# *HOLDING SCHOOLS ACCOUNTABLE FOR EQUITY IN CHARACTER MEASUREMENT*
## *Bridging Theory and Practice*

**Madora Soutter**
*Villanova University*

**Shelby Clark**
*Harvard University*

**Alison Lee**
*EL Education*

When educators are committed to centering character education in their classrooms, how can they be certain that their initiatives are working? Even though a plethora of research points to the positive benefits of character and socioemotional learning[1] in schools (e.g., Berkowitz, 2011; Mahoney et al., 2018), how can individual teachers and school communities be assured that their own approach and implementation is having the intended impact? Just as importantly, how can they know if students of all identities and demographics (e.g., race/ethnicity, gender, age, socioeconomic status, and religion) are benefiting as intended in the same way?

Our aim here is to focus specifically on how educators can hold themselves accountable for the effectiveness of their character education equity goals. We begin by laying a foundation of the field of character measurement and highlight some of the associated critiques and concerns. We then hone in on two specific areas that have the potential to begin to address some of these concerns: measurement that focuses on contextual factors and measurement for continuous improvement. We then highlight one specific organization, EL Education, to show how their approach to character education and measurement exemplifies these two areas.

Our hope is that this article will be useful for educators who are working to center character in their schools and classrooms, but who also want to hold themselves accountable for doing so effectively, critically, and equitably.

## *EQUITY CONCERNS IN CHARACTER MEASUREMENT*

Although character education has been a foundational part of American schooling since its origins (Campbell, 2012; Dewey, 1916/1966, 1938; Johanek, 2012; McClellan, 1999; Seider, 2012), recent years have seen an explosion of a new focus on character and social

• **Correspondence concerning this article should be addressed to:** Madora Soutter, madora.soutter@villanova.edu

Journal of Character Education, Volume 18(1), 2022, pp. 113–134 ISSN 1543-1223

emotional learning frameworks, education, and measurement (e.g., CASEL, Character Lab, 21st Century Skills, etc.; Assessment Work Group, 2019; EASEL Lab, 2019). In fact, one study found that over 71% of American principals are currently implementing, or plan to implement, socioemotional learning (SEL) programs in their schools (Atwell & Bridgeland, 2019); moreover, schools and school districts are estimated to spend approximately $640 million on such education annually (Assessment Work Group, 2019; Krachman & LaRocca, 2017). This emphasis on character education and SEL in schools is unsurprising, given that studies continually demonstrate associations between character or SEL and positive life outcomes; for example, a 2011 meta-analysis of 213 studies and more than 270,000 students indicated that social and emotional programming yields significant positive effects on prosocial behaviors, self-esteem, and academic performance, as well as reduced behavioral incidents and reduced emotional distress such as depression and anxiety (Durlak et al., 2011).

Despite this growing emphasis on character and social emotional learning throughout educational settings, how to best measure character in schools and evaluate the effectiveness of character interventions has remained a complex question. Concerns around reliability (e.g., will results differ depending on a student's particular mood when completing a survey? Might students and teachers misinterpret a survey item or inaccurately report on internal emotions?), face validity (e.g., does a particular survey on student empathy actually measure how a school is defining empathy? How can we account for students who will rate themselves higher to appear more desirable?) and reference bias (e.g., does a student's frame of reference impact how they interpret elements of character) abound. Questions around what constitutes possessing a certain character strength remain as well: for example, should we trust student self-reports on their own beliefs and dispositions, or must we rely on measuring only external manifestations such as tangible behaviors that can be seen and documented?

