

# Measuring MESH:

Student and Teacher Surveys Curated for the CORE Districts

A downloadable guide from Transforming Education

---

**v1.0 April 2016**



In partnership with the CORE Districts



## Table of Contents

<b>Overview</b> .....	<b>3</b>
<b>Survey Research and Development</b> .....	<b>4</b>
<b>Background on the CORE Districts</b> .....	<b>4</b>
<b>Defining the Competencies</b> .....	<b>5</b>
Self-Management .....	5
Self-Efficacy .....	5
Growth Mindset.....	6
Social Awareness .....	6
<b>Curating the Measures</b> .....	<b>7</b>
<b>Pilot Testing</b> .....	<b>8</b>
<b>Evidence Related to Survey Quality</b> .....	<b>9</b>
<b>Reliability</b> .....	<b>9</b>
<b>Validity</b> .....	<b>12</b>
<b>Comparability Across Schools</b> .....	<b>15</b>
<b>Concluding Remarks</b> .....	<b>17</b>
<b>FAQs</b> .....	<b>17</b>
<b>Surveys</b> .....	<b>21</b>
<b>Sample Administration Protocols</b> .....	<b>21</b>
<b>Student Self-Report Survey on MESH Competencies</b> .....	<b>23</b>
<b>Teacher Report on Students' MESH Competencies</b> .....	<b>26</b>

## OVERVIEW

Rigorous longitudinal research demonstrates that certain Mindsets, Essential Skills, and Habits (MESH) can help students succeed in college, career, and life. Transforming Education (TransformEd) believes that assessing students' MESH competencies can give educators valuable information that helps them understand students' strengths and identify the areas in which students need additional support. When used in this way, surveys assessing students' MESH competencies can help educators tailor their instruction and classroom environment to serve students more effectively. Such surveys can also help educators identify which of the many practices already in use to build students' MESH skills seem to be working best.

In partnership with California's [CORE Districts](#), TransformEd has curated a set of survey-based MESH measures that were originally developed by leading researchers and later field tested with nearly half a million students. The data from that field test were analyzed by our partners at the [Harvard Center for Education Policy Research](#), who concluded that the measures used were internally reliable and were significantly correlated with students' GPA, test scores, attendance, and suspension rates. (See the FAQ section below and ["Should non-cognitive skills be included in school accountability systems?"](#) for more details.)

TransformEd fully supports the use of MESH measures to inform instruction and programming. We also recognize that some states and districts – like the CORE Districts - may choose to include results from these surveys in their school-level accountability systems. It is important to note that CORE's accountability system is unique in that it is a "low stakes" system developed by the districts themselves to focus on supporting continuous school improvement. As Rick Miller, Executive Director of the CORE Districts describes, "Our model is about getting better, not doling out punishments." (See p. 4 for more details on CORE's model and their use of MESH measures.)

Both TransformEd and CORE recommend that the surveys NOT be used for high-stakes decisions about individual students or educators. This set of measures is still in the early stages of development and that evidence regarding its validity is just beginning to emerge. Additionally, it is important to note that, as with all survey-based measures, these measures may be subject to various forms of intentional gaming, particularly when used in a high-stakes setting. Finally, we acknowledge that validation is an ongoing process and will continue to work with our school partners and research partners to collect, assess, and publish further data on the validity of these measures for various purposes.

To inquire about how TransformEd can assist your school, district, or state in identifying, field testing, and employing measures of students' MESH competencies, please reach out to us at 617-378-3939 or [MeasuringMESH@transformingeducation.org](mailto:MeasuringMESH@transformingeducation.org).

If you have questions or feedback about the measures or this user guide, or if you want to share your own approach to incorporating MESH in your district or school, please e-mail us at [MeasuringMESH@transformingeducation.org](mailto:MeasuringMESH@transformingeducation.org). We'd love to hear from you!

To stay abreast of new developments in the MESH field, please [subscribe](#) to our newsletter and follow us on Twitter ([@Transforming\\_Ed](#)).

### BACKGROUND ON THE CORE DISTRICTS

The CORE Districts (CORE) are a group of nine California school districts that are committed to working together to identify, implement, and scale new strategies that help students succeed. Together, CORE represents over one million students and 1,500 schools across Fresno, Garden Grove, Long Beach, Los Angeles, Oakland, Sacramento, San Francisco, Sanger, and Santa Ana Unified School Districts. Six of these districts are operating under a No Child Left Behind (NCLB) waiver granted in August 2013 to implement a holistic system of accountability and continuous improvement.<sup>1</sup> At the heart of this system lies the School Quality Improvement Index (SQII), which measures school quality as a function of students' academic outcomes, students' social-emotional (or MESH) skills, and school culture/climate. Survey-based MESH measures count for 8% of the total 100-point score assigned to each school under CORE's SQII. A school that performs poorly on the overall Index is paired with a higher performing school that provides mentorship and support. As such, CORE's SQII focuses on capacity-building and continuous improvement, rather than the punitive consequences that many traditional accountability systems employ.

After receiving their NCLB waiver in August 2013, CORE sought to determine which set of social-emotional (or MESH) competencies to include in their accountability system. In November 2013, CORE convened representatives from each of the participating districts, ranging from superintendents to directors of student support, directors of social-emotional learning, and directors of special education. Content experts in MESH also joined this convening from the Collaborative for Academic, Social and Emotional Learning (CASEL), the John W. Gardner Center for Youth at Stanford, and TransformEd.

TransformEd proposed three key criteria for the CORE Districts to consider when selecting MESH competencies: each competency needed to be meaningful, measureable, and malleable (i.e. pass the "3M's test").<sup>2</sup> Meaningful indicates that a particular competency is predictive of important academic, career, and life outcomes for students. Measureable indicates that the competency can be measured reliably through an assessment that is feasible to administer at scale in school settings. Malleable means that there is evidence to suggest that the competency can be developed through in-school interventions. Additionally, CORE prioritized identifying at least one *intrapersonal* skill and one *interpersonal* skill in its initial set of MESH competencies to ensure that there was a broad range of competencies represented.

Using these criteria, the CORE Districts voted to prioritize four specific MESH competencies for inclusion in the School Quality Improvement Index: growth mindset, self-efficacy, self-management, and social awareness. While neither CORE nor TransformEd believed this to be a comprehensive set of MESH competencies, both felt that it was a strong starting point for the CORE Districts' collective work in that it was informed by research and reflects the priorities of the participating districts.

---

<sup>1</sup> These districts include Fresno, Long Beach, Los Angeles, Oakland, San Francisco, and Santa Ana.

<sup>2</sup> CORE uses the terminology meaningful, measurable, and actionable.

## DEFINING THE COMPETENCIES

A definition of the four MESH competencies prioritized by the CORE Districts is included below, along with a brief summary of the research related to each competency. For a more detailed summary of the research on the importance of MESH skills, please see [Ready to Be Counted](#).

### Self-Management

Self-management, also referred to as “self-control” or “self-regulation,” is the ability to regulate one’s emotions, thoughts, and behaviors effectively in different situations. This includes managing stress, delaying gratification, motivating oneself, and setting and working toward personal and academic goals.<sup>3</sup> Students with strong self-management skills arrive to class prepared, pay attention, follow directions, allow others to speak without interruption, and work independently with focus.

Self-control in children has been linked to a wide range of adult outcomes ranging from high school and college completion to physical health and financial stability. For example, one recent study demonstrates that students with high self-control are significantly more likely to earn a high school diploma and be financially stable as adults. Those with high self-control are also significantly less likely to struggle with depression, obesity, and substance abuse as adults.<sup>4</sup>

### Self-Efficacy

Self-efficacy is the belief in one’s ability to succeed in achieving an outcome or reaching a goal. Self-efficacy is related to whether a student believes that s/he has sufficient control over his/her environment in order to succeed. High self-efficacy reflects confidence in the ability to exert control over one’s own motivation, behavior, and environment and allows students to become effective advocates for themselves.<sup>5</sup>

Decades of research show that self-efficacy is a valid predictor of students’ motivation and learning. Students with high levels of self-efficacy participate more in class, work harder, persist longer, and have fewer adverse emotional reactions when encountering difficulties than students with lower self-efficacy.<sup>6</sup> High self-efficacy can also motivate students to use specific learning strategies and to engage in self-directed learning.<sup>7</sup>

---

<sup>3</sup> CASEL.org (<http://www.casel.org/social-and-emotional-learning/core-competencies/>)

<sup>4</sup> See for example Moffitt et al. (2011) A gradient of childhood self-control predicts health, wealth, and public safety; Knudsen, Heckman, Cameron & Shonkoff (2006) Economic, neurobiological, and behavioral perspectives on building America’s future workforce; McClelland, Piccinin, Acock & Stallings (2011) Relations between preschool attention and later school achievement and educational outcomes

<sup>5</sup> Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman; Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary educational psychology*, 25(1), 82-91.

