

Key Design Principles for Direct Assessments of SEL: Lessons Learned from the First Design Challenge

Background

Social and emotional learning (SEL) has been defined many ways. CASEL defines SEL as the process through which children and adults acquire and effectively apply the knowledge, attitudes, and skills necessary to understand and manage emotions, set and achieve positive goals, feel and show empathy for others, establish and maintain positive relationships, and make responsible decisions. SEL encompasses a broad range of skills and behaviors that generally fall into two categories: (a) intrapersonal (e.g., self-management/self-control/self-regulation, beliefs/mindsets/attitudes) and (b) interpersonal (e.g., social awareness, perspective-taking).

Some of these skills and behaviors are not easily inferred from a person's actions, making them a challenge to assess. As a result, few assessments are available that measure social and emotional skills directly through student performance on challenging social-emotional tasks. Instead, most existing social-emotional assessments rely on a reporter who judges behaviors or tendencies related to a child's social and emotional learning skill. Teacher report, for example, measures a teacher's view of the frequency of child behaviors. This method is well-suited to assessing the frequency of clearly observable behaviors. Self-report is well-suited to assessing attitudes in children old enough to understand questionnaires. However, both teacher report and self-report are susceptible to rater biases. Teachers may have different reference groups in mind when rating a child, and children may respond in ways that reflect social expectations more than actual skill level when completing self-report questionnaires. Direct assessments are well-suited to measuring dimensions of social and emotional learning that are not easily observed, such as the thinking skills required to understand what people are thinking and feeling. At the same time, not much work has been done to develop and test direct assessment for broad applied adoption.

AUTHORS

Clark McKown
Rush University
Medical Center

Lindsay Read
CASEL

Noah Bookman
CORE Districts

AWG Design Challenge

A key challenge in the field is, therefore, how to design direct assessments that complement, build upon, and extend the existing measurement approaches. To address this challenge, a committee of the Establishing Practical Social-Emotional Competence Assessments of Preschool to High School Students Work Group project, also called the Assessment Work Group (AWG), created a design challenge seeking submissions of innovative direct assessments of social-emotional skills. Direct assessment of these skills includes all forms of assessment in which social and emotional learning is measured from a child's performance on a challenging task. This includes but is not limited to naturalistic technology-enhanced simulations, performance tasks, game-based data, and live structured social simulations. Direct assessment is distinct from other forms of assessment, such as observations, student self-report, or teacher or parent ratings of students' skills. The design challenge will occur annually, in 2017, 2018, and 2019. Through the design challenge, the committee aims to learn from innovative methods of assessing social-emotional skills and construct and refine a set of key design principles that should guide the development of novel direct assessments and stimulate the development and adoption of social-emotional assessments that support effective instruction and positive student development.

A preliminary vision for a set of design principles was laid out in the first [call for submissions](#), released in spring 2017. Specifically, the committee sought submissions that were: (1) direct assessments of social-emotional learning, which includes all forms of assessment in which SEL is measured from a child's performance on a challenging task, (2) innovative, and (3) aimed at any age range between pre-kindergarten and 12th grade.

The committee was open to assessments in any stage of development, from very early work in assessment development to more fully developed assessment systems. We specified that submissions would be evaluated for:

- The innovativeness of the assessment.
- The clarity of the assessment goals.
- The usefulness of the assessment data for teaching and learning.
- The potential of the assessment to be administered at scale.
- The usefulness of the data reporting.
- The developmental and cultural appropriateness of the assessment.
- The potential of the assessment to be engaging to users.
- The assessment's technical properties, or, for early stage concepts, the potential for those technical properties to be evaluated.

Applicants were asked to submit a narrative proposal of up to five pages and a prototype that was sufficiently detailed to permit the reader to visualize how the assessment works. Proposals were eligible for a cash prize of up to \$5,000, public recognition (featured pieces in blogs, websites, and other media outlets), and inclusion in a professional network dedicated to advancing social-emotional competence assessment.

Results and Analysis of Emerging Themes

In response to the first annual call for submissions, we received 20 submissions. Of these, 15 qualified as direct assessments upon initial review and were reviewed by a panel of outside reviewers consisting of eight researchers and practitioners. Reviewers evaluated and rated each submission on the criteria outlined above. We then held a virtual review meeting to discuss four submissions with scores on the border of winning and with highly discrepant reviewer scores. Final scores for those proposals were submitted during the meeting. Eight design challenge committee members reviewed final scores and selected the top seven submissions for awards. [Winning submissions](#) included:

- A stealth assessment of self-regulation and self-management for grades six through ten in which the test developers quantify rapid guessing while children complete a standardized computerized academic test (Student Assessment Engagement).
- A video-based assessment of perspective-taking for high school in which students in grades six through twelve view brief videos and then answer questions about the peoples' thoughts and feelings (Social Detective).
- A direct assessment for grades six through ten that uses challenging puzzles to measure persistence, effort, resilience, and challenge-seeking (PERC).
- An animated social simulation for third through fifth graders in which children's manner of responding to situations assesses their communication, cooperation, emotion regulation, empathy, impulse control, and social initiation (Zoo U Social Emotional Skills Assessment).
- An online calendaring application for grades six through twelve that measures executive functioning skills such as organization, time management, prioritization, planning, and scheduling by asking students to prioritize and organize a complex set of tasks (The Calendar Task).
- A brief tablet-based battery of tests of inhibitory control, cognitive flexibility skills, and working memory, to yield a broad measure of executive functioning skills for use with PreK through twelfth grade students (PLUS Executive Functioning Assessment).
- A task for students in fourth through twelfth grade that assesses grit through students' choice of easy or difficult math tasks and their response to those tasks (An Incentivized Method for Measuring Grit).

Several themes emerged in our review and evaluation of these submissions:

- There is work going on to develop assessments across the developmental span from PreK through high school.

- Creative work is underway with high potential to fill important SEL needs.
- Submissions reflected a variety of conceptualizations of social-emotional learning, but all focused mainly on social thinking and regulatory processes rather than behavior.
- Most of the winning submissions focused on the assessment of self-regulation.
- Only one assessment measured multiple dimensions of social-emotional skill beyond self-regulation.
- Although all of the winning submissions demonstrated usability, scalability, ease of administration, and implications for teaching and learning, most of the submissions did not demonstrate that implementation is feasible across a wide range of school districts.

We believe that the design challenge successfully sampled innovative approaches to the direct assessment of SEL skills. In so doing, it illustrated vividly the state of the field of direct SEL assessment development. By shining a bright light on some very creative and promising assessment approaches, the design challenge demonstrated that good work is underway. The very existence of these efforts serves as recognition of the need for efficient and practical school-based SEL assessments. This sample of our winning submissions suggests that work remains to create direct assessment systems that reflect a shared and coherent understanding of what SEL is, that assess multiple components of SEL, and that are practically useful for educators.

Updated Design Principles Based on First Annual Design Challenge

The design challenge call for submissions reflected an initial set of design principles, developed by the design challenge committee in consultation with the broader AWG. These principles seemed to the design challenge committee and the broader AWG to reflect highly desirable features in future SEL assessment systems. In reviewing the submissions, and particularly in discussion with our peer review panel, we learned important lessons that led us not so much to add or remove design features to our conceptualization of what is most urgently needed, but to clarify which of those features is most important for an SEL assessment to be ready for broad adoption in applied settings. Three issues rose to the top of the design principle priority list.

First, the assessment should be transparently relevant. While all of the submissions reflected innovative and promising designs, it was not always clear how the skills they measured were relevant to life in school, or tied to the broader move to implement SEL policies and practices. Therefore, we believe that future design challenges in particular, and SEL assessment development initiatives more broadly, ought to be keenly focused on clearly communicating what is being measured and how that thing is related to the reality of classroom life and the broader SEL movement.

Second, assessments should be practical to use. Members of the design challenge committee and the peer reviewers agreed that most of the submissions did not reflect a sufficient degree of usability and feasibility to be used at large scale. The barriers to usability and feasibility reflected in the design challenge submissions were varied. Some submissions required expertise to administer; others took too long to administer; others had challenging scoring systems; others did not clearly report easily understood assessment reports. For many, it was not totally clear that what was being assessed was relevant to the classroom setting. From these insights, the design challenge committee identified important dimensions of usability and feasibility, including:

- Test duration of 30 minutes or less.
- Administration can be completed in the classroom.
- Assessments clearly reflect skills that are used in the classroom.
- Assessments can easily be scaled for use in multiple school settings.
- Administration can be completed easily, with little or no training.
- Assessments can be scored automatically and provide immediate feedback.

Third, the assessment data should guide practice. Despite their innovative qualities, most of the submissions, including the winning submissions, were not yet proven to inform practice. The best assessment system, for example, would identify children's social-emotional strengths and needs in a way that teachers can use to make decisions about curriculum use and instructional practice. The best assessment systems would tie assessment results directly to resources such as lesson plans and online training opportunities designed to help teachers address needs identified by the assessment.

Conclusion and Next Steps

The first year of the AWG Design Challenge was a success, in particular in shining a light on the state of development of direct measures of social-emotional competence. As we move forward, we will work to integrate lessons learned about design principles from the first year into the next round of design challenges. In particular, we envision revising the call for submissions to target submissions that meet the field's need for relevance, practicality, and links to practice.