Perhaps most importantly, there is also a growing call for a need to be attuned to issues of character assessment and equity, with equity referring to the concept that "each child receives what he or she needs to develop to his or her full academic and social potential" (nationalequityproject.org). For example, to date, research has found that students of color score lower on traditional summative surveys of character, signaling a need to shift the way that schools conceptualize and evaluate these constructs (e.g., Fiarman et al., 2021). Similarly, character and SEL education has not been found to benefit all students in a fair and equitable manner. For example, in a recent review of 136 SEL competency frameworks, fewer than 20% took culturally or linguistically diverse samples into consideration, fewer than 20% considered youth with disabilities' experiences, and fewer than 6% considered students' potential experience with trauma (Berg et al., 2017). A study by RAND has likewise found that schools that implement SEL use such frameworks to control and manage students' individual behaviors rather than to foster student belonging or to explore contextual aspects of students' lives (The Education Trust, 2020; Hamilton et al., 2019). In fact, in focus groups with students of color, Black families, and Latinx families, The Education Trust (2020) found that these participants did not trust their educational institutions to teach the SEL skills and strengths that they value, such as a knowledge of one's self and one's cultural identity, and did not believe that these institutions were set up to create safe, unbiased spaces where students of color could explore their authentic selves.

One specific equity concern relates to the continued focus on the need for individual students to "fix" themselves and their own character rather than for schools to provide adequate environments and relationships to help nurture student character (e.g., Fisher, 2020; Kaler-Jones, 2020; Kirshner, 2015; Love, 2019; Simmons, 2019; Soutter, 2020;

Zaino, 2016). For example, Fisher (2020) noted that in one recent prominent character report the word "skills" appeared 140 times whereas the words "relationship" and "network" appeared only 23 and 14 times, respectively. As Fisher stated,

> Students certainly need to develop a host of critical *skills* to manage their emotions and interact productively with others…. But students also need to be exposed to relationship-rich environments and offered *access* to a range of supportive connections who can help them get by and get ahead. (emphasis in original, para. 4)

In a similar critique of the preeminent SEL focus on teaching students of all backgrounds to persist or "have grit" despite whatever systemic challenges they might encounter, one scholar noted, "can you imagine the outcry if, let's say, an old toxic dump was discovered near Scarsdale or Beverly Hills and the National Institutes of Health undertook a program to teach kids strategies to lessen the impact of toxins but didn't do anything to address the toxic dump itself?" (Rose, 2013). Clearly, a "one size fits all" approach to character and SEL measurement will not adequately capture the experiences and development of all students; measuring character and SEL equitably means looking outwards towards our environments and contexts and understanding how they, too, can impact student development.

A second equity concern relates to how character measurements are *used* by teachers and institutions. Character education measurement traditionally grew out of trait-based personality measures (e.g., the Big 5 trait taxonomy; Costa & McCrae, 1992), that originally aimed to assess how stable an individual's personality remained over time (Lerner et al., 2015). As such, these tools, although often used to measure student character (Lerner et al., 2015), were not optimally designed to be responsive to teachers' need for developmentally appropriate measures that could both assess students' character change and growth and subsequently inform teacher practice (Assessment Work Group, 2019). Based on this history, current character summative assessments—those that attempt to measure student or teacher results or progress after a period of time—tend to be higher stake (e.g., student or teacher outcomes, such as performance ratings, might rely on the results) and/or they do not provide enough information from which to make actionable, real-time adjustments in the classroom (Assessment Work Group, 2019; Marsh et al., 2018; McKown, 2019). Given the rise in character education in schools across the United States and elsewhere, a new focus on *low-stakes*, formative character measurement in schools is now becoming increasingly widespread (Bryk et al., 2015; McKown, 2019). That is, more and more educators and administrators want character and social emotional learning measures that can be used for "continuous improvement" efforts in classrooms in order to help better guide instruction and respond to students' character and social emotional learning needs (Assessment Work Group, 2019; Fiarman et al., 2021; Lee & Riordan, 2018; Schlund et al., 2020). Still, the question of how character measurements are interpreted and used remains an important concern. If a student receives a low mark on an end-of-year character report card, for example, that score might become internalized (by themselves and others) as a final judgment of that student's character. If, on the other hand, measurements are explicitly used to gather information so as to make instructional and institutional improvements, these measures have the potential to be interpreted with a more strengths-based lens (e.g., what can I as the teacher do to support students in these ways) and to serve a very different purpose.

