. Considering the importance of students' voice in SEL programs, the development and use of reliable and valid assessments like the SSIS SEL*b*-S is a step forward.

References

Eklund, K., Kilpatrick, K. D., Kilgus, S. P., & Haider, A. (2018). A Systematic Review of State-Level Social–Emotional Learning Standards: Implications for Practice and Research. *School Psychology Review*, *47*, 316-326.

Elias, M. J., Leverett, L., Duffell, J. C., Humphrey, N., Stepney, C., & Ferrito, J. (2015). Integrating SEL with related prevention and youth development approaches. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook for social and emotional learning: Research and practice* (pp. 33–49). New York, NY: Guilford.

Denham, S. A. (2015). Assessment of SEL in educational contexts. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook for social and emotional learning: Research and practice* (pp. 285–300). New York, NY: Guilford.

Elliott, S.N., DiPerna, J.C., Anthony, C.J., Lei, P., & Gresham, F.M. (2020a). *SSIS SEL Brief Scales-Student Form.* Scottsdale, AZ: SAIL CoLab.

Elliott, S.N., DiPerna, J.C., Anthony, C.J., Lei, P., & Gresham, F.M. (2020b). *SSIS SEL Brief Scales-Teacher Form.* Scottsdale, AZ: SAIL CoLab.

Elliott, S.N., Frey, J.R., & Davies, M. (2015). Systems for assessing and improving students' social skills to achieve academic competence. In J. Durlak, C. Domitrovich, R. Weissberg, & T. Gullotta (Eds.). *Handbook of Social & Emotional Learning: Research & Practice,* (pp. 301-319) New York: Guilford Press. doi 10.1093/oxfordhb/9780199935291.013.70.

Elliott, S.N., & Gresham, F.M. (2017). *SSIS SEL Classwide Intervention Program (CIP).* Scottsdale, AZ: SAIL CoLab.

Gresham, F. M., & Elliott, S. N. (2017). *SSIS SEL Edition Rating Form-Student*. Bloomington, MN: Pearson.

McKown, C. (2017). Social-emotional assessment, performance, and standards. *The Future of Children*, 157-178.

Learn More …

Assessment Simplified Briefs are available at https://ssiscolab.com/

Founder & Senior Scientist-Practitioner:
Stephen N. Elliott, PhD
Mickelson Foundation Professor
 Arizona State University
Contact information: snelliott25@gmail.com.

BRIEF #4:

Screening the Whole Social Emotional Child for SEL Programs

Students' Self-Ratings of Their SEL Strengths & Emotional Behavior Concerns

Social emotional learning (SEL) programs contribute to educating the whole child and are being implemented in school throughout the country. Generally, these programs are universal by design and focus on teaching or increasing all students' desired social emotional skills in five core competency domains (i.e., self-awareness, self-management, social awareness, responsible decision making, and relationship skills). Programs like the *SSIS SEL Classwide Intervention Program* (CIP; Elliott & Gresham, 2017), a CASEL SELect Program and one that focuses on teaching these five core competencies, consistently has been shown to increase elementary students' frequency of SEL skills. Interestingly, the CIP also has been shown to decrease many problem behaviors and increase academic engagement of elementary students (e.g., DiPerna, Lei, Bellinger, & Cheng, 2015, 2016; DiPerna, Lei, Cheng, Hart, & Bellinger, 2017). Some students, perhaps 10% to 20%, however, continue to display undesirable or negative emotional behaviors even if they participate in an effective universal SEL program or behavior management program (e.g., Forness, Kim, & Walker, 2012). Common undesirable behaviors that are likely to persist involve both internalizing concerns (e.g., anxiousness, depression) and externalizing concerns (e.g., aggressiveness, bullying) (Vidair, Sauro, Blocher, Scudellari, & Hoagwood, 2014).

> *What if you could, with less than 10 minutes of student time, get reliable data on the social emotional health of all your students? Would you want the data for only their SEL skills or would you prefer also to have data on any significant emotional behavior concerns?*

Identifying students with emotional behavior concerns (EBC) as early as possible is a sound intervention practice that is being overlooked or ignored in some SEL programs. Many SEL advocates have argued that a stronger focus on positive student strengths is needed in today's schools. Although such a perspective has unquestionable merit, when taken to an extreme, there is potential to neglect students' needs in other areas, such as emotional behavior concerns, that can be present in students with low, medium, and high levels of SEL skills. Current models of mental health conceptualize complete mental health as being composed of two distinct dimensions: one dimension involves *psychosocial wellbeing/positive experiences* and the other dimension *psychosocial distress/negative experiences*. In summary, although you may be primarily interested in improving students' SEL competencies, it seems wise to consider screening and monitoring for indicators of psychological distress as part of all multi-tiered screening assessments and intervention efforts. Thus, it is important to consider how effective and efficient a narrow band (SEL skills only) versus a broadband (SEL + emotional behavior concerns) assessment of students is relative to SEL program goals that stress strength focused, "whole" child intervention.

At present, many schools that are concerned about the whole social emotional child use teacher referrals, rather than formal assessments, to identify students at risk for social, emotional, and behavioral difficulties (e.g., Bruhn, et al., 2014). Although teachers generally identify students with

externalizing difficulties accurately, they routinely have been found to under-identify students experiencing internalizing problems (e.g., Dowdy, Doane, Eklund, & Dever, 2013). Teachers in middle and high schools, compared to elementary teachers, have more difficulty identifying students with internalizing concerns because they generally interact less time with their students.

Student Self-Assessment of Their SEL and Emotional Behavior Concerns
The first step in solving the problem of effectively serving the social emotional needs of all students is a reliable and valid screening process that can be completed by students to efficiently identify their strengths and areas of concern. Self-assessment via rating scales is a very common, if not the most widely used approach to documenting students' SEL competencies. Indeed, McKown (2015, p. 323) noted that self-report measures are "irreplaceable sources of information about children's views of themselves." An examination of the CASEL Assessment Guide in 2019 indicated that 15 of the 26 SEL assessments reviewed are student self-rating scales. Most of these assessments broadly focus on social-emotional competencies and are completed by students in grades 6-12. A few of the most recently developed scales focus more narrowly on the SEL competency framework advanced by CASEL while covering students in grades 3 to 12. Interestingly, none of the 26 SEL assessments include scales involved emotional behavior concerns or problem behaviors. Thus, it is fair to say that the vast majority of SEL assessments in use today only screen for desired social emotional skills, ignoring undesirable emotional behavior concerns. This situation, however, has changed with the publication of the *SSIS SEL Brief + Mental Health Scales* (Elliott, Anthony, DiPerna, Lei, & Gresham, 2020).