We will also consider segmenting the design challenge into two or more focus areas. We may wish to prioritize submissions that focus on a particular age group, skill or competency, task type, or phase of development. For example, we may wish to explicitly encourage assessments that measure interpersonal competencies such as social awareness and perspective-taking. In addition, we received very few submissions from practitioners (e.g. school and district administrators and teachers). Nearly all of our winners were researchers. Given that practitioners are selecting and using assessments, we will consider whether and how to reach out to practitioners who may be in a position to contribute to or lead a design challenge submission.

Ultimately, we believe that breakthrough systems of assessing social-emotional competence will require sufficient resources and collaborative contributions from the best minds in SEL and assessment development. Beyond contributing a set of design principles, therefore, we wish to determine what conditions would be necessary to engender true breakthroughs in assessment development and application. There is precedent for such an effort. The National Institutes of Health invested significant resources to bring together the best minds in SEL, assessment development, and education to develop the NIH Toolbox, a set of common neurodevelopmental assessment instruments for clinical and research purposes that have outstanding psychometric properties and are purpose-built for research and clinical practice.

Beyond design principles, the field needs an effort of this magnitude and level of ambition if we are to fulfill the promise that this year's winning submissions reflect. The AWG is laying the groundwork for such an investment by catalyzing the development of innovative and practical direct assessments through its annual design challenge and the creation of guiding principles of SEL direct assessment design.

The AWG is the working group for the Establishing Practical Social-Emotional Competence Assessments of Preschool to High School Students project, which was created to advance progress toward establishing practical SEL assessments that are scientifically sound, feasible to use, and actionable as a key priority for the field. The project's work group is a multidisciplinary collaborative of leading researchers and practitioners in the fields of PreK-12 education, assessment, social and emotional learning (SEL), and related fields. As part of a three-year effort, members of the work group will work in partnership to make key advancements in student SEL assessment.

To learn more, please visit <http://measuringSEL.casel.org/design-challenge/>

Table 1. Description of winning Design Challenge submissions

Title	Primary Applicant Name and Affiliation	Skill Area(s) Assessed	Age Range	Description of Assessment
<u>Student Assessment Engagement</u>	James Soland, NWEA	Self-regulation and self-management	6 th - 10 th grade	When students take an achievement test on a computer, metadata like the amount of time spent on each item are often collected. Research shows that students who often respond extremely fast – so quickly they could not have understood the item’s content – are likely disengaged from the test. Our measure quantifies how often students respond extremely quickly over the course of a test, which is strongly correlated with scores from measures of social-emotional learning constructs like self-regulation and self-management.
<u>Social Detective</u>	Panorama Education	Perspective-taking	6 th - 12 th grade	Panorama’s Social Detective is designed to measure and help students practice social perspective-taking, a malleable and central social competency that underlies a vast range of social-emotional functioning at school and in life. In this performance task, students are challenged to be a “social detective” whose job is to figure out other people’s values, interests, and perspectives. After watching short video interviews, students answer a series of questions to gauge how well they perceive and understand each person.
<u>PERC</u>	Tenelle Porter, University of California, Davis	Persistence, effort, resilience, and challenge-seeking	6 th - 10 th grade	The PERC is a computer-based tool that assesses students’ Persistence, Effort, Resilience, and Challenge-seeking behavior. These are key behavioral expressions of a growth mindset of intelligence.

Table 1. Description of winning Design Challenge submissions

Title	Primary Applicant Name and Affiliation	Skill Area(s) Assessed	Age Range	Description of Assessment
<u>Zoo U Social Emotional Skills Assessment</u>	Melissa E. De-Rosier, 3C Institute and Centervention	Communication, cooperation, emotion regulation, empathy, impulse control, and social initiation	3 rd - 5 th grade	Zoo U provides a game platform for performance-based formative assessment of social-emotional skills in upper elementary grades.
<u>The Calendar Task</u>	Patricia Inglese, Educational Testing Service	Self-management, specifically: organization, time management, prioritization, planning, and scheduling	6 th - 12 th grade	The calendar task (in development) will potentially provide a flexible and naturalistic platform to evaluate student self-management skills with the potential for multiple use cases, both summative and formative in nature.
<u>PLUS Executive Functioning Assessment</u>	Jelena Obradovic, Stanford University	Executive function skills as represented by inhibitory control, working memory, and cognitive flexibility	PreK - 12 th grade	In order to employ direct assessments of executive function (EF) skills at scale, we developed a group-based assessment procedure that is time-efficient and cost-effective. We adapted four developmentally appropriate, widely used EF tasks for administration on tablet computers in a classroom setting. Our classroom protocol allows a minimally disruptive assessment of EF skills in all students at the same time.
<u>Measuring Grit</u>	Seda Ertac, Koc University	Grit	4 th - 12 th grade	This assessment measures grit, i.e., the propensity to set ambitious goals, persevere in the face of failures, and put effort to build skill. We use an incentivized methodology that involves rewarding successful outcomes.