Thus, if we care about finding ways to measure character equitably in schools it behooves us to find ways to *holistically* and *formatively* assess the character and context of all students. Below, we delve further into these two areas related to character measurement and equity, followed by a description of the ways in which one organization—EL Education—has

approached character measurement around these two areas of focus through their co-construction and initial implementation of their walkthrough observation tool. This approach looks beyond the use of survey tools alone and works to create an equitable school environment through assessing students' learning contexts and using these data for the purpose of continuous improvement. We are hopeful that such a case study might prove useful to others interested in creating formative measures for assessing student character development, but, perhaps more importantly, for those working with or in schools attempting to understand new ways to think about equitable character assessment and implementation.

## *ADDRESSING EQUITY CONCERNS IN CHARACTER MEASUREMENT*

Given these circumstances and the concerns outlined above, scholars have outlined several ways that the field can address these two overarching equity concerns regarding contextual considerations and summative vs. formative measurement. We review the various ways these concerns have been addressed in the sections below.

### *Equity in Measurement: Contextual Considerations*

**Individual Versus Contextual Conceptualizations.** Many scholars have advocated for a reconceptualization of character and SEL. For example, Jagers and colleagues (2018) have argued that there is a need for "transformative SEL," a form of SEL that acknowledges that students can have vastly different experiences in the classroom due to racial/ethnic or gender discrimination, economic status, or past traumatic events. As such, they noted that transformative SEL "build[s] strong, respectful relationships founded on an appreciation of similarities and differences; learn[s] to critically examine root causes of inequity; and develop[s] collaborative solutions to community and social problems" (p. 2). Similarly, The Education Trust (2020) champions a vision of SEL education that similarly focuses on recognizing contextual and cultural differences across individuals, challenges students to reach their full potential, aims to reduce bias, works to foster student belonging, and focuses on the whole student.

Within the character education field scholars have similarly criticized the primary focus on developing students' individual character strengths to the detriment of acknowledging other contextual and systemic factors that hinder or promote character growth. For example, Seider and Graves (2020) argue that focusing on cultivating character strengths like grit or growth mindset in students "can have powerful and positive effects, but doing so can also unintentionally convey that students' ability to thrive is entirely dependent upon their own internal qualities" (para. 2). They argue, instead, for a critically conscious framework for character education—one that includes teaching students about structural inequities and how to challenge such inequities while drawing on moral, intellectual, and civic character strengths (Seider et al., 2017). Kirshner (2015) has similarly argued that youth development, and youth character development, needs to be understood from a sociocultural perspective wherein youth development is understood holistically within its ecologies and contexts. As others have noted, equitable social and emotional learning needs to move away from focusing on how to "fix kids" and more towards focusing on the contexts within which children live and develop, including their cultural, familial, socioeconomic, educational, individual, and societal contexts (The Education Trust, 2020). This shift is complex, of course, as character development is most commonly understood as resting neither solely with an individual *nor* solely within the influence of one's contextual surroundings (e.g., Lerner & Callina, 2014; Search Institute, 1997). Indeed, consistent with ecological systems theory (Bronfenbrenner, 1979, 1993) and relational developmental systems metatheory (Lerner & Schmid Callina, 2014), character development is most recently understood as

resulting from reciprocal interactions between the individual and their culture(s), peers, schools, families, sociopolitical climate, and other and contexts (Osher et al., 2020). So while we cannot attribute socioemotional development to context alone, it is nonetheless critically important to recognize these factors to avoid an overly individualistic vision of character that ignores external and systemic influences.

What does transformative SEL (Jagers et al., 2018) or a critically conscious approach to character (Seider et al., 2017) mean for character measurement? How can we *measure* character equitably? As Carter and colleagues professed (2017): "You can't fix what you don't look at" (p. 207); if part of equitable character education is being attuned to students' contexts and how they impact student development (e.g., The Education Trust, 2020), then, in order to build more equitable learning environments we need to *assess and analyze* these learning contexts and environments rather than only students' thoughts, attitudes, and behaviors themselves (Schlund et al., 2020).

