

## **MOOSES: A framework for categorizing SEL Measures**

### **Background**

In recent years, interest in social and emotional learning (SEL) has burgeoned as both education policy and research have underscored the importance of SEL skills for success in school, employment, and civic life (Heckman & Rubinstein, 2001; Heckman, Stixrud, & Urzua, 2006). While much research has been done on effective SEL design and implementation (Corcoran, Cheung, Kim, & Xie, 2018; Graczyk et al., 2000), less is known about effective SEL measurement, which is essential for program evaluation.

### **Purpose**

The purpose of this paper is to present a framework for categorizing SEL measures: Measureable, Observable, Objective Social Emotional Skills (MOOSES). While traditional academic outcomes are often measured with standardized and norm-referenced tests and measures of specific skills, SEL outcomes are often more conceptually complex and thus difficult to measure. The MOOSES framework proposes typologies for SEL measurement that are based on how measurable, observable, and objective these measures are. A multitude of SEL measures exist, and the MOOSES framework is an approach for practitioners and researchers to determine appropriate SEL measures for different contexts.

### **Approaches to measuring social emotional skills**

There are a number of ways researchers and practitioners measure social emotional skills. World Learning (World Learning, 2018) summarizes common methods of measuring social and emotional skills as observations, project-based, and written or oral assessments such as knowledge-based, psychometric assessment of personality or attitude, and scenario-based measures. Compendia of SEL measures that highlight administrative features and psychometric properties have also been developed including the RAND Education Assessment Finder and Youth Power Action Inventory of Soft Skills Measurement.

However, to date, there is little guidance on which types of measures are least prone to bias, especially bias in program evaluations that may favor treatment groups, leading to misleading program effects and increased Type 1 error. Some of the most common SEL measures are self-reported questionnaires (Duckworth & Yeager, 2015, Galloway et al., 2017), but these suffer from a number of well-known limitations including heterogeneous frames of reference, social desirability bias, and acquiescence bias (Galla et al., 2014). In particular, teachers and students know they are in an experiment and may report positive perceptions to enhance their status with the experimenters. To address these limitations – especially those biases that may systematically favor the experimental group in program evaluations – this paper proposes the MOOSES framework for categorizing SEL measures according to their potential to assess students' SEL skills with meaningful outcomes and minimal possibilities of bias.

### **A conceptual framework**

MOOSES are quantifiable measures that can be observed objectively and are unlikely to change simply because staff or students know there is an experiment taking place. For example, records of disciplinary referrals or blinded observations of student behavior are acceptable indicators of behavior because they are observed objectively regardless of treatment status; conversely, teacher reports or student self-reports of behavior may not be objective if they are

not independent of the treatment. Teachers implementing the program may be more prone to rate students favorably for the sake of the program evaluators (social desirability) or because of the time and effort they invested in the program, even if there is no actual perceived improvement in students' behaviors.

Measures may also be overaligned to the treatment students receive such that they are simply repeating back what they have been taught during the intervention instead of actually undergoing a behavior change. For instance, in studies of programs designed to improve students' social skills, intervention teachers who are taught to identify specific social skills may be more likely than teachers in the control group to report that their students display these social skills. Such ratings may not necessarily reflect greater social skills, but may be reported because teachers in intervention groups simply have learned a way of identifying a specific set of social behaviors. This could be seen as an example of "parroting," in which subjects or involved raters (such as teachers or parents) have learned what to say or identify narrow behaviors at the time of rating, instead of truly changing behaviors or attitudes.

In response to these issues, the MOOSES framework categorizes types of measures that are meaningful to measure in school settings based on how closely they are meaningful, observable, and objective social emotional skills that are least prone to biases such as parroting or social desirability. The following lists the different levels, starting from the least MOOSES types of measures:

1. **Knowledge of SEL skills:** assess respondents' self-reported knowledge of specific skills (e.g., "I know how to tell when my friend is sad")
2. **Attitudes towards SEL:** assess respondents' beliefs and attitudes towards SEL competencies (e.g., "Aggression is wrong")
3. **Intention for SEL behaviors:** assess what respondents would do in hypothetical or future scenarios (e.g., scenario-based assessments)
4. **Self-reports of SEL behaviors:** assess self-reported behaviors indicating SEL competencies. This category is further divided into two subcategories based on whether the respondent has potential reasons to systematically favor positive program outcomes.
  - a. **Interested.** These could include teachers or principals as assessors who may have developed a positive relationship with researchers or need to positively report on the intervention to increase chances for extended funding. (e.g. "This student treats others with respect")
  - b. **Disinterested.** Those who have no vested interest in showing the program evaluation in a positive light, such as parents who are not involved in treatment implementation or students themselves (e.g., "I treat others with respect")
5. **Observations and performance tasks:** assess behaviors as observed by independent observers, ideally those who are unaware of treatment assignment, such as trained observers or school staff who would report on student behaviors as part of their normal practice (e.g., reports of disciplinary action or standardized assessments)

### **Practical Application**

We acknowledge that bias can exist at any level of measurement. There are also obvious tradeoffs between more rigorous methods of measurement and the time and resources required. However, the MOOSES framework attempts to provide practitioners and researchers with an

approach to evaluating and selecting appropriate measurement methods for program monitoring and evaluation of SEL programs that maximize meaning and minimize bias.

## References

- Corcoran, R. P., Cheung, A. C., Kim, E., & Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: a systematic review and meta-analysis of 50 years of research. *Educational Research Review, 25*, 56-72.
- Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher, 44*(4), 237-251.
- Galla, B. M., Plummer, B. D., White, R. E., Meketon, D., D'Mello, S. K., & Duckworth, A. L. (2014). The Academic Diligence Task (ADT): Assessing individual differences in effort on tedious but important schoolwork. *Contemporary educational psychology, 39*(4), 314-325.
- Galloway, T., Lippan, L., Burke, H., Diener, O, & Gates, S. (2017). Measuring Soft Skills & Life Skills in International Youth Development Programs: A Review and Inventory of Tools. Washington DC: USAID's YouthPower Implementation IDIQ- Task Order 1, YouthPower Action.
- Graczyk, P. A., Weissberg, R. P., Payton, J. W., Elias, M. J., Greenberg, M. T., & Zins, J. E. (2000). Criteria for evaluating the quality of school-based social and emotional learning programs.
- Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review, 91*(2), 145-149.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics, 24*(3), 411-482.
- World Learning (2018). *Soft Skills Development: Guiding Notes for Project and Curriculum Design and Evaluation*. Retrieved from <https://www.youthpower.org/resources/soft-skills-development-guiding-notes-project-and-curriculum-design-and-evaluation>