The SSIS SEL Brief + Mental Health Scales
The SSIS SEL*b* + MHS consists of Teacher, Parent, and Student Forms that each measure both SEL skills (*Self-Awareness, Self-Management, Social Awareness, Relationship Skills, and Responsible Decision Making*) and EBCs (*Internalizing and Externalizing*). Each of these brief measures is based on its SSIS Rating Scale (Gresham & Elliott, 2008) counterpart and decades of research on the inter-relationships among social skills and internalizing and externalizing social behaviors of children (Elliott, Davies, & Frey, 2015). Each of these forms is comprised of 30 items (4 items for each of the 5 SEL scales and 5 items for the 2 EBC scales) and requires informants 8 minutes to complete online. The accompanying figure provides a sampling of both SEL and EBC items from the Student Form of the SSIS SEL*b* + MHS; items 21, 23, 25, and 27 are externalizing (aggressive and bullying) behaviors, while items 22, 24, 26, and 28 are internalizing behaviors that are common concerns. The same 10 EBCs are on the SSIS SEL*b* + MHS Teacher, Parent, and Student Forms, allowing for a comprehensive multi-informant assessment of

	Social Emotional Health Behaviors	Not TRUE	A Little TRUE	A Lot TRUE	Very TRUE
1	I pay attention when others present their ideas.	Not	A Little	A Lot	Very
2	I try to forgive others when they say "sorry."	Not	A Little	A Lot	Very
3	I am careful when I use things that aren't mine.	Not	A Little	A Lot	Very
4	I stand up for others when they are not treated well.	Not	A Little	A Lot	Very
16	I keep my promises.	Not	A Little	A Lot	Very
17	I stay calm when dealing with problems.	Not	A Little	A Lot	Very
18	I work well with my classmates.	Not	A Little	A Lot	Very
19	I ask for help when I need it.	Not	A Little	A Lot	Very
20	I stay calm when I disagree with others.	Not	A Little	A Lot	Very
21	I make people do what I want them to do.	Not	A Little	A Lot	Very
22	I think bad things will happen to me.	Not	A Little	A Lot	Very
23	I do not let others join my group of friends.	Not	A Little	A Lot	Very
24	I think no one cares about me.	Not	A Little	A Lot	Very
25	I try to make others afraid of me.	Not	A Little	A Lot	Very
26	I feel nervous with my classmates.	Not	A Little	A Lot	Very
27	I say things to hurt people's feelings.	Not	A Little	A Lot	Very
28	I feel sad.	Not	A Little	A Lot	Very

key emotional behavior concerns. Thus, in 5 minutes or less these screening measures yield reliable SEL performance level information (e.g., Emerging, Developing, Competent, and Advanced levels) to help establish pre-intervention baselines for all students and for 2 or 3 additional minutes of assessment

time, they can also screen for those 10% to 20% of students experiencing emotional behavior concerns. The result: A more complete picture of the health of the whole social emotional child!

An Example of What You Miss when Only Screening for SEL Skills

As part of the development and validation research for the SSIS SEL*b* + MHS, we conducted studies in several schools. Most of the schools wanted a brief SEL only screener, while a couple schools with Mental Health Teams were motived to know more about their students. For purposes of this report, we want to highlight SSIS SEL*b* + MHS-Student results from one school's fall schoolwide screening in five 5th grade classrooms. This screening involved 113 students.

The accompanying figure summarizes the results of the screening. Specifically, it features (in the center) the SEL Performance Levels of all students with 35 (31%) at the Advanced level, 70 (62%) at the Competent level, 8 (7%) at the Developing level, and 0 at the Emerging level. These performance levels are based on the transformation of the Composite Raw SEL score via the SEL Competency-Referenced Performance Framework (See *Assessment Simplified Brief #5* for more details). Many educators and parents would be very happy with these results given 93% of students indicated themselves to be at a Competent or Advanced level of social emotional functioning. Yet, when we add in the results from the EBC scales notice that 21 (18.6%) students self-identified at the Concern level for either an internalizing or an externalizing EBC. Not the vast majority (17 of 21) of students were considered to be

Internalizing Concerns		SEL Performance Levels	Externalizing Concerns	
Concerns	No Concerns		No Concerns	Concerns
2	33	35 Advanced	32	3
5	65	70 Competent	63	7
2	6	8 Developing	6	2
0	0	0 Emerging	0	0
9 8%	104 92%	113 5th Grade	101 89%	12 11%

functioning at the Competent or Advanced SEL performance levels, not the lowest levels of SEL performance. Specifically, the data indicate 9 (8%) of the 5th graders characterized themselves with an Internalizing emotional behavior concern and 12 (11%) with an Externalizing emotional behavior concern.

Summary

Currently, none of the widely used screening assessments in use with SEL programs measure any emotional behavior concerns or problem behaviors of students. It is understandable for SEL programs to focus on developing students' strengths, but to concurrently risk overlooking behaviors indicative of aggression, bullying, anxiety, or depression is unnecessary and unwise. These emotional behavior concerns can be efficiently and effectively assessed and, in many cases, addressed within some SEL intervention program. If SEL programs are to meaningfully contribute to developing the whole social emotional child, they can't afford to assess only the SEL part!

References

Bruhn, A.L., Woods-Groves, S., & Huggle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education and Treatment of Children, 37,* 611-634.

DiPerna, J. C., Lei, P., Bellinger, J., & Cheng, W. (2015). Efficacy of the Social Skills Improvement System Classwide Intervention Program (SSIS-CIP) primary version. *School Psychology Quarterly, 30*(1), 123–141. doi:10.1037/spq0000079

DiPerna, J. C., Lei, P., Bellinger, J., & Cheng, W. (2016). Effects of a universal positive classroom behavior program on student learning. *Psychology in the Schools, 53*(2), 189–203. doi:10.1002/pits.21891

DiPerna, J. C., Lei, P., Cheng, W., Hart, S. C., & Bellinger, J. (2017). A cluster randomized trial of the Social Skills Improvement System-Classwide Intervention Program (SSIS-CIP) in first grade. *Journal of Educational Psychology, 110*(1), 1–16. doi:10.1037/edu0000191

Dowdy, E., Doane, K., Eklund, K., Dever, B. (2013). A comparison of teacher nomination and screening to identify behavioral and emotional risk within a sample of underrepresented students. *Journal of Emotional and Behavioral Disorders, 21,* 127-137.

Elliott, S.N., DiPerna, J.C., Anthony, C.J., Lei, P., & Gresham, F.M. (2020a). *SSIS SEL Brief Scales-Student Form.* Scottsdale, AZ: SAIL CoLab.