**Measuring Contextual Elements.** One example of this focus on contextual elements comes from Paunesku and Farrington's (2020) work with practitioners to develop an educational and developmental framework focused on students' socioemotional and academic development (SEAD). In this project, Paunesku and Farrington (2020) examined how contextual, "upstream," elements, such as feelings of classroom belonging, could influence "downstream" elements such as students' social emotional outcomes and levels of academic attainment. Accordingly, these scholars argued that it is critical that scholars and practitioners move away from the more typical focus of accountability measures (such as test scores) and focus more on measuring the ways that students feel about the "upstream" elements that are related to the development of their social emotional and academic strengths (Paunesku & Farrington, 2020). Importantly, Paunesku and Farrington (2020) pointed out that such strengths tend to not be something that an individual "has" or does not "have," but rather are something that one displays depending upon the surrounding context, which, they argued, further indicates the importance of measuring educational contexts and how they are affecting students.

Notice that when speaking about character and SEL measurements in schools we often refer almost exclusively to *surveys*. In fact, in a recent report on the state of K–12 SEL education, Tyton Partners (2020) reported that 52% of school districts have engaged in some form of SEL measurement, with the Panorama Student Survey, Panorama Teacher and Staff Survey, Panorama Family-School Relationships Survey, and Panorama SEL survey being the most popular forms of measurement in schools. Other core SEL measurement providers and assessments identified in the report include the CORE district surveys, YouthTruth, Perts, Emote, and the Pearson Social Skills Improvement System Rating Scales (SSIS), and xSEL Labs. Of these, xSEL Labs and the Pearson SSIS are the only non-survey options; xSEL labs offers behavioral measures (those that directly measure observable skills) in addition to teacher and peer rating forms, while the Pearson SSIS offers a way for teachers to directly rate their students.

Indeed, many questions exist about whether surveys are the most appropriate way to measure character, particularly from an equity perspective. Many, for example, worry that character surveys are reductive of nonacademic outcomes (or, in one school leader's words, "suck the joy out of character"). Others fear that character assessment may be used to further alienate and penalize stakeholders (leaders, teachers, and children) in a system already rife with inequitable and punitive consequences connected with data (Soutter, 2020). Self-report surveys also suffer from being prone to a variety of biases; for example, reference bias (King & Wand, 2007), where individuals use different frames of reference to measure themselves or others when making a judgment; social desirability bias (answering

in ways that the participant thinks the researcher wants to see or hear; Arthur et al., 2016; Duckworth & Yeager, 2015; Park & Peterson, 2005); or, when rating others, implicit or personal bias (Peterson, Rubie-Davies, Osborne, & Sibley, 2016).

Given this measurement landscape, some organizations and researchers have attempted to move towards observation measures as a form of character and SEL assessment in educational settings in order to capture what is happening contextually in the classroom or school more broadly. Observation measures focus on classroom or school climate rather than student-level thoughts or behaviors, and often ask an administrator or teams to rate constructs such as teacher-student relationships or elements of classroom management. Holahan and Batey (2019) noted that classroom observations have been effectively used as a means of evaluating and improving adult-student relationships in out-of-school time activities, and that such tools are beneficial because they elevate "the importance of context, the learning environment, and adult-student relationships" (p. 18). To date, several character-oriented classroom observation tools have emerged from out-of-school time research or school-climate focused observation questions that are part of larger teacher assessment frameworks (Holahan & Batey, 2019). For example, the Weikart Center for Youth Program Quality (n.d.) has created an SEL program quality assessment for out-of-school time programs that assesses elements such as creation of safe spaces, scaffolded learning, and growth mindset ethos. Teacher frameworks that include school climate observational elements include the Danielson Framework for Teaching as well as the Marzano Observation Protocol (Yoder, 2014). More school specific SEL observation protocols include The Classroom Assessment Scoring System (CLASS) as well as CASEL's recent walkthrough protocol. CLASS specifically looks at elements such as the emotional climate of the classroom as well as behavior management (Allen et al., n.d.). CASEL's SEL walkthrough protocol offers a 4-point rubric that assesses classroom climate and practices (e.g., teacher-student relationships; cultural responsiveness) as well as school-wide climate and practices (e.g., youth voice and engagement; authentic family partnerships; CASEL, 2019). The walkthrough tool is specifically designed to help SEL teams or observers assess areas for improvement and to help schools focus on areas for "feedback and development" (CASEL, 2019, p. 1).