<sup>6</sup> Bandura, A. (1997); Zimmerman, B. J. (2000).

<sup>7</sup> Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of educational Psychology*, 82(1), 51; Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American educational research journal*, 29(3), 663-676.

## Growth Mindset

Growth mindset is the belief that one's abilities can grow with effort. Students with a growth mindset believe that they can develop their skills through effort, practice, and perseverance. These students embrace challenges, see mistakes as opportunities to learn, and persist in the face of setbacks. By contrast, students with a fixed mindset believe that their own intelligence and talent are innate traits that don't change. These students typically worry about not looking smart, get upset by mistakes, and give up sooner on tough tasks.<sup>8</sup>

Multiple longitudinal studies have shown that growth mindset is linked to increased motivation, better grades, and higher test scores.<sup>9</sup> Studies have also shown that students with a growth mindset work harder and spend more time on a subject instead of giving up when things get difficult. Teachers report that students who receive an intervention to develop growth mindset demonstrate greater effort and interest in the classroom. Research suggests that having a growth mindset is particularly important during transitions (e.g. from elementary to middle school or middle to high school), since students often meet with new challenges during these times.<sup>10</sup>

## Social Awareness

Social Awareness is 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.<sup>11</sup>

Social awareness and the related skill "social competence" increase students' chance of succeeding academically. Students who demonstrate strong social awareness are able to engage in constructive communication with their peers and resolve conflicts when they arise. These students benefit from peer learning and know how to take advantage of social supports.<sup>12</sup> A recent longitudinal study demonstrated that, even after controlling for socioeconomic status and early academic ability, higher social competence in kindergarten led to higher odds of graduating from high school and college.<sup>13</sup> Social awareness and social competence are also widely established as important factors in workforce success. Higher social competence in kindergarteners predicted stable employment at age 25,<sup>14</sup> and a recent employer survey conducted by the Partnership for 21<sup>st</sup> Century Skills demonstrates that four of the five most important skills for high school graduates entering the work force are linked to social awareness: professionalism, collaboration, communication, and social responsibility.<sup>15</sup>

---

<sup>8</sup> Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, 78(1), 246-263.

<sup>9</sup> *Ibid.*

<sup>10</sup> Blackwell, L. S., Trzesniewski, K. H. & Dweck, C. S. (2007); Dweck, C. S., Walton, G. M., & Cohen, G. L. (2011) Academic Tenacity: Mindsets and Skills that Promote Long-Term Learning. Bill and Melinda Gates Foundation. Retrieved from: <https://ed.stanford.edu/sites/default/files/manual/dweck-walton-cohen-2014.pdf>

<sup>11</sup> CASEL.org (<http://www.casel.org/social-and-emotional-learning/core-competencies/>)

<sup>12</sup> Gehlbach, H., Young, L. V., & Roan, L. K. (2012). Teaching social perspective taking: how educators might learn from the Army. *Educational Psychology*, 32(3), 295-309.

<sup>13</sup> Jones, D. E., Greenberg, M., & Crowley, M. (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. *American Journal of Public Health* 105(11), 2283–2290.

<sup>14</sup> Jones, D. E., Greenberg, M., & Crowley, M. (2015).

<sup>15</sup> Casner-Lotto, J., & Barrington, L. (2006). Are They Really Ready to Work? Employers' Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century US Workforce. *Partnership for 21st Century Skills*. 1 Massachusetts Avenue NW Suite 700, Washington, DC 20001.

## CURATING THE MEASURES

In order to identify the most promising measures of these four competencies, TransformEd sought guidance from leading experts in the MESH field, including Lisa Blackwell, Clancy Blair, Eduardo Briceno, Celene Domitrovich, Angela Duckworth, Carol Dweck, Camille Farrington, Greg Walton, Roger Weissberg, and David Yeager. We then curated a set of measures that had been developed by researchers and that were free and practical for the CORE Districts to administer at scale. Finally, TransformEd and CORE worked with Dr. Hunter Gehlbach to modify the original scales in ways that would mitigate common survey issues and then compared the performance of the original and modified versions during an initial pilot test in 2014.

Exhibit 1 below shows which researchers' work influenced the current student self-report and teacher report scales most significantly. Please note that while these experts were consulted, any errors in fact and interpretation are the responsibility of Transforming Education. We at Transforming Education take sole responsibility for the views we express and the materials we develop. Additionally, these researchers have continued to develop and refine their own measures over the past few years, while the CORE Districts have chosen to continue using the exact survey items they piloted in 2014 and field tested in 2015 in order to ensure that their year over year data would be comparable. The measures included in this document are the ones that CORE is currently using, which may not reflect the most recent updates made by the original contributors of each scale.

**Exhibit 1. Key Experts Consulted for Scales**

	Student Self-Report (Grades 5-12)	Teacher Report (Grades K-12)	Key Experts Consulted
Self-Management	x	x	Angela Duckworth (University of Pennsylvania)  Clancy Blair (New York University)
Social Awareness	x	x	CASEL  American Institutes of Research
Self-Efficacy	x		Camille Farrington (Consortium on Chicago School Research - CCSR)
Growth Mindset	x		Camille Farrington (CCSR)  Carol Dweck (Stanford University)

## PILOT TESTING

CORE conducted a pilot test of the MESH measures in spring 2014 with approximately 9,000 students and over 300 teachers. During this pilot, two different forms of the student self-report survey and teacher survey were randomly assigned to participants. For each competency, one of the forms used the original measure developed by a researcher and the other form provided a modified version developed in partnership with Dr. Gehlbach to reflect emerging best practices in survey design (e.g. removing double-barreled items, translating statements into questions, etc.). CORE, TransformEd, and researchers at Harvard CEPR compared the two forms of each measure to identify the more promising form based on the following criteria: 1) correlations with validating scales (construct validity); 2) correlations with external variables (criterion validity); and 3) internal reliability. The scales that performed the best across these three criteria were then tested at much greater scale during the 2015 field test. (See page 10 for detailed validity evidence from the field test.)<sup>16</sup>

The pilot test also presented an opportunity to address three different forms of potential bias that commonly affect survey-based measures. First, the CORE Districts were concerned about reference bias, or the tendency for respondents to interpret a given survey scale differently depending on their individual *frame of reference*, which may be related to the culture of their school. Along with the surveys themselves, districts piloted anchoring vignettes for self-management and social awareness that had been developed by ETS to mitigate the effects of reference bias. Anchoring vignettes is a technique that uses students' ratings of hypothetical individuals who exhibit varying levels of the target competency as a tool for norming and adjusting students' self-ratings. Research suggests that observable behaviors or competencies (e.g., self-management and social awareness) are more subject to reference bias, whereas, self-efficacy and growth mindset are understood to be internal mindsets or beliefs that cannot be observed. As such, the latter are less susceptible to reference bias, and anchoring vignettes are less likely to be necessary for these competencies.

In addition to reference bias, CORE was also concerned about the potential for stereotype threat to impact students' responses. Stereotype threat refers to the tendency for survey responses to be influenced by the respondents' perception of how people in their group (e.g. racial, ethnic, or socio-economic class) are believed by others to perform in a particular domain.<sup>17</sup> For example, when female students were asked to report their gender before taking a math test, they performed worse on that test than their peers with similar levels of math skill, which provides evidence that the participating students may have internalized negative stereotypes about women's performance in math.<sup>18</sup> The CORE districts addressed this concern by including demographic questions only at the *end* of the survey or by removing all demographic questions from the survey and using a bar code as a confidential student identifier.

Finally, CORE was concerned with social desirability bias, or the tendency for survey responses to be influenced by social pressures.<sup>19</sup> For example, if a survey asks how often a student is polite to adults, the student may answer "almost all the time," even if the response "almost never" more accurately reflects her behavior because she knows that it is socially desirable to be polite to

---

<sup>16</sup> Please contact us if you'd like more information on pilot test results for each measure and criterion.

<sup>17</sup> See e.g., Spencer, S., Steele, C., & Quinn, D. (1999). Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, 35(1), 4-28. doi:10.1006/jesp.1998.1373

<sup>18</sup> *Ibid.*

<sup>19</sup> See e.g., Podsakoff, P., MacKenzie, S., Lee, J., & Podsakoff, N. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88(5), 879-903. doi:10.1037/0021-9010.88.5.879



adults. The CORE Districts attempted to mitigate this issue in two ways: 1) by communicating that all students' survey responses would remain confidential and would not influence grades or other assessments of the students' performance; and 2) by asking the adults who proctored survey administration to stand at the back of the classroom so that students felt they could complete their survey without being judged by a teacher or other adult.