Elliott, S.N., DiPerna, J.C., Anthony, C.J., Lei, P., & Gresham, F.M. (2020b). *SSIS SEL Brief Scales-Teacher Form.* Scottsdale, AZ: SAIL CoLab.

Elliott, S.N., Frey, J.R., & Davies, M. (2015). Systems for assessing and improving students' social skills to achieve academic competence. In J. Durlak, C. Domitrovich, R. Weissberg, & T. Gullotta (Eds.). *Handbook of Social & Emotional Learning: Research & Practice,* (301-319) New York: Guilford.

Elliott, S.N., & Gresham, F.M. (2017). *SSIS SEL Classwide Intervention Program (CIP).* Scottsdale, AZ: SAIL CoLab.

Forness, S.R., Kim, J., & Walker, H.M. (2012). Prevalence of students with EBD: Impact on general education. *Beyond Behavior, 21,* 3-9.

LeBuff, P.A., Shapiro, V.B., & Naglieri, J.A. (2014). *Devereux Student Strengths Assessment K-8th grade.* Charlotte, NC: Aperture Education.

McKown, C. (2017). Challenges and opportunities in the direct assessment of children's social and emotional competence. In J. Durlak, C. Domitrovich, R. Weissberg, & T. Gullotta (Eds.). *Handbook of Social & Emotional Learning: Research & Practice,* (320- 335) New York: Guilford.

Vidair, H.B., Sauro, D., Blocher, J.B., Scudellari, L.A., & Hoagwood, K.E. (2014). Empirically supported school-based mental health programs targeting academic and mental health functioning: An update. In H.M. Walker & F.M. Gresham (Eds.), *Handbook of evidence-based practices for emotional and behavioral disorders: Applications in schools.* (pp. 15-53). New York, NY: Guilford.

Learn More …

Assessment Simplified Briefs are available at https://ssiscolab.com

Authors
Stephen N. Elliott, Arizona State University
Christopher J. Anthony, University of Florida
Christopher Huzenic, Pearson Assessment
Contact information: snelliott25@gmail.com.

ASSESSMENT SIMPLIFIED

BRIEF #5: The SEL Competency-Referenced Performance Framework
Documenting the Status and Progress of All Students in CASEL Aligned Programs

The central goal of social emotional learning (SEL) programs is the improvement of all students' social emotional competencies. To know if this improvement goal is accomplished, one needs a clear definition of SEL competencies, evidence of positive change in the competencies, and criteria for evaluating the amount or quality of change.

Many of the SEL competencies that society values have been articulated in the popular CASEL Competency Framework (www.CASEL.org). This theoretical framework privileges five

competency domains that represent the intra-personal competencies of self-awareness and self-management, the inter-personal competencies of social awareness and relationship skills, and a fifth domain that is considered both an inter- and intra-personal competency, responsible decision making (Figure 1). These competency domains, and the many skills representative of them, can be assessed, taught, and improved; thus, advancing children's chances of functioning well and preventing social emotional problems at school, home, and in their communities.

To document evidence for the improvement of students' SEL competencies most educators use some type of an assessment. Most current assessments of SEL competencies and skills are behavior rating scales. To review dozens of such assessments, visit the CASEL Assessment Guide (https://measuringsel.casel.org/access-assessment-guide/) or the RAND Assessment Finder (https://www.rand.org/education-and-labor/projects/assessments.html) websites, both online resources for persons interested in finding the best assessment for their information needs. A careful examination of the SEL assessments described on these websites, however, indicates the majority are poorly aligned content-wise to the CASEL competency framework and provide results as standard scores and/or percentile ranks based on normative samples. These types of assessments, when used as pre- and post-intervention measures can be used to quantify changes in students' functioning, but the metric of change is general in standard score units, a difficult type of score to comprehend by many educators. For example, knowing a student has improved 10 standard score points is not particularly informative, even when you know that the average student in the same SEL program improved 7 standard score points. Many educators want more practical information from assessments than just standard scores and percentile ranks. They want information about students' developmental status and progress against an understandable criterion. A new competency-referenced framework has been developed to

address educators' need for practical information about students' SEL functioning as measured by assessments.

The SEL Competency-Referenced Performance Framework (CPRL)

The SEL CRPF facilitates a competency-referenced interpretive approach based on the CASEL competency framework. It was developed initially to facilitate the interpretation of raw score ratings from the SSIS SEL Brief Scales, but has broader applications to other assessments (Elliott, Anthony, DiPerna, Lei, & Gresham, 2020). Specifically, this strength-focused, competency-referenced approach characterizes clusters of self-awareness, self-management, social awareness, relationship, and responsible decision making skills into four performance levels: *Emerging*, *Developing*, *Competent*, and *Advanced*. The Competent level of performance of the SEL CRPF is presented Figure 2.

Content-wise the five competence domains (illustrated in the figure by bullet points) that comprise each of the four performance levels align 100% with the SEL competencies in the CASEL Framework. A close read of the Competent Performance Level description (and each of the other Performance Levels) indicates three aspects of a skill are featured: the observed *frequency* (i.e., never, seldom, often, and almost always) of the skill; the range of *support* (i.e., from direct prompting, to cueing, minimal prompting, and independently) needed to elicit the skill, and; the *generalized use* (i.e., range of social situations) of the skill. The complete CRPF is available at SSIScolab.com.

Competent Level

Students at this performance level **often** exhibit most of the following competencies appropriately and with minimal or no prompting in social situations:

- Accurately recognize one's emotions, thoughts, and their influence on behavior. This includes accurately assessing one's strengths and limitations, and possessing a well-grounded sense of confidence and optimism.

- Regulate one's emotions, thoughts, and behaviors effectively in different situations. This includes managing stress, controlling impulses, motivating oneself, and setting and working toward achieving personal and academic goals.

- Take the perspective of and empathize with others from diverse backgrounds and cultures, understand social and ethical norms for behavior, and recognize family, school, and community resources and supports.

- Establish and maintain healthy and rewarding relationships with diverse individuals and groups. This includes communicating clearly, listening actively, cooperating, resisting inappropriate social pressure, negotiating conflict constructively, and seeking and offering help when needed.

- Make constructive and respectful choices about personal behavior and social interactions based on consideration of safety concerns, social norms, the realistic evaluation of consequences of various actions, and the well-being of self and others.

Each competence performance level is intended to vary developmentally in comparison to the next higher level. Thus, across the performance levels from Emerging to Advanced, one observes a progression of fundamental SEL skills that occur more frequently, with less support, and in more social situations or environments. Figure 3 (on the next page) illustrates these key conceptual dimensions of this CRPF used to interpret the SEL raw score for the Self-Management Scale on the SSIS SEL Brief Scales – Teacher K-12 Form. The same general progressions of skills are expected for each competence domain and collectively contribute to one's composite SEL performance level.