Importantly, classroom observations, too, suffer from drawbacks. Since observations are aggregated at the classroom or school level, this means that their results cannot be disaggregated by student group identities or demographics for either purposes of formative improvement or federal data reporting. Holahan and Batey (2019) have also noted the importance of using trained observers in order to ensure valid, reliable data, which, they note, can sometimes be a challenging bar to achieve for some school systems. In addition, as noted earlier, it is important to not measure contextual factors *alone* when measuring character as these cannot be used as proxies for character development; indeed, without a clear theoretical framework connecting contextual influences and character strengths, these kinds of observations can also be problematic (e.g., bias or noise, see Kahneman et al., 2021). Ultimately, triangulation between different forms of measurement—drawing upon several forms of measurement in order to draw patterns of interpretation between them—is one of the primary ways of dealing with the limitations of character measures while simultaneously allowing educators to capture and examine variations in student outcomes (Arthur et al., 2017, p. 112).

## *Equity in Measurement: Formative Assessment and Continuous Improvement*

**Delineating Between Summative Versus Formative Measures**. Beyond focusing on context, evidence also suggests that placing *learning* at the center of assessment practices—through formative assessment and a

focus on improving context and classroom practices—is associated with more equitable outcomes for students (Black & Williams, 2010; Kalinec-Craig, 2017; Pryor & Crossouard, 2008; Shepard & Penuel, 2018).

Indeed, although there are numerous types of character and SEL measures and assessments, here we are most concerned with the differences between formative and summative assessments and between high-stakes and low-stakes assessments (McKown, 2019). Summative and high-stakes measures are often longer, tend to come at the end of a defined period of learning (e.g., the end of the semester or school year) and may be used in an evaluative manner either by researchers or school systems to assess students' character strengths or character development. Formative and low-stakes measures are often shorter, may be used at any time throughout the learning period, are used to guide and improve instruction, and are meant to be authentic to the setting or behaviors in question (Duckworth & Yeager, 2015; McKown, 2017a, 2017b, 2019; Yeager et al., 2013). Importantly, there is no clear delineation or groupings of types of character measurements (e.g., surveys, teacher ratings) by whether or not they are high or low stakes, formative or summative. In addition, the terms high stakes and summative are not necessarily synonymous, nor are low stakes and formative. As such, in this manuscript we are focused on the *goals* for the use of the assessment rather than *type* of assessment or whether or not it is summative or formative. For example, McKown (2019) points out that assessing student character negatively on report cards, a summative assessment, could be interpreted as a form of high-stakes assessment should parents feel this could become a long-lasting stain on a students' permanent record. On the other hand, assessing students' character negatively on report cards could be used as a form of formative assessment if it is used to help encourage student growth throughout the school year and as a way to guide teachers in thinking about ways to improve instructions for student character growth. That is, the same form of assessment, character report cards, can be used for both high-stakes, summative and low-stakes, formative purposes, depending on the way it is used by the educator.