## EVIDENCE RELATED TO SURVEY QUALITY

The validity of an assessment is a function not only of the properties of the assessment but also of the intended uses of the assessment results. Therefore, when validating an assessment, many different sources of evidence should be brought to bear.<sup>20</sup> We present preliminary evidence on the following aspects of technical quality of the measures assessing students' MESH skills:

- **Reliability:** consistency of student scores;
- **Validity:** the degree to which an assessment measures what it purports to measure; and
- **Comparability Across Schools:** the degree to which student scores are interchangeable across schools.

We recognize that validation is an ongoing process. Therefore, we will continue to collect, analyze, and publish additional data on the technical quality of these measures, with a particular focus on the validity of measures across different subgroups and as used for different purposes.

Below, we summarize the validity, reliability, and comparability evidence on student and teacher surveys administered in CORE districts in school years 2013-14 and 2014-15. This evidence comes primarily from the field test conducted with more than 450,000 students in California's CORE districts during spring 2015. The sample of students included in the analyses shown here was further limited to those who completed the full survey and those for whom complementary administrative data was available. This included students in grades 3-12 across five districts in California.<sup>21</sup> We also provide evidence from the pilot test (conducted in spring 2014), where applicable.

### RELIABILITY

With respect to reliability, we examine the extent to which student ratings of each individual skill are internally consistent, as well as the consistency of student and teacher ratings of the same skill. The most common method for evaluating internal consistency is with a statistic called "Cronbach's Alpha". Cronbach's alpha can range from 0 to 1, with higher values suggesting that student results across items are more consistent (i.e. higher reliability). With standardized academic assessments, which typically contain a large number of items and are typically used for high-stakes purposes, we expect to find internal consistency values of between 0.80 and 0.90. Values lower than 0.70 can suggest that the assessment results contain more error variance than

---

<sup>20</sup> Kane, M. (2006). Validity in Educational Assessments. In Brennan, R. L., (ed) *Educational Measurement, 4<sup>th</sup> Edition*. Portsmouth NH: Praeger

<sup>21</sup> Note that surveys were administered primarily in grades 5-12, with the exception of 2 districts that administered the surveys in earlier grades. Administrative data were not available for one of the CORE waiver districts, which was therefore excluded from the analyses.

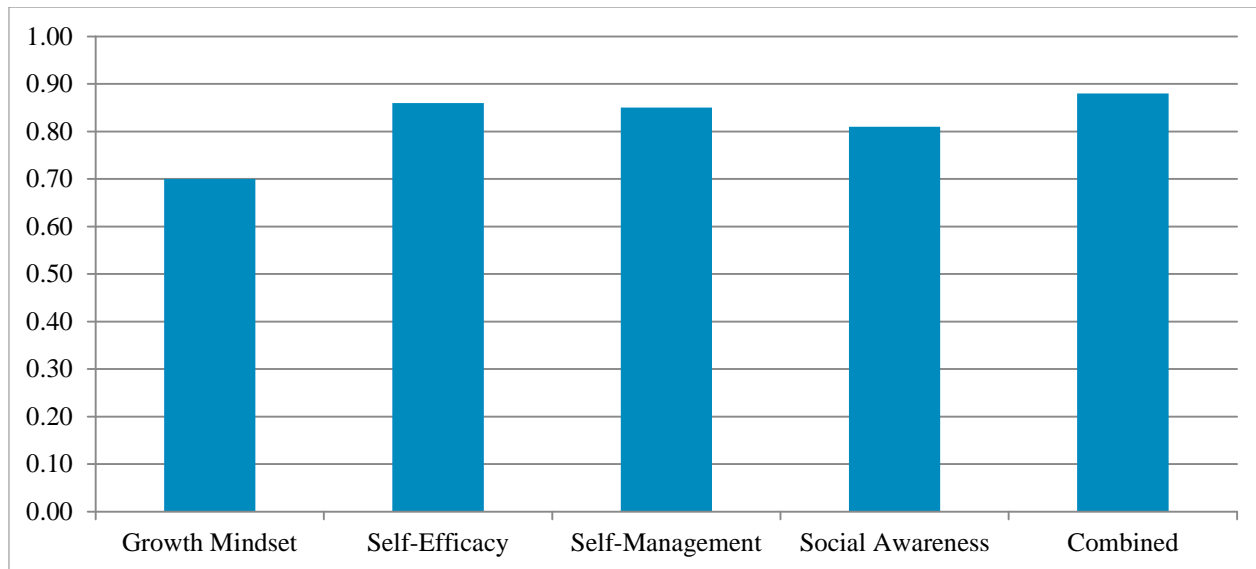
is often desirable; however low estimates can often be a function of short assessments with few items.

Interrater reliability is another form of reliability and is used to assess the degree of convergence among multiple raters of a particular competency. We can use the fact that (a) both teacher and student surveys assessed some of the same student competencies, and (b) multiple teachers assessed an individual student’s MESH competencies, to provide insight into the consistency in the way different raters interpret each measure and thus into the amount of consistency in scoring of student competencies based on these measures. Below, we summarize the reliability of the measures assessing student MESH skills in the CORE Districts’ 2014-15 field test.

### INTERNAL CONSISTENCY

As demonstrated in Exhibit 5, below, the measure for each competency meets internal consistency standards with coefficients at or above 0.70 based on student survey ratings. The self-efficacy measure demonstrates the highest level of reliability across the four measures, with an internal consistency estimate of 0.87, while the growth mindset measure exhibits the lowest level, with an internal consistency estimate of 0.70. The lower internal consistency among items on the growth mindset survey appears to be driven by the surveys administered in the lower elementary grades (i.e., 3<sup>rd</sup> and 4<sup>th</sup> grade). For this reason, we do not recommend that the measures be used for students younger than 5<sup>th</sup> grade. The average rating across the four measures “combined” has an internal reliability estimate in line with that of standardized assessments widely used in schools, 0.88.

**Exhibit 5. Overall sample reliability of student survey measures<sup>22,23</sup>**



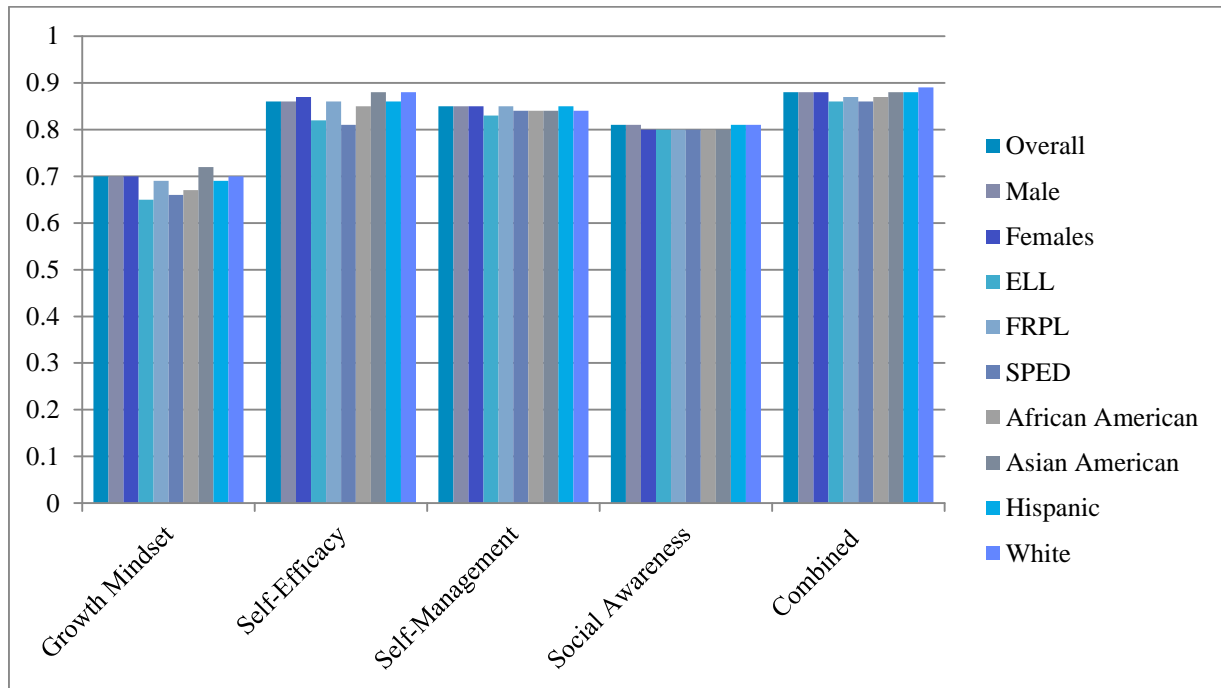
<sup>22</sup> Results are based on field tests across all but one CORE district, with a total student sample size of 378,456. The reliability estimates are derived using Cronbach’s alpha. Each survey measure consists of 4 to 9 items with the exception of the combined MESH measure, where the estimate is calculated using every item from each of the four measures.