Use of the SEL CRPF with the SSIS SEL Brief Scales Assessments and Beyond

To operationalize each of the four performance levels for use with the SSIS SEL Brief Scales, we needed to determined score ranges that defined each performance level. These score ranges were primarily influenced by the item response anchors of *never, seldom, often,* and *almost always* and a learning progression theory of SEL competency development. With these factors in mind, we initially proposed that the children's Emerging level performances across 20 items typically would result in frequency ratings of *never (0 points)* or *seldom (1 point).* Children at the Developing level typically would be expected to have frequency ratings of *seldom (1 point)* or *often (2 points).* Children at the Competent level performance would consistently have frequency ratings of *often (2 points)* or *almost always (3 points).* Finally, children at the Advanced level performance would consistently have frequency ratings of *Almost Always (3 points)* on nearly all items. A detailed account of the development of the SEL CRPF is provided in the User Guide & Technical Manual of the SSIS SEL Brief Scales (Elliott, Anthony, DiPerna, Lei, & Gresham, 2020).
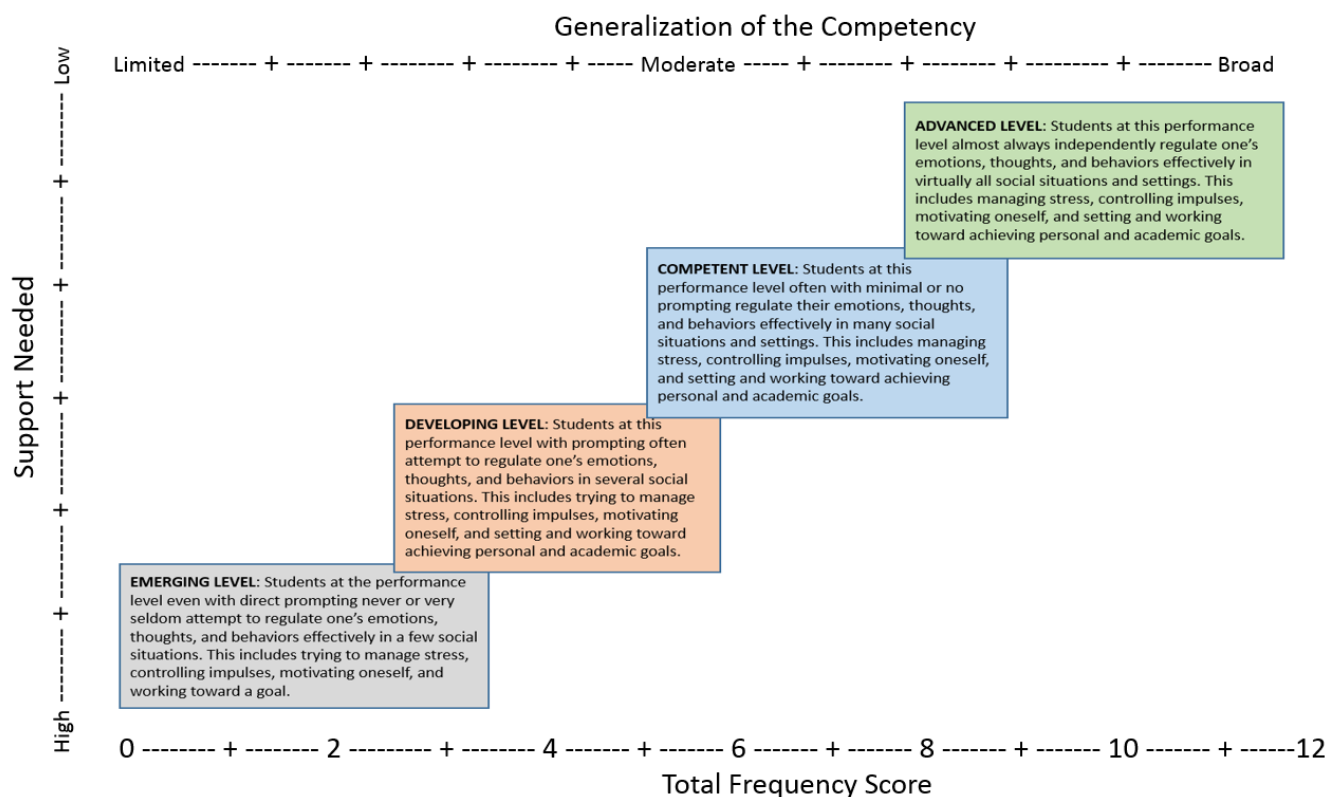
Using this logic regarding the typical item scores for children at each performance level, we generated several possible cut scores to differentiate the four levels. We then tested these various cut scores to determine their impact on the percentage of students from our Performance Cut Score Samples assigned to each performance level. After testing the impact of possible sets of cut scores on the proportion of students in each of our five grade clusters (PreK, K-2, 3-5, 6-8, & 9-12) and racial/ethnic subgroups (Black, Hispanic, Other, & White), we

determined the score ranges to define our competency-referenced performance levels. This approach to determining cut scores and defining score ranges borrows heavily from performance standard setting methods used with many statewide achievement tests (e.g., Cizek & Earnest, 2016).

Conclusions
The SEL CRPF can be used to transform scores from assessments that are (a) aligned substantially with the content of the CASEL Competency Framework and (b) used to evaluate instructional programs designed to improve children's SEL competencies. For example, the scores from each of the SSIS SEL family of multi-informant assessments (i.e., SSIS SEL Rating Forms, SSIS SEL Screening/Progress Monitoring Scales, and SSIS SEL Brief Scales) can be transformed to the SEL CRPF performance levels because they each are aligned with the CASEL Competency Framework and yield reliable and valid scores for the five competency domains.

The SEL CRPF, however, does more than transform scores. Specifically, it can be used to integrate various forms of assessment evidence (e.g., rating scale scores, direct assessment scores, observations, performance ratings) into a common framework that is descriptive, strength-focused, and developmentally relevant. In effect, the CRPF functions as a "grading" framework (A-C-D-E rather than the traditional A-B-C-D-E used in schools across the country). As a result, practical communications are advanced regarding students' functioning status and progress in an SEL program.

Finally, it is important to note that competency-referenced results are not designed to replace norm-referenced assessment results and probably should not be used when making high stakes decision without extensive predictive validity evidence to support them. However, for many educators charged with implementing SEL program, a rich description of students' current performance capabilities with regarding self-management or responsible decision making skills is more meaningful than a standard score of 55 shaded in blue or a percentile rank of 70 on an assessment with a normalized standard score scale.