The improvement science movement in education is an example of the ways in which formative assessment principles are being used to create more equitable character education and SEL learning environments (The Carnegie Foundation, n.d.). Park et al. (2013) have described improvement science as "discern[ing what works for addressing a particular problem, for whom, and under what set of specific conditions" (p. 4). The six core principles of improvement science include: (1) make the problem specific and user centered; (2) address the variation in the problem; (3) see the system producing the outcomes; (4) "we cannot improve at scale what we cannot measure"; (5) engage in inquiry cycles (e.g., Plan, Do, Study, Act cycles) to improve; and (6) use networked communities to help accelerate improvements (Bryk et al., 2010). Continuous improvement then means making these principles of improvement regular and constant within one's organization, making them deeply interwoven into the daily work of the individuals of the organization, and working with individuals to continually situate their work within systems and contexts (Park et al., 2013). The California CORE educational districts specifically draw on the ideas of continuous improvement, using a nine-step continuous improvement cycle and drawing on networked communities to spread effective practices throughout the districts (Gallagher et al., 2019; Gallagher & Cottingham, 2019). The CORE districts began measuring four SEL competencies via student self-report survey in 2016 (growth mindset, self-efficacy, self-management, and social awareness) and additionally measured areas of school climate and belonging via student self-report survey (coredistricts.org). However, these CORE measures were used not only as measures of high-stakes accountability under the Every Student Succeeds Act, but, as noted here, also as formative

measures to help teachers think about ways to continuously improve their practice. For example, one CORE principal noted that they would look at the survey data "at the end of the year to 'do a cycle of continuous improvement, saying where did we rank, what can we do to improve participation and such'" (Marsh et al., 2018, p. 57).

**Using Summative Versus Formative Measures.** Thus, a school's character and SEL assessment journey should not end with collecting and analyzing data. Collecting and analyzing data is one element of the continuous improvement cycle and should serve to inform conversations about how to improve instruction and school climate and ethos (Annie E. Casey Foundation, 2014; Bryk et al., 2015; Schlund et al., 2020). For example, in their work with educators, Paunesku and Farrington (2020) developed an online survey system where educators can deploy surveys to their students regarding their SEL learning experiences; teachers then receive rapid feedback from the online system regarding their students' experiences surrounding feeling cared for, receiving feedback, and feeling engaged in their work. The system subsequently recommends appropriate strategies to help the teacher improve their students' experiences. Once the teacher tries out the new strategies the students can then be later assessed via the same online survey system to see if any changes occur. Thus, Paunesku and Farrington's (2020) model serves as an example of using continuous improvement strategies—via a formative survey assessment and rapid feedback—to help build more equitable classrooms. Carter and colleagues (2017) likewise recommend that data analysis in schools lead to continuous improvement and further conversation related to equity concerns. They noted that data should

> open a door to reflective and critical conversations about the ways in which school processes, adult actions, and adult interactions with students; about how race factors into how adults react to students and how students then react to adults; about which false or harmful notions about "races" we carry around with us … and even when and how thinking of other human beings in terms of race is helpful. (p. 220)

## *Equity in Measurement: Other Considerations*

Apart from attending to the contextual aspects of the school environment and using data formatively to improve learning and development, other data collection equity concerns arise when measuring character and SEL in schools. For example, assessments often use biased language, or the assessment method itself may have been found to be biased for or against some demographic groups (Jagers et al., 2018). Similarly, teachers may have implicit or explicit biases that emerge when rating or otherwise assessing students (Schlund et al., 2020). Accordingly, it is important to consider power dynamics within a school, classroom, or community, and to take into account who is collecting data from whom and to whom data will be shared (Schlund et al., 2020). In light of this concern, some schools and districts engage students, families, and/or their local communities in the data development and/or collection processes. For example, the Minneapolis Public Schools' Youth Participatory Evaluation teams, made up of middle and high school students, are involved in making recommendations regarding school climate and student engagement measures in their district (Schlund et al., 2020).

In addition, there are ways to *analyze* character and SEL data in a more equitable manner. For example, data should be disaggregated by race, gender, disability and other relevant demographics in educational environments in order to help identify which additional resources might be needed for which student groups (Annie E. Casey Foundation, 2014; Carter et al., 2017; Fiarman et al., 2021; National Forum on Education Statistics, 2016). For example, in Washoe County, Nevada, principals are provided with "data books" that guide educators regarding how to