<sup>23</sup> Source: Analyses conducted by CEPR for CORE Districts and Transforming Education.

## INTERNAL CONSISTENCY ACROSS SUBGROUPS

The reliability estimates remain high and fairly consistent across student subgroups (Exhibit 6). The exception is with the growth mindset measure, for which the internal consistency within select student subgroups --including students eligible for free and reduced-priced lunch, English Language Learners, students with disabilities, African American students, and Hispanic students - falls just below the 0.70 benchmark.

**Exhibit 6. Overall sample reliability of student self-report measures by student subgroup<sup>24,25</sup>**



## INTERRATER RELIABILITY BETWEEN STUDENT AND TEACHER RATINGS

In our preliminary examination of interrater reliability, we report correlations between student and teacher scores. While in future analyses we plan to add robustness to our investigation of interrater reliability, correlational analyses can provide preliminary estimates of the consistency between student and teacher reports (i.e., the extent to which the two sets of scores vary together). There were 31,828 cases across 166 schools in which students and teachers were linked (i.e., at least one teacher reported on a student who also provided a self-report). We find a moderate to strong relationship between student self-reports and teacher reports, as demonstrated in Exhibit 7. At the school level, the correlation between student and teacher reports on the self-management competency is 0.40 among elementary schools in the sample, and 0.74 among middle and high schools. Similarly, for social awareness, the school-level correlation between student and teacher self-reports is 0.35 at the elementary school level, 0.64

<sup>24</sup> Results are based on field tests in each of the CORE districts. The reliability estimates are derived using Cronbach's alpha. Each survey measure consists of 4 to 9 items with the exception of the combined MESH measure, where the estimate is calculated using every item from each of the four measures.

<sup>25</sup> Source: Analyses conducted by CEPR for CORE Districts and Transforming Education.

at the middle school, and 0.73 at the high school level. The lower correlations found in elementary school appear may be driven by the participation of 3<sup>rd</sup> and 4<sup>th</sup> graders, which we do not recommend. The stronger correlations in middle and high school suggest that student survey measures and teacher survey measures in these grades provide complementary information on student MESH competencies.

**Exhibit 7. The School-Level Relationship between Student Ratings and Teacher Ratings<sup>26</sup>**

	Self-Management	Social Awareness	N
Elementary Schools	.40	.35	76
Middle Schools	.74	.64	76
High Schools	.74	.73	14

**INTERRATER RELIABILITY AMONG MULTIPLE TEACHER RATINGS**

A sub-sample of students (9,165) had multiple teachers rate their MESH competencies, which enables us to examine the interrater reliability of teacher ratings (i.e. the degree to which one teacher’s assessment correlates with another teacher’s assessment of a given student’s competencies). Among students with multiple teacher ratings, the interrater reliability of teacher reports was 0.50 for the self-management measure and 0.38 for the social awareness measure. Further analyses of rater agreement will contribute to determinations of the consistency of scoring across teachers.

**VALIDITY**

**CONVERGENT VALIDITY EVIDENCE**

Evidence of validity, or the degree to which each MESH measure is assessing the underlying construct it was developed to assess, comes primarily from the 2014 pilot. Students responded to Form A or Form B of the primary MESH surveys being considered for inclusion in CORE’s School Quality Improvement Index and to complementary validating scales that were intended to assess related competencies. For example,

- A survey of emotional regulation was used as an validating scale for self-management;
- A survey of social perspective taking effort was used as a validating scale for social awareness;
- A survey of classroom effort was used as a validating scale for growth mindset; and

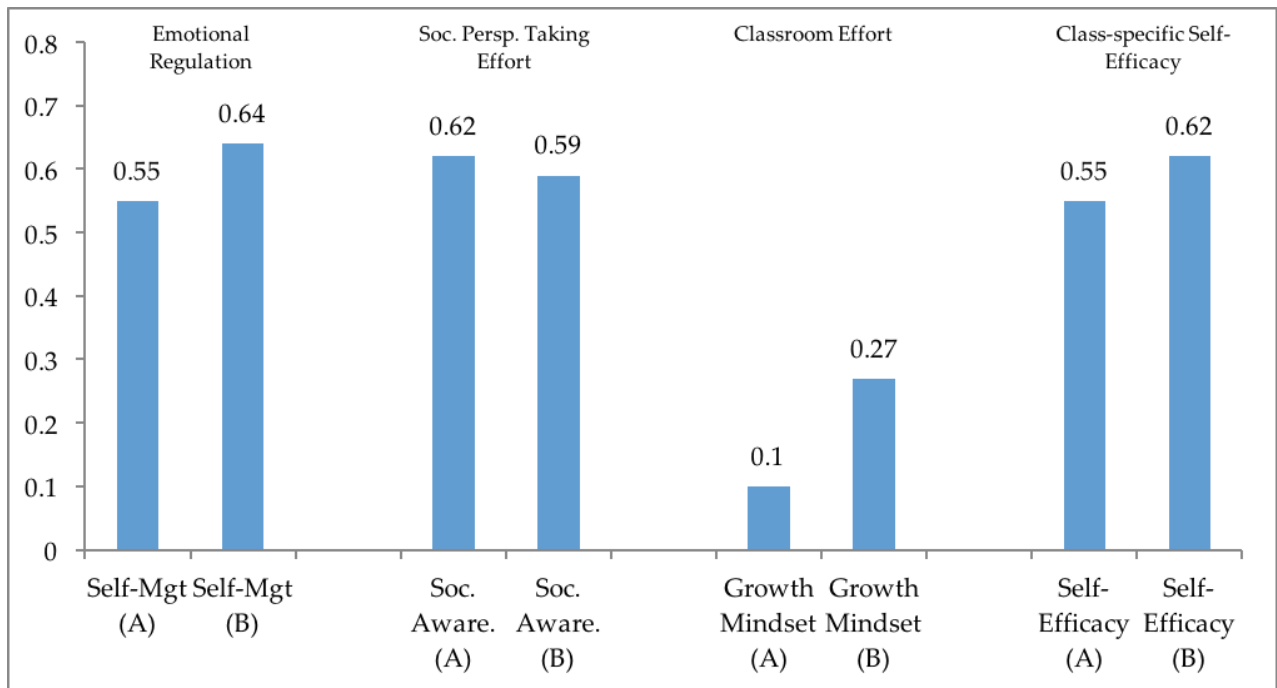
---

<sup>26</sup> Source: Analyses conducted by CEPR for CORE Districts and Transforming Education. Note: All correlations shown here are significant at p<.05.

- A survey of class specific self-efficacy was used as a validating scale for global self-efficacy.

The validating measures for self-management, social awareness and self-efficacy provided strong convergent validity evidence for each form of the primary MESH measures administered during the pilot (Exhibit 2). However, there was a weak relationship between classroom effort and growth mindset, indicating that the two surveys, growth mindset and classroom effort, may be measuring different underlying constructs. Since the growth mindset scale had previously been validated by Camille Farrington and her colleagues at CCSR, we concluded that “classroom effort” scale, which asked students to self-report the amount of effort they put into their math class may not have been an appropriate validating scale.