The goal of improving all students' SEL competencies is fundamental to all SEL programs. Determining if this goal is achieved requires sound assessments and a developmental conceptualization of improvement. The new CRPF contributes to integrating assessment results within just such a context.

References

Cizek, G.J., & Earnest, D.S. (2016). Setting performance standards on tests. In S. Lane, M.R. Raymond, & T.M. Haladyna (Eds.), *Handbook of Test Development (2nd ed.) (pages 212-238).* New York: Routledge.

Elliott, S.N., Anthony, C.J., DiPerna, J.C., Lei, P., & Gresham, F.M. (2020). *SSIS SEL Brief Scales User Guide & Technical Manual*. Scottsdale, AZ: SAIL CoLab.

Elliott, S.N., DiPerna, J.C., Anthony, C.J., Lei, P., & Gresham, F.M. (2020). *SSIS SEL Brief Scales-Teacher Form.* Scottsdale, AZ: SAIL CoLab.

Learn More …

Assessment Simplified Briefs are available at https://ssiscolab.com

Author
Stephen N. Elliott, Arizona State University
Contact information: snelliott25@gmail.com.

ASSESSMENT SIMPLIFIED

BRIEF #6: The SSIS SEL Family of Assessments for Intervention
Fundamental Characteristics and Uses with SEL Programs

A number of researchers have advanced the practices of social behavior assessments or intervention programs. Two scientist-practitioners, Frank Gresham and Steve Elliott, have championed the design of both assessments <u>and</u> intervention programs. Their assessments are for use by multi-informants (teacher, parent, and student), are strength-focused, and directly lead to an intervention program for teaching social emotional skills when identified as in need of improvement.

The SSIS SEL Family of Assessments for Intervention
The family of assessments Gresham and Elliott have invented and validated include the *Social Skills Improvement System Rating Scales* (Gresham & Elliott, 2007), the *SSIS Social Emotional Learning Edition Rating Forms* (Gresham & Elliott, 2017), the *SSIS SEL Screening/Progress Monitoring Scales* (Elliott & Gresham, 2017a), and most recently the *SSIS SEL Brief Scales* (Elliott, Anthony, DiPerna, Lei, & Gresham, 2020). The SSIS SEL Rating Forms, Brief Scales, and Screening/Progress Monitoring Scales each address the growing need for assessments aligned with the highly regarded CASEL SEL Competency Framework (CASEL, 2015). The figure below summarizes aspects of the SSIS SEL family of assessments that align with the CASEL five competency domains of the CASEL framework.

Assessment Name (Publisher/date)	# Items & Time	Alignment	Interpretation	Tiers & Uses
SSIS SEL Edition Rating Forms – Teacher, Parent, Student (Pearson, 2017)	46-51 15 min	CASEL SSIS-CIP	Norm-Referenced	Tiers 2+3 Identify S+W Plan Intervention
SSIS SEL Edition Screening/Progress Monitoring - Teacher (Pearson, 2017)	8 rubrics 30 min for class	CASEL SSIS-CIP	Criterion-Referenced	Tiers 1, 2, 3 Monitor Intervention Progress
SSIS SEL Brief Scales – Teacher K-12 Form (SSIS CoLab, 2020)	20 5 min	CASEL SSIS-CIP	Competency-Referenced	Baseline + End of Year Screening Program Evaluation
SSIS SEL Brief Scales – Parent K-12 Form (SSIS CoLab, 2020)	20 5 min	CASEL SSIS-CIP	Competency-Referenced	Baseline + End of Year Screening Program Evaluation
SSIS SEL Brief Scales – Student 3-12 Form (SSIS CoLab, 2020)	20 5 min	CASEL SSIS-CIP	Competency-Referenced	Baseline + End of Year Screening Program Evaluation
SSIS SEL Brief Scales – Teacher Preschool Form (SSIS CoLab, 2020)	20 5 min	CASEL SSIS-CIP	Competency-Referenced	Baseline + End of Year Screening Program Evaluation
SSIS SEL Brief Scales – Parent Preschool Form (SSIS CoLab, 2020)	20 5 min	CASEL SSIS-CIP	Competency-Referenced	Baseline + End of Year Screening Program Evaluation

Note. Each of these assessments is offered online and provides reports that link assessment results to specific skill units in the SSIS Classwide Intervention Program (CIP). The CIP is an evidenced based program covering 30 SEL skills and is recognized by CASEL as a SELect Program. The CIP is available at SSIScolab.com and all resources are online.
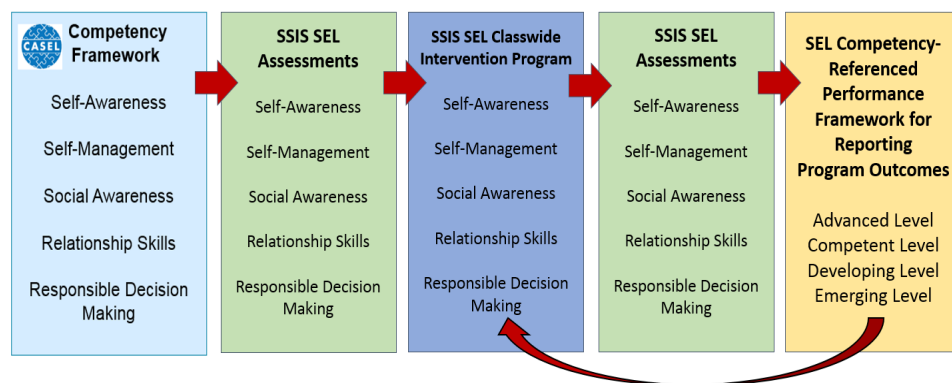
The SSIS SEL Classwide Intervention Program (CIP)

Concurrent with the development of assessments, Elliott and Gresham developed intervention programs based on effective social behavioral methods to teach the social emotional skills identified by these assessments as in need of improvement. The *SSIS SEL Classwide Intervention Program* (CIP; Elliott & Gresham, 2017b), their most recent program, is fully aligned with the competency content of the CASEL Framework. Specifically, this program focuses on 30 key SEL skills and teaches them using a 6-step structured process of Tell → Show → Do → Practice → Monitor Progress → Generalize. With empirical evidence in support of its efficacy (e.g., DiPerna, Lei, Bellinger, & Cheng, 2015; DiPerna, Lei, Cheng, Hart, & Bellinger, 2018) and growing interest in universal social emotional learning (SEL) programs, the SSIS CIP has been recognized by CASEL as a SELect Intervention Program for elementary school children.