interpret data related to school climate, engagement, discipline, and academic achievement "with an equity lens" (Schlund et al., 2020, p. 22). Accordingly, educators in Washoe County are asked to think about their data in a disaggregated manner and to consider whether both their educational strategies and their assessments are being applied in an equitable manner. Another example: over the 2015–2016 and 2016–2017 school years, California's Oakland School Unified District developed a school performance framework that drew upon an end of year school climate survey and other behavioral indicators (e.g., absences) in order to assess student growth and performance. This data was disaggregated by specific groups (e.g., ELL learners) and schools were given a guide in order to help them use the data to set priorities to lessen the achievement gap and battle inequities in their schools. In disaggregating data, however, it is of utmost importance to caution that disparities are not viewed in discriminatory ways (e.g., seeing a particular subgroup as lacking or less-than) but rather as a recognition of contextual changes that must be made to support these differences. For example, if data show that a certain group of students report a lower sense of belonging, this should not be seen as a deficit of that group of students, but rather as an area that the school or classroom itself must attend to (e.g., Kaler-Jones, 2020; Simmons, 2019; Soutter, 2020; Zaino, 2016).

## *EL EDUCATION AND EQUITABLE MEASURES OF CHARACTER*

Like others in the field of character education and social emotional learning, EL Education —a national network of primary and secondary schools—has become increasingly interested in finding and creating measures that assess the impact of their educational model on students' character change and growth while informing school improvement efforts; to date they have pursued both summative, higher stakes character assessments as well as lower stakes, formative assessments. However, intertwined within EL Education's character assessment model is a deep commitment to developing equitable measures of character that not only assess the student and their character development, but also the contexts within which EL Education students are studying in order to improve upon them.

Below, we describe EL Education's character framework followed by its approach to measurement as a way of illustrating the ways in which this organization is using an equity-driven, socioculturally grounded, practitioner-informed approach to character measurement for equity and continuous improvement. We first outline EL Education's conceptualization of character education and equity, and then describe their character measurement goals. Subsequently, we describe the ways in which EL Education's approach to measurement is aligned with a commitment to equity through its prioritization of contextual factors and continuous improvement cycles, and a co-construction of measures that values practitioner input.

It is important to note in the sections below that the process embarked upon by EL Education is grounded in the basic principles of measurement design, but was intended for practical use. That is, their primary goal was to compile and create tools that were streamlined and actually usable in schools. So while their process was rigorous, our aim in outlining aspects of their undertaking is not necessarily to focus on the particulars of their psychometric design, but rather to focus on how school leaders might embark on a similar process of creating equitable, formative measurement tools that are actually practical and not prohibitively time consuming.

In addition, it is not necessarily the uniqueness of EL Education's measures that we aim to highlight here; though we emphasize their use of a walkthrough tool, we recognize (as described above) that other scholars have aimed to similarly take stock of students' classroom contexts (e.g., CASEL, 2020; Paunesku & Farrington, 2020). Rather it is the

explicit commitment to measuring character in such a way that not only prioritizes character as a critical indicator of success, but that aims to do so in an equitable way. EL Education's efforts to account for systemic and contextual factors, and to use disaggregated data for continuous improvement is a relatively new endeavor, and one that centers equity concerns in a way that so many other efforts do not. Our hope is that practitioners and school leaders alike can learn from this case study when conceptualizing their own approach to equitable character measurement.

## *EL Education's Conceptualization of Character Education and Equity*

Grounded in a belief that the purpose of school is to prepare students not to perform on tests but to be critical thinkers with a strong sense of character, EL Education champions a broader definition of student achievement that prioritizes character as a critical, interlinking dimension. More specifically, successful achievement is defined as a combination of mastery of knowledge and skills, high quality student work, *and* character (see Figure 1).

In order to paint a clearer picture of what they mean by *character,* this piece is expanded upon in EL Education's Character Framework that is grounded in three core dimensions (effective learners, ethical people, contributing to a better world) surrounded by the necessary conditions for these areas to thrive (see Figure 2). As evidenced below, EL Education maintains an explicit focus on equity in each foundational piece of their conceptualization of character, a commitment that remains a driving force of their measurement goals as well.

### *Effective Learners, Ethical People, Contribute to a Better World*