**Exhibit 2. The correlation between student ratings on the CORE measures and ratings on complementary validating measures of related skills<sup>27</sup>**



### TEST-CRITERION VALIDITY EVIDENCE

An important question when validating assessments is whether the ratings are predictive of outcomes hypothesized to be related to the underlying construct. As discussed in the Survey Research and Development section, research has shown that students with stronger MESH competencies are likely to have higher GPAs and test scores, as well as lower rates of absenteeism and suspension. Therefore, we would expect student and teacher ratings from the

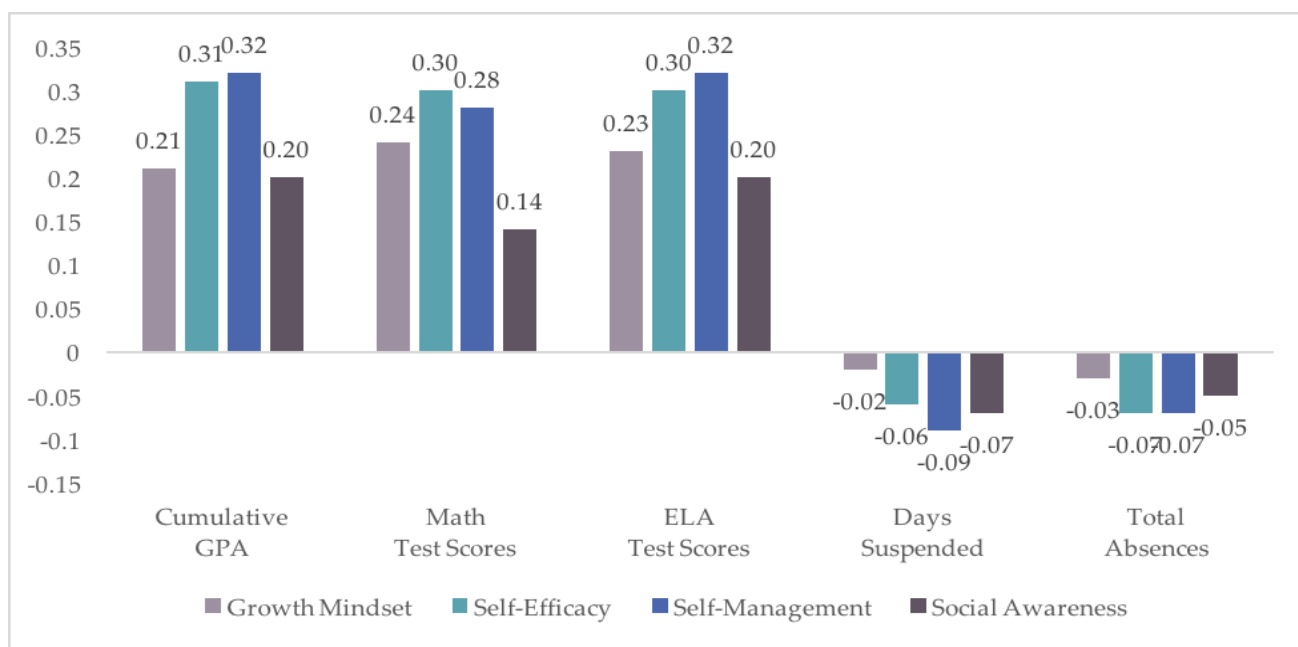
<sup>27</sup> Source: Analyses conducted by the Center for Education Policy Research (CEPR) at Harvard University for CORE Districts and Transforming Education.

MESH measures used in CORE districts to be positively related to student GPA and standardized test scores, and negatively related to absenteeism and suspension rates.

### CRITERION-RELATED VALIDITY EVIDENCE OF STUDENT SURVEYS

As Exhibit 3 shows, student ratings on each of the measures correlate in the expected direction with other academic and behavioral outcomes. For example, students’ self-efficacy ratings correlate positively with GPA and standardized math and ELA scores (with a correlation of approximately 0.3). They correlate negatively with the number of days a student was suspended as well as the total number of days s/he was absent (with a correlation of approximately -0.06). These correlations are all statistically significant.<sup>28</sup>

**Exhibit 3: Student-level correlation between self-reported MESH skills and other student outcomes<sup>29</sup>**



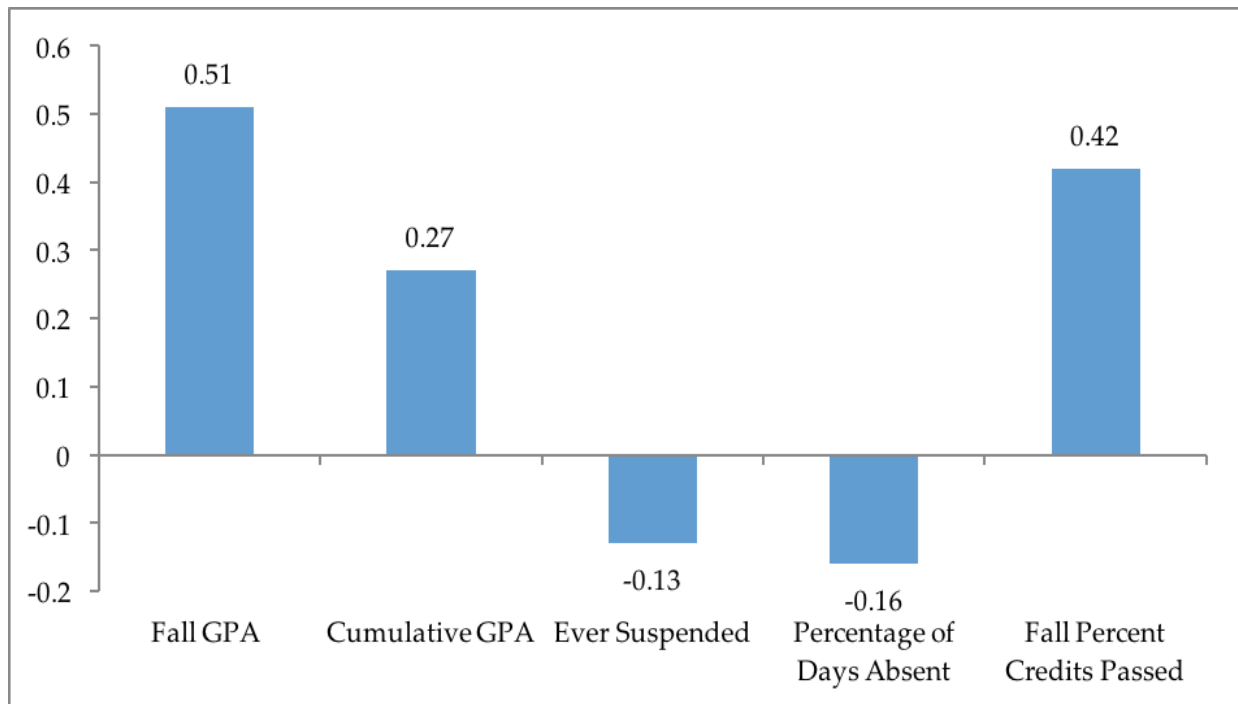
### CRITERION-RELATED VALIDITY EVIDENCE OF TEACHER SURVEYS

Teacher survey ratings of student skills should also be predictive of student outcomes. As Exhibit 4 shows, teacher ratings of students’ MESH competencies correlate in the expected direction with student outcomes of interest. All of the correlations are statistically significant. These findings suggest that because students and teachers have overlapping but distinct perspectives of students’ skills, it may be useful to include teacher reports in order to help triangulate students’ skills with greater accuracy.

<sup>28</sup> All of the correlations were statistically significant at the .001 level, with the exception of the correlation between growth mindset and suspensions, which was significant at the .01 level.

<sup>29</sup> Source: Analyses conducted by CEPR for CORE Districts and Transforming Education

**Exhibit 4: Student-level correlation between teacher-reported MESH skills and other student outcomes<sup>30</sup>**



## COMPARABILITY ACROSS SCHOOLS

As discussed in the Survey Research and Development section, a primary concern with regard to self-report surveys is that of reference bias, in which students' rate themselves relative to peers in their school or group rather than against an absolute criterion.<sup>31</sup> The concern is that students whose peer group has stronger MESH skills would rate themselves lower than students whose peer group has weaker MESH skills. In such a scenario, the comparability of score interpretations across schools would be compromised.

Our investigation of the impact of adjusting students ratings based on their responses to anchoring vignettes included in the 2014 pilot (see p. 8) did not improve the performance of the measures with respect to criterion validity. This provides an initial indication that reference bias may not be an important phenomenon within the CORE districts. We also examined the extent to which evidence of reference bias was observed in the field test results. To do so, we identified the relationship between students' MESH self-ratings and other student outcomes (e.g., GPA and standardized test rating) *across all schools* and compared these results to those based on the relationship between self-ratings and other outcomes among students attending the *same* school. Finding that the relationship is consistently stronger within a school would suggest the existence of reference bias. On the other hand, finding that the overall relationship between MESH skills

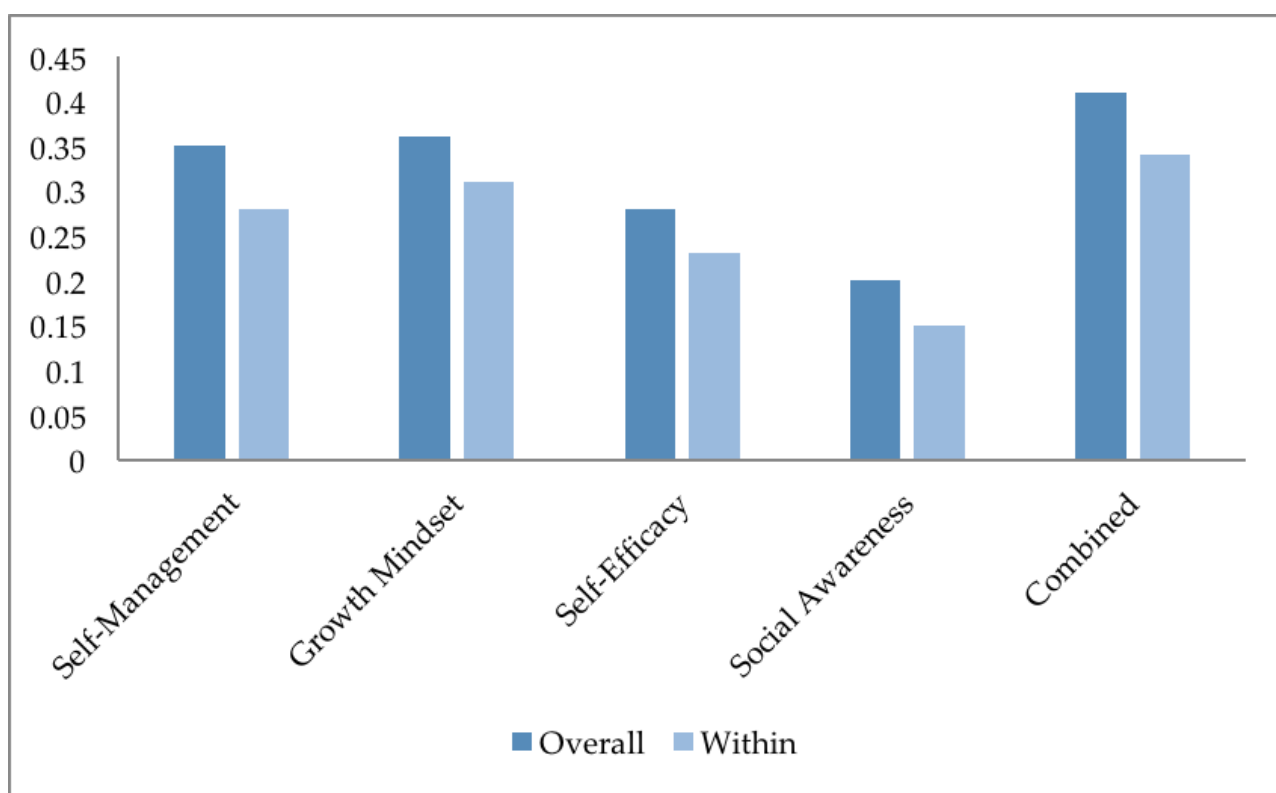
<sup>30</sup> Source: Analyses conducted by CEPR for CORE Districts and Transforming Education.

<sup>31</sup> See West, M.R., et al., *Promise and Paradox: Measuring Student's Non-Cognitive Skills and the Impact of Schooling*, Educational Research and Policy Analysis. August 2015; and 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

and other student outcomes is stronger than the within-school relationship provides evidence that reference bias may not be a significant concern in this data set.

Exhibit 8 shows the student-level correlations between the MESH competencies and English language arts (ELA) standardized test scores based on the overall analysis and the within-school analysis for middle schools. It is evident that the relationship between MESH self-ratings and ELA test scores across the four competencies is consistently stronger for the overall-school analysis compared to the within-school analysis. These correlations are all statistically significant, and similar patterns are found at the elementary school level and high school level, as well as with GPA and mathematics test scores.

**Exhibit 8. Student-level correlations between MESH skills and English language arts (ELA) test scores in CORE District middle schools, overall and within schools<sup>32</sup>**



These results provide suggestive evidence that reference bias due to differences in school environments may not be an important phenomenon in the CORE Districts. This issue is discussed in greater detail in a [Brookings Institution Brief by Martin West](#) and in the FAQs section of this document on p. 19. Of course, these findings do not rule out the possibility that other forms of bias may still be at play. For example, in future analyses, we intend to examine whether the individual items or scales may be interpreted differently by students in different subgroups.

<sup>32</sup> Source: Analyses conducted by Harvard CEPR for CORE Districts and Transforming Education.



## CONCLUDING REMARKS

Overall, the evidence presented above provides promising support for the validity, reliability, and comparability of the MESH measures developed by researchers, curated by TransformEd, and administered by the CORE Districts. This evidence is based on ratings that were collected without any stakes attached; therefore, the findings could change if stakes are attached to these measures.

Based on CORE's thoughtful, phased approach to this work, we will have some initial empirical evidence as early as summer 2016 about what happens when low stakes are attached to the measures. Under the auspices of their NCLB waiver, CORE Districts were slated to include student self-report surveys on MESH skills in their Fall 2016 School Quality Improvement Index (using data collected in Spring 2016). With the passage of ESSA and the ending of waivers in August 2016, the role of these measures in formal school accountability is currently uncertain for the CORE Districts. That said, if we see a significant increase in students' self-reported MESH skills between school year 2015 (when no stakes were attached) and school year 2016 (when the intention was to attach low stakes at the school level), then we will know that further exploration is needed to ensure that these measures are providing accurate information about students' competencies.

In 2016, TransformEd will also be releasing a case study on CORE's work and a policy brief that will contain additional evidence on the validity and reliability of these measures as well as examples of how they can be used in practice.

## FAQS

### **Why assess students' MESH competencies?**

We believe that preparing students for college, career, and life goes beyond teaching core academic skills. MESH competencies have been shown to predict students' academic achievement as well as their success in the workplace and their well-being throughout life. Assessing these competencies systematically can help build awareness about the importance of MESH skills and provide students, educators, and parents with baseline information that can help students improve their MESH skills over time.

### **Can my school focus on a subset of the four competencies covered by these measures?**

Yes. The four competencies prioritized by the CORE Districts do not constitute a perfect or exhaustive set of MESH competencies. Rather, CORE and its advisors selected these competencies as a starting point for their work together, acknowledging that the list of competencies and the measures might evolve over time. The individual scales for growth mindset, self-efficacy, self-management, and social awareness could be used separately and/or combined with additional scales to assess other competencies.

### **Can I use a subset of the items listed under each competency?**

Each of the measures included here was tested and validated as a survey scale, or a series of items administered together to assess a particular competency. As such, we recommend using each scale in its entirety rather than selecting a subset of the items that make up that scale. Any changes to the scales could affect the validity and reliability of the survey.

**In what grades can these measures be used?**

The student self-report survey is intended for students in grades 5-12. The teacher survey is intended for teachers of students in grades K-12.

**Should we administer both student and teacher surveys?**

Analyses from the CORE field test suggest that students and teachers have overlapping but distinct perspectives of students' skill levels. No measure is perfect, so educators and researchers often find that using multiple measures can help triangulate students' skills with greater accuracy. You might also consider cross-referencing the data from these surveys with other information you already collect (e.g. homework completion, attendance, data from a PBIS system, etc.).

**How much time will this survey take to complete?**

The student survey takes, on average, 15-20 minutes for students to complete. The teacher survey takes approximately one minute per student on which teachers report. At the secondary level, where educators may teach 150+ students over the course of a day, we recommend capping the number of students on which a teacher reports. For example, some of our district partners have used a 60-student cap in order to limit the total time teachers spend reporting on students' SE skills to approximately one hour.

**When should these measures be administered?**

Student self-reports can be administered at any time, but teacher reports generally require that the teacher have had time to get to know the student. For this reason, spring may be more appropriate time than fall to administer surveys. Administering the surveys at approximately the same time each year helps ensure that the year-over-year data is as comparable as possible. If you plan to administer the measures twice per year (e.g. as a pre- and post-intervention assessment), then it is best to wait until approximately one month into the school year (e.g. because some of the scales reference the "past 30 days" specifically).

**What are the costs associated with administering these measures?**

The surveys themselves are free: TransformEd curated these measures for use in the CORE Districts, and both organizations are committed to sharing our work with a broad audience in order to improve MESH outcomes for students nationwide. These surveys can be administered on paper or online. Some of our partner districts have chosen to administer the surveys themselves (which can be done at low or no cost if the district already has experience administering surveys and analyzing data). Other partner districts have chosen to pay an external survey administration and reporting partner to streamline the administration process. For example, several of the CORE Districts worked with Panorama Education to administer their surveys.

**How should these measures be scored?**

There are multiple ways to summarize the data, but the CORE Districts have chosen to score each item on a scale from 1 to 5, with 5 indicating the "best" response (e.g. "almost all the time," "completely true," etc.) Please note that the four Growth Mindset items are reverse scored, meaning that the "best" response is "not at all true," so this response would be scored as a 5.) The CORE Districts considered 4's and 5's as "positive" and reported aggregated data in terms of the percentage of all responses that were "positive" for each individual competency.

**Will benchmark data be available?**

If you're interested in learning more about benchmark data, please contact us by phone (617-378-3939) or email: [MeasuringMESH@transformingeducation.org](mailto:MeasuringMESH@transformingeducation.org).