The SSIS SEL Brief Scales & Their Uses

The newest members of the SSIS SEL family of assessments are the Brief Scales created using IRT methods for selecting the most informative items from the comprehensive SSIS SEL Rating Forms. They are name Brief Scales to stress that they are time-efficient rating scales comprised of 20 strength-focused items collectively measuring the five CASEL competency domains of *Self-Awareness,* S*elf-Management, Social Awareness, Relationship Skills,* and *Responsible Decision Making*. Substantial evidence supports the claim that each of the SSIS SEL Brief Scales yield reliable, valid, and fair scores for their intended purposes. As a result, these online assessments can be used by educational professionals, parents, and students for purposes of (1) *Screening* groups of students to identify their overall SEL competency level, (2) *Identifying* students' SEL strengths and skills in need of improvement, (3) *Planning* which SSIS SEL CIP skill units to teach for improving students' skills, (4) *Monitoring and documenting* change in students' SEL skills over time, and (5) *Evaluating* the effects of intervention programs on students' level of SEL skills. Each of the Brief Scales yields a Composite SEL score along with scores for each SEL competency domain. These scores are interpreted via the 4-level (Emerging, Developing, Competent, Advanced) SEL Competency-Referenced Performance Framework (see *Assessment Simplified Brief #5* for details about the CRPF)*.*

The scores from each of the SSIS SEL Brief Scales or any of the other members in the SSIS SEL family of assessments all can be transformed via the CRPF to descriptive Performance Levels*.* As a result, an SEL Competency-Referenced Assessment-Intervention-Outcomes model is created and can be used to guide the evaluation of SEL programs that embrace the CASEL Competency Framework. The figure summarizes the relations among input, output, and outcome components commonly featured in a logic model (McLaughlin & Jordon, 2015).

Conclusions

CASEL's theoretical SEL Competency Framework featuring five skill domains inspired the SSIS SEL multi-informant assessments. These assessments include comprehensive norm-referenced SEL Rating Forms, a criterion-referenced classwide SEL Screening/Progress Monitoring Scales (teacher only), and most recently competency-referenced SEL Brief Scales. Collectively, this group of assessments is a "family" because they all measure the same competencies, share a number of common items, and now have the ability to transform their composite scores to a common performance framework. Such family attributes make the SSIS SEL assessments particularly well-suited for use in MTSS-situated SEL programs where assessments are often repeated and used for a variety of purposes. With the development of a CRPF fully aligned with the CASEL SEL competencies, multiple SSIS SEL assessments can be used by multiple informants to yield scores interpreted via a common set of performance levels. This makes the SEL family of assessments uniquely valuable for evaluating SEL program outcomes.

References

Collaborative for Social, Emotional, and Academic Learning (2015). *The CASEL Competency Framework*. (www.CASEL.org).

DiPerna, J. C., Lei, P., Bellinger, J., & Cheng, W. (2015). Efficacy of the Social Skills Improvement System Classwide Intervention Program (SSIS-CIP) primary version. *School Psychology Quarterly, 30*(1), 123–141. doi:10.1037/spq0000079

DiPerna, J. C., Lei, P., Bellinger, J., & Cheng, W. (2016). Effects of a universal positive classroom behavior program on student learning. *Psychology in the Schools, 53*(2), 189–203. doi:10.1002/pits.21891

DiPerna, J. C., Lei, P., Cheng, W., Hart, S. C., & Bellinger, J. (2017). A cluster randomized trial of the Social Skills Improvement System-Classwide Intervention Program (SSIS-CIP) in first grade. *Journal of Educational Psychology, 110*(1), 1–16*.* doi:10.1037/edu0000191

Elliott, S.N., Anthony, C.J., DiPerna, J.C., Lei, P., & Gresham, F.M. (2020). *SSIS SEL Brief Scales User Guide & Technical Manual.* Scottsdale, AZ: SAIL CoLab.

Elliott, S.N., & Gresham, F.M. (2017a). *SSIS SEL Edition Screening/Progress Monitoring Scales.* Bloomington, MN: Pearson Assessment.

Elliott, S.N., & Gresham, F.M. (2017b). *SSIS SEL Classwide Intervention Program (CIP).* Scottsdale, AZ: SAIL CoLab.

Gresham, F.M., & Elliott, S.N. (2008). *SSIS Rating Scales.* Bloomington, MN: Pearson Assessment.

Gresham, F.M., & Elliott, S.N. (2017). *SSIS SEL Edition Rating Forms.* Bloomington, MN: Pearson.

McLaughlin, J.A., & Jordan, G.B. (2015. Using logic models. K.E. Newcomer, H.P. Hatry, & J.S. Wholey (Eds.). In Handbook of Practical Program Evaluation (4th ed.), (Chapter 3). New York: John Wiley.

Learn More …

Assessment Simplified Briefs are available at https://ssiscolab.com

Author
Stephen N. Elliott, Arizona State University
Contact information: snelliott25@gmail.com.

ASSESSMENT SIMPLIFIED

SSIS SEL
Social–Emotional Learning Edition

BRIEF #7
Fair Assessment of Children and Youth's SEL Competencies
A Design Imperative for the SSIS SEL Brief Scales

Fairness is an essential quality of social emotional learning (SEL) assessments, yet it receives far less coverage than the related technical qualities of reliability and validity. As highlighted in the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014), a common view of fairness in public discourse is that it involves "the equality of testing outcomes for relevant test-taker subgroups." The Standards',

however, explicitly exclude this common view from its examination of fairness and note that "group differences in outcomes do not in themselves indicate that a testing application is biased or unfair" (p. 54).

The fairness of SEL assessments is a challenging topic that is intertwined with discussions of educational equity, a central goal of many SEL

> *"Group differences in outcomes do not in themselves indicate that a testing application is biased or unfair."*
> (AERA, APA, & NCME, 2014)

programs. Such programs typically are designed to provide students access and opportunities to experiences and instructional support to improve SEL competences known to enhance both social and academic performance. Thus, many educators have championed SEL programs as an important tool to advance interpersonal understanding and relationships among students and educators with the goal of improving educational equity (Jagers, Rivas-Drake, & Borowski, 2018). Equity in an educational context means that personal or social circumstances such as gender, ethnic/racial origin or socio-economic background, are not obstacles to achieving educational potential. Specific to educational assessments, equity also means all students have access to participate in the assessment in meaningful ways, their resulting scores are free of bias, and the use of the results do not have a disparate impact for groups of students (Ercikan & Elliott, 2015).

In designing the SSIS SEL Brief Scales (Elliott, Anthony, DiPerna, Lei, & Gresham, 2020), a core design imperative was *Fair and Unbiased Assessment of All Students*. In response to this imperative, we established five fairness objectives for our assessments. These objectives were:
1. The skills and behaviors measured must be determined to be socially important by a diverse sample of stakeholders,
2. The responses to our items must be representative of a large sample of U.S. students with diverse racial/ethnic backgrounds,
3. The collection of items for each SEL competency must be empirically determined to be unbiased,
4. The online assessment must be highly accessible for English speakers, and
5. The score interpretation framework must not have evidence of disparate impact on students from diverse racial/ethnic backgrounds.

Evidence is provided about each of these objectives to support the claim that the SSIS SEL Brief Scales are fair and can be used with racially/ethnically diverse groups of students in schools in the United States.

The Skills or Behaviors Measured Must be Socially Important to Diverse Stakeholders
The items on the SSIS SEL Brief Scales have all been determined to align with one of the SEL competencies in the CASEL Framework (2015), but are the behaviors and skills measured accessible to all children and culturally relevant across diverse groups of students in US schools and communities? The social validity or importance of the content of these items is a fundamental aspect of the fairness of the assessments and provides evidence to address the likelihood that all students, regardless of their background, have an opportunity to learn and use these skills in their cultures and communities. To determine the social importance of all SSIS SEL items, we asked a racially/ethnically diverse sample (343 students, 726 teachers, and 1,680 parents) of users to evaluate the social importance of the item content of each item on each of the Brief Scales using a 0 = *Not Important,* 1 = *Important,* 2 = *Critical* rating. An examination of the resulting data indicated that for the total sample the Mean Importance ratings for the 20 items comprising the Composite Scale as rated by students, teachers, and parents were almost exactly the same: 1.27, 1.25, and 1.26, respectively. Item Importance ratings greater than 1.0 indicate that the skill measured by an item was perceived to be *Important* or *Critical* by the majority of raters in the total sample. This overall pattern of Importance ratings held across virtually all grade clusters and raters. Detailed importance ratings by teachers, parents, and students representing different racial/ethnic groups and males and females across all grade clusters indicated some variability amongst ratings, but no consistent patterns that suggest differences in perceived importance of the skills for students from different subgroups. Thus, the item content focus of the SSIS SEL Brief Scales is regarded as important to social functioning in schools and community cultures by a diverse group of student, teacher, and parent raters.

The Sample of Students Must Be Representative
One of the most common questions asked by users of an assessment concerns the racial/ethnic representativeness of the samples of students used to develop the assessment and its interpretation of results. For the development of the SSIS SEL Brief Scales, we used different combinations of what Gresham and Elliott (2008) called the norming cases and "extra" cases from their national standardization. Although the original SSIS SEL samples were large and comprehensive, they matched demographic estimates of students in 2006-2007. Thus, we generated new samples for the K-12 and Preschool assessments for purposes of evaluating the impact of the SSIS SEL Competency-Referenced Performance Level cut scores resulting in a sample of 1,091 student cases that matched demographic targets for sex and race/ethnicity projected for the US in 2025 (Snyder, de Brey, & Dillow, 2019). This resulted in a sample that was 51% female, 50% White, 15% Black, 25% Hispanic, and 10% Other racial/ethnic backgrounds to match demographic projections for each grade cluster (PreK, K-2; 3-5; 6-8; 9-12) to ensure roughly even representation of these characteristics across developmental levels.

The Items Must be Unbiased
As mentioned, a fundamental tenet of test bias is that group differences, in and of themselves, are not necessarily indicative of test bias. Average group differences can arise for many reasons. Test bias refers to the notion that average group differences arise from *artificial* differences in test functioning across groups. For this to occur, a test would need to function differently for

members of different groups not based on the assessed competence, but rather based on group membership. One prominent method of assessing this involves Differential Item Functioning (DIF) analyses, which evaluate whether individual items of an assessment function differently across groups. For example, the figure to

Shows Concern for Others
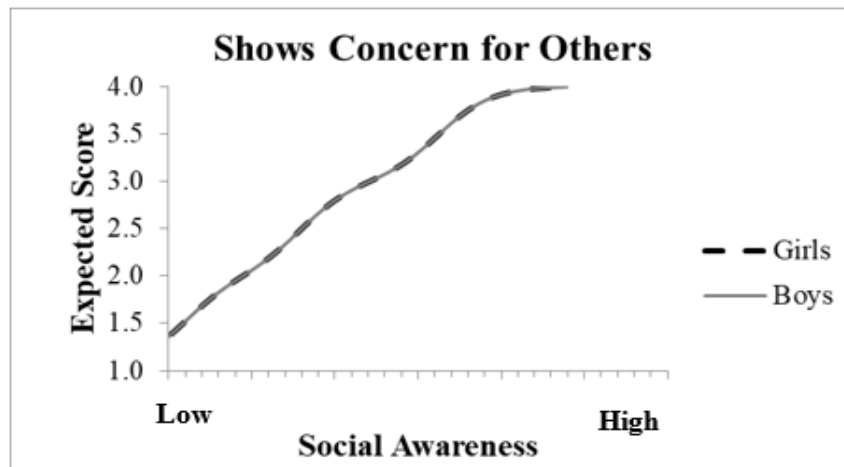
the right shows the results of some DIF analyses we conducted in the development of the SSIS SEL Brief Parent Scale. This graph shows that for both high and low levels of social awareness, boys and girls have the same expected rating on this item (the Y-axis), indicated by the fact that the lines for boys and girls completely overlap. Unsurprisingly, parents rate students with high social awareness higher on this item than those with low social awareness, but importantly, expected ratings do not differ for boys and girls at the same levels of social awareness. Thus, this item does not function differently for boys and girls and is thus, according to this definition of test bias, unbiased.