### **Are you worried that these measures are “gameable”?**

As with any survey-based measure, it is possible for respondents to provide ratings that are not accurate or truthful (e.g. for students or teachers to inflate their ratings of students’ MESH competencies). The risk of this increases when high stakes are attached to the measures. As such, we do not recommend using these measures in a high-stakes setting where there are punitive consequences for schools. Moreover, these surveys should not be used for high-stakes decisions about individual students or educators.

### **What is reference bias?**

Reference bias refers to the tendency for survey responses to be influenced by the context in which the survey is administered. Experts in reference bias point out that the culture of a school might influence a students’ frame of reference, which may in turn cause the student to interpret a survey scale differently than students attending other schools.<sup>33</sup> For example, students attending a school that heavily emphasizes self-management might develop higher internal standards for self-management. Such students might then rate themselves lower on a self-reported measure of self-management than students in a school that does not emphasize this competency. This phenomenon is often marked by a significant difference in the correlation between self-reported skills and related academic and behavioral indicators within schools versus between schools.

To mitigate concerns about reference bias, the CORE Districts partnered with ETS to pilot anchoring vignettes, a technique that uses brief descriptions of sample students who exhibit varying levels of the target competency as a tool for norming students’ responses to ensure comparability. Ultimately, the analyses conducted by Harvard Center for Education Policy Research (CEPR) showed that using the anchoring vignettes did not improve the quality of the survey data. One possible interpretation of this finding is that the school cultures vis-à-vis the relevant MESH competencies were not dramatically different across schools, so students taking the surveys at different schools had relatively similar frames of reference. Ultimately, CORE decided not to use anchoring vignettes during the 2014-15 field test because the potential research benefit to including them was outweighed by educators’ desire to minimize testing time. However, we continued to use the field test data to examine the possibility that reference bias could undermine cross-school comparisons of student skills by comparing overall and within-school correlations of student self-reported skills and academic performance as described on p.15.

### **How are states, districts, and schools using the data from these measures?**

We will soon be releasing a policy brief and case study outlining how these data are currently being used or could be used by schools, districts, and states. We also want to spotlight unique approaches developed by schools, districts, and states across the country. If you would like to share your approach to using MESH data to improve student and school outcomes, please contact us at [MeasuringMESH@transformingeducation.org](mailto:MeasuringMESH@transformingeducation.org).

### **Should these measures be used in an accountability system?**

It is clear that the word “accountability” means different things to different people. When it comes to accountability, one of the most relevant questions is what the consequences are for schools identified as lower-performing. In the case of CORE, the consequences are quite intentionally supportive rather than punitive. Specifically, low-performing schools are paired with higher-performing schools, which provide support and mentorship to help build educators’ capacity and

---

<sup>33</sup> See e.g., West, M.R., et al., *Promise and Paradox: Measuring Student’s Non-Cognitive Skills and the Impact of Schooling*, Educational Research and Policy Analysis. August 2015. (The paper explores this paradoxical effect based on research that Transforming Education and its partners conducted in Boston schools).

improve student outcomes over time. Perhaps most importantly, there are no consequences attached to the survey-based MESH measures for individual teachers or students.

With the passage of the Every Student Succeeds Act (ESSA), the federal government has empowered states to develop their own accountability systems. ESSA also acknowledges that schools and students are more than just test scores by requiring states to choose at least one other measure of school quality or student success as part of their state-determined systems. Furthermore, ESSA gives states the most latitude they've had in recent years to determine what the consequences of their accountability systems will be. We hope that states will use these new flexibilities to articulate a broader, more holistic definition of student success and school performance. We also hope that states will find ways to prioritize MESH skills and assess these skills systematically so that educators have access to the data they need to make informed decisions about how to support students most effectively. We think there are multiple ways to accomplish this goal, including incorporating these measures into a state-determined accountability system or collecting the data systematically for formative purposes only.

We understand that people have questions and concerns about how attaching even low stakes to survey-based MESH measures will play out in practice. Based on CORE's thoughtful, phased approach to this work, we will have some initial empirical evidence on this topic as early as summer 2016. Under the auspices of their NCLB waiver, CORE Districts were slated to include student self-report surveys on MESH skills in their Fall 2016 School Quality Improvement Index (using data collected during Spring 2016). With the passage of ESSA and the ending of waivers in August 2016, the role of these measures in formal school accountability is currently uncertain for the CORE Districts. That said, if we see a significant increase in students' self-reported MESH skills between school year 2015 (when no stakes were attached) and school year 2016 (when the intention was to attach low stakes at the school level), then we'll know that further exploration is needed to ensure that these measures are providing accurate information about students' competencies.

## SURVEYS

Below are the teacher and student surveys on social emotional competencies. The student survey contains four scales: self-management, self-efficacy, growth mindset, and social-awareness. Each scale consists of 5-9 items with a 5-point answer scale. The entire survey (with all four measures) typically takes students 15-20 minutes to complete. The scales are intended to be administered to students in grades 5-12.

Following the four student scales, we provide questions on student demographics that may be helpful to educators as they consider using the survey data for school improvement purposes. Demographic questions, if included, should always be administered at the *end* of the survey in order to mitigate the risk of stereotype threat.

The teacher survey contains three scales: self-management (school work), self-management (interpersonal), and social awareness. These are based on the two competencies that are believed to be externally observable. The surveys are designed to be administered to teachers of students in grades K-12. The entire survey usually takes approximately one minute to complete per student. Each scale contains 4-5 descriptors of the relevant competency and a single 5-point answer scale. Please note that teachers are asked to consider the full set of descriptors and provide a single overall rating for each competency. In other words, teachers should provide three ratings per student: one for self-management (school work), one for self-management (interpersonal), and one for social awareness. Teachers do NOT need to rate students on the individual descriptors listed under each competency.

To inquire about how TransformEd can assist your school, district, or state in identifying, field testing, and employing measures of students' MESH competencies, please reach out to us at **617-378-3939** or [MeasuringMESH@transformingeducation.org](mailto:MeasuringMESH@transformingeducation.org).

## SAMPLE ADMINISTRATION PROTOCOLS

When administering the student and teacher surveys, we recommend adapting the sample protocols shown here to your particular needs. Administering the measures at approximately the same time each year helps ensure that the year-over-year data is as comparable as possible. If you plan to administer the measures twice per year (e.g. as a pre- and post-intervention assessment), then it is best to wait until approximately a month into the school year (e.g. because some of the scales reference the "past 30 days" specifically).

Each of the measures included here was tested and validated as a survey scale, or a series of items administered together to assess a particular competency. As such, we recommend using each scale in its entirety rather than selecting a subset of the items that make up each scale. Any changes to the scales could affect the validity and reliability of the survey. Additionally, if you choose to use the standard demographic questions included with the surveys, these should always be administered at the end of the survey in order to decrease the possibility of stereotype threat.

If you would like to discuss our recommendations on how to use these measures or inquire about how TransformEd can assist your school, district, or state in identifying, field testing, and employing measures of students' MESH competencies, please reach out to us at **617-378-3939** or [MeasuringMESH@transformingeducation.org](mailto:MeasuringMESH@transformingeducation.org).

## FOR STUDENT SURVEYS:

**Sample explanation to students on the purpose of the survey:** This introduction to students is an example; it should be customized based on how each district or school decides to administer the surveys.

*“Today you will be taking a survey about how you approach your classes and how you interact with others. For all questions on today’s survey, please think back to the past 30 days of school. The results of this survey will give us feedback that helps us improve our school, so please respond honestly. There are no wrong answers, and once you have discarded the cover sheet with your name on it, I will not be able to identify your responses. Your responses will not impact your grades in any way.”*

**Oversee students completing the survey:** We recommend standing in one place in the classroom where you can see all students (to ensure that students are on task) but where you cannot read students’ responses. Having a teacher circulate around the room or stand close by may make students concerned about the confidentiality of their responses and lead them to respond differently.

Each student should read and respond to the survey independently. \*If a student has a question, feel free to define a word that they do not understand, but please do not reinterpret an entire statement or respond to a question in a way that may influence their answer. If the student is still having trouble, simply ask them to answer as best they can or leave the question blank.

**\*Offer appropriate accommodations to students with special needs:** When taking the survey, students with special needs should be offered accommodations consistent with their IEP. Such accommodations may include but are not limited to the following:

- Student is given additional time to complete the survey
- Survey is read aloud by a staff member
- Responses are filled in by a staff member on behalf of the student

**When students finish the survey,** please instruct them to complete a quiet activity at their desks until all other students have finished.

## FOR TEACHER SURVEYS:

**Sample explanation to teachers on the purpose of the survey:** This introduction is just an example; it should be customized based on how each district or school decides to administer the surveys.

*“As part of this pilot, you will be asked to assess students’ competencies through a confidential, individualized survey. The results of this assessment will have no bearing on students’ grades or on teacher evaluations; rather, they will be used to provide more targeted resources to students and schools. Our goal is to ensure that these assessments are feasible, valid, reliable, and useful to educators before we move forward with them. As such, neither your principal nor anyone in your district will see how you respond. The results of the assessment will be reported back at the school level, with breakdowns by grade or subject given only when such breakdowns do not jeopardize the anonymity of respondents.”*

*The survey will take approximately 30 – 60 minutes to complete. Please note that the survey window will close on (date), and anyone who has not yet completed the survey will receive periodic reminders via email during the survey window.*

*Please contact (contact name and phone number) if you need technical support.”*

**Additional reminders:** In customizing the introductory text, we recommend that:

- Administrators be clear about the purpose of the survey and how the resulting data will be used.
- Administrators inform teachers of the expected length of time the survey will take and the length of time they have to complete the survey. The teacher survey takes approximately one minute per student on which teachers report. At the secondary level, where educators may teach 150+ students over the course of a day, we recommend capping the number of students on which a teacher reports. For example, some of our district partners have used a 60-student cap in order to limit the total time teachers spend reporting on students’ MESH skills to approximately one hour.
- If the teacher survey is administered online, administrators should also provide clear instructions for technical support, whether that support is to be offered by local school / district / state staff or provided by an external survey administration partner. If the survey is administered on paper, administrators should provide clear instruction on where to submit completed surveys.

## **STUDENT SELF-REPORT SURVEY ON MESH COMPETENCIES**

***This survey asks about your behavior, experiences, and attitudes related to school. We look forward to using your feedback to try to make schools better.***

***Some of the survey questions will ask you about specific periods of time (such as the past 30 days). Please pay careful attention to these time periods and classes when you respond.***

***Thank you for taking this survey!***

### **Self-Management<sup>34</sup>**

***First, we’d like to learn more about your behavior, experiences, and attitudes related to school.***

***Please answer how often you did the following during the past 30 days. During the past 30 days...***

1. I came to class prepared.
2. I remembered and followed directions.
3. I got my work done right away instead of waiting until the last minute.
4. I paid attention, even when there were distractions.
5. I worked independently with focus.
6. I stayed calm even when others bothered or criticized me.

---

<sup>34</sup> Adapted from Patrick & Duckworth (2013, May) *Empirical support for a tripartite taxonomy of character in adolescents*. Poster presented at the 25th annual convention of the Association for Psychological Science.



7. I allowed others to speak without interruption.
8. I was polite to adults and peers.
9. I kept my temper in check.

*(Almost Never, Once in a While, Sometimes, Often, Almost All the Time)*

### **Growth Mindset<sup>35</sup>**

***In this section, please think about your learning in general.***

***Please indicate how true each of the following statements is for you:***

10. My intelligence is something that I can't change very much.
11. Challenging myself won't make me any smarter.
12. There are some things I am not capable of learning.
13. If I am not naturally smart in a subject, I will never do well in it.

*(Not At All True, A Little True, Somewhat True, Mostly True, Completely True)*

### **Self-Efficacy<sup>36</sup>**

***How confident are you about the following at school?***

14. I can earn an A in my classes.
15. I can do well on all my tests, even when they're difficult.
16. I can master the hardest topics in my classes.
17. I can meet all the learning goals my teachers set.

*(Not At All Confident, A Little Confident, Somewhat Confident, Mostly Confident, Completely Confident)*

### **Social Awareness<sup>37</sup>**

***In this section, please help us better understand your thoughts and actions when you are with other people.***

***Please answer how often you did the following during the past 30 days. During the past 30 days...***

18. How carefully did you listen to other people's points of view?  
*(Not Carefully At All, Slightly Carefully, Somewhat Carefully, Quite Carefully, Extremely Carefully)*
19. How much did you care about other people's feelings?  
*(Did Not Care At All, Cared A Little Bit, Cared Somewhat, Cared Quite A Bit, Cared A Tremendous Amount)*
20. How often did you compliment others' accomplishments?  
*(Almost Never, Once in a while, Sometimes, Often, Almost all the time)*
21. How well did you get along with students who are different from you?

---

<sup>35</sup> Farrington et al. (2013) Becoming Effective Learners Survey Development Project, Chicago Consortium for School Research.

<sup>36</sup> Adapted from Farrington et al. (2014) *Becoming Effective Learners Survey Development Project, Chicago Consortium for School Research.*

<sup>37</sup> Adapted from AIR and CASEL (2013) Student self-report of social and emotional competencies.



*(Did Not Get Along At All, Got Along A Little Bit, Got Along Somewhat, Got Along Pretty Well, Got Along Extremely Well)*

22. How clearly were you able to describe your feelings?  
*(Not At All Clearly, Slightly Clearly, Somewhat Clearly, Quite Clearly, Extremely Clearly)*
23. When others disagreed with you, how respectful were you of their views?  
*(Not At All Respectful, Slightly Respectful, Somewhat Respectful, Quite Respectful, Extremely Respectful)*
24. To what extent were you able to stand up for yourself without putting others down?  
*(Not At All, A Little Bit, Somewhat, Quite A Bit, A Tremendous Amount)*
25. To what extent were you able to disagree with others without starting an argument?  
*(Not At All, A Little Bit, Somewhat, Quite A Bit, A Tremendous Amount)*

### **Demographic Questions**

***Finally, we'd like some background information about you.***

#### **Are you male or female?**

- Male
- Female
- I'd rather not say

#### **What grade are you in?**

- 3<sup>rd</sup> grade
- 4<sup>th</sup> grade
- 5<sup>th</sup> grade
- 6<sup>th</sup> grade
- 7<sup>th</sup> grade
- 8<sup>th</sup> grade
- 9<sup>th</sup> grade
- 10<sup>th</sup> grade
- 11<sup>th</sup> grade
- 12<sup>th</sup> grade
- Other grade
- Ungraded

#### **What is your race or ethnicity? (Please pick only one answer.)**

- American Indian or Alaska Native
- Asian
- Black or African American
- Hispanic or Latino
- Filipino
- Native Hawaiian or Other Pacific Islander
- White
- Two or More Races/Ethnicities

## TEACHER REPORT ON STUDENTS' MESH COMPETENCIES

***We'd like you to think about how well your students were able to manage themselves with regard to their school work. How frequently did each student display the following set of behaviors during the past 30 days?***

***Please consider the full set of behaviors and provide a single overall rating for each student. In other words, you will provide three ratings per student: one for self-management (school work), one for self-management (interpersonal), and one for social awareness. You do NOT need to rate students on each individual behavior listed under self-management (school work), e.g. "came to class ready to learn."***

### **Self-Management (School Work)<sup>38</sup>**

- Came to class ready to learn
- Remembered and followed directions
- Persisted when tasks became challenging
- Paid attention and maintained focus
- Resisted distractions

*(Almost Never, Once in a While, Sometimes, Often, Almost All the Time)*

***Next, please think about how well your students were able to manage themselves with regard to classroom behavior. How frequently did each student display the following set of behaviors during the past 30 days? (Please consider the full set of behaviors and provide a single overall rating for each student.)***

### **Self-Management (Interpersonal)<sup>39</sup>**

- Remained calm even when under stress
- Allowed others to speak without interruption
- Got along well with others
- Kept his/her temper in check

*(Almost Never, Once in a While, Sometimes, Often, Almost All the Time)*

***Finally, please think about how well your students were able to manage themselves with regard to their social interactions in class. How frequently did each student display the following set of behaviors during the past 30 days? (Please consider the full set of behaviors and provide a single overall rating for each student.)***

### **Social Awareness<sup>40</sup>**

- Listened carefully to other people's points of view.
- Got along with students who were different from him/her.
- Disagreed with someone without starting an argument.
- Stood up for him/herself without putting others down.
- Noticed and complimented others' accomplishments.

*(Almost Never, Once in a While, Sometimes, Often, Almost All the Time)*

---

<sup>38</sup> Adapted from Patrick, S. D., & Duckworth, A. L. (2013, May). *Empirical support for a tripartite taxonomy of character in adolescents*. Poster presented at the 25th annual convention of the Association for Psychological Science

<sup>39</sup> Adapted from Patrick, S. D., & Duckworth, A. L. (2013, May). *Empirical support for a tripartite taxonomy of character in adolescents*. Poster presented at the 25th annual convention of the Association for Psychological Science.

<sup>40</sup> Adapted from AIR and CASEL (2013). *Teacher rating of student social and emotional competencies*.