Not all items, however, show no evidence of bias. For example, the same item (*Shows Concern for Others*) showed very slight evidence of DIF when evaluated for race/ethnicity (White vs. Nonwhite). This difference is shown in the figure below. Note in the shaded region, white students are rated slightly higher than non-white students, on average, despite the fact that these students are equivalent on overall levels of social awareness. This last point is critical, an item

Shows Concern for Others
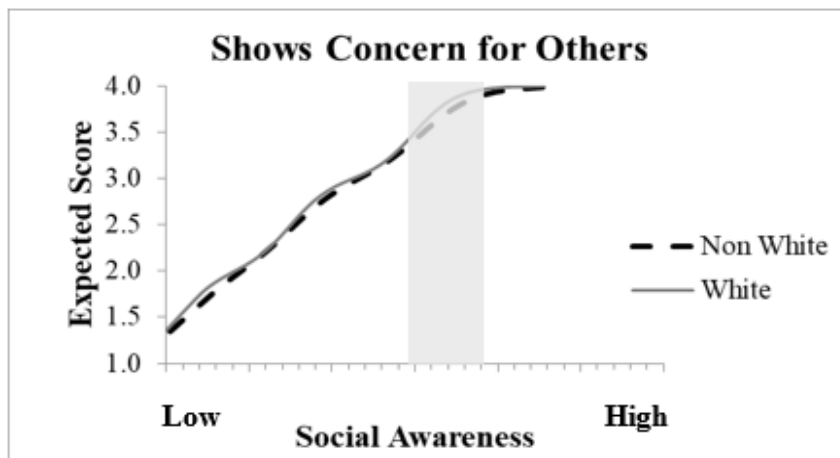
shows evidence of DIF only if it functions differently for individuals who are equivalent with regard to the competence assessed, in this case social awareness. Some groups do evidence differences on these sets of skills and indeed some of these differences are critical intervention targets. For example, based on the definitional features of Autism Spectrum Disorder (ASD), students with ASD should have lower social awareness on average, than students without ASD. Merely observing an average difference would not indicate a biased test unless items were functioning differently for students *who had equivalent levels of the targeted set of skills.*

Another important point about DIF illustrated by the second figure is that DIF is not an "all-or-none" property and it is not evaluated by purely statistical grounds. For example, the figure shows an item with exceedingly low DIF. Although some differences are statistically apparent, they are so small as to likely have no meaningful effect on the groups in question. Furthermore, showing concern for others is a core feature of social awareness. To omit it because of such minimal DIF would be to severely limit the meaningfulness and utility of the resulting assessment. For these reasons, this item was retained for inclusion of the SSIS SEL Brief Scales – Parent Form despite the small amount of statistical DIF it evidenced (formally evaluated with effect size measures; the Expected Score Standardized Difference; Meade, 2010).

The Assessment Must be Accessible

In accordance with the fairness guidance from the Testing Standards (AERA, APA, & NCME, 2014), the online instructions and rating scale content were designed according to Universal Design principles and are highly accessible. Specifically, the user experience (UX) and user interface (UI) of the Resonant Education software assessment application used to deliver the Brief Scales was independently evaluated and found to meet Level AA of the Web Content Accessibility Guidelines (WCAG 2.1). The readability levels (using the Flesch-Kincaid Grade Level formula) for the Teacher and Parent versions are 6.2 and 2.5 for the Student version. The application has read-aloud in multiple languages to support adult and student users. Thus, the SSIS SEL Brief Scales are considered highly accessible for students from diverse backgrounds with very basic reading abilities.

The Interpretation of Scores Must Not Result in Disparate Impact

A fair interpretation approach to scores from an assessment should not systematically favor one group over another. With this principle in mind, we opted for a criterion-referenced rather than a norm-referenced approach to score interpretation. Norm-referenced approaches to score interpretation by design compare students to students statistically, whereas with a criterion-referenced approach all students are compared to the same performance criteria not each other. Specifically, as documented in *Assessment Simplified Brief #5,* we created the SEL Competency-Referenced Performance Framework with four performance levels based on the CASEL competencies and common expectations for social behavior skill development. In addition to evaluating the overall impact of our Competency-Referenced Performance Level cut score ranges on the percentage of students at each level, we determined whether these cut score ranges would lead to disparate impact across gender and racial/ethnic subgroups. To do so, we compared percentages of students representing different sexes and different racial/ethnic groups falling within each level and determined whether these differences were meaningful. The results indicated there was no discernable pattern that consistently arises

> Norm-referenced approaches to score interpretation compare students to students to provide a sense of their relative performances. Conversely, our *SEL Competency-Referenced Performance Framework* is based on a criterion-referenced approach, where all students are compared to the same criteria, is more informative about students' competences and is fair because all students have an equal opportunity to achieve at the highest levels.

indicating disproportionate impact of our performance level cut across race/ethnicity or gender groups. This provides evidence that scores from the SSIS SEL Brief are fair and have substantial consequential validity.

Conclusion

The design imperative of *Fair and Unbiased for All Students* launched a comprehensive effort to examine whether or not the SSIS SEL Brief Scales met a high standard for fairness for female and male students from diverse racial/ethnic backgrounds. Based on the definition of fairness and practice standards advanced in the AERA, APA, and NCME (2014) *Testing Standards*, we focused on design and use aspects of the assessments. We learned that the SEL items represented skills and behaviors deemed socially valued and culturally relevant for teachers, parents, and students from diverse groups and they were highly accessible via a universal designed online testing application. We also learned these items exhibited little or no bias based on the ratings by multiple informants of a robust nationally representative sample of students ages 3 to 18. Finally, and perhaps most importantly, we found no systematic disparate impact of the Competency-Referenced Performance Level reporting system on racially and ethnically diverse groups of students. With this initial evidence, we confidently claim that the SSIS SEL Brief Scales yield fair and unbiased results for children ages 3 to 18 from diverse racial and ethnic groups in the United States. As such, these Brief Scales offer efficient alternatives for use in SEL programs designed to advance educational equity and excellence!

References

AERA, APA, & NCME (2014). *Standards for Educational and Psychological Testing.* Washington, D.C.: American Educational Research Association.

Elliott, S.N., Anthony, C.J., DiPerna, J.C., Lei, P-W, & Gresham, F.M. (2020). *SSIS SEL Brief Scales User Guide & Technical Manual.* Scottsdale, AZ: SAIL CoLab.

Ercikan, K., & Elliott, S.N. (2015). *Assessment as a tool for communication and improving educational equity.* Paper commissioned by Smarter Balanced Assessment Consortium. University of California at Los Angeles.

Gresham, F.M., & Elliott, S.N. (2008). *Social Skills Rating Scales: Technical Manual.* Bloomington, MN: Pearson Assessment.

Jagers, R.J., Rivas-Drake, D., & Borowski, T. (2018, November). Equity & social and emotional learning: A cultural analysis. *Measuring SEL.* Chicago: CASEL.

Meade, A. W. (2010). A taxonomy of effect size measures for the differential functioning of items and scales. *Journal of Applied Psychology, 95*(4), 728-743. doi: 10.1037/a0018966

Snyder, T.D., de Brey, C., and Dillow, S.A. (2018). *Digest of Education Statistics 2016* (NCES 2017-094). NCES, IES, U.S. Department of Education. Washington, DC.

Learn More …

Other *Assessment Simplified Briefs* are available at https://ssiscolab.com/

Authors
Stephen N. Elliott, PhD
    Arizona State University
Christopher J. Anthony, PhD
    University of Florida
Contact information: steve_elliott@asu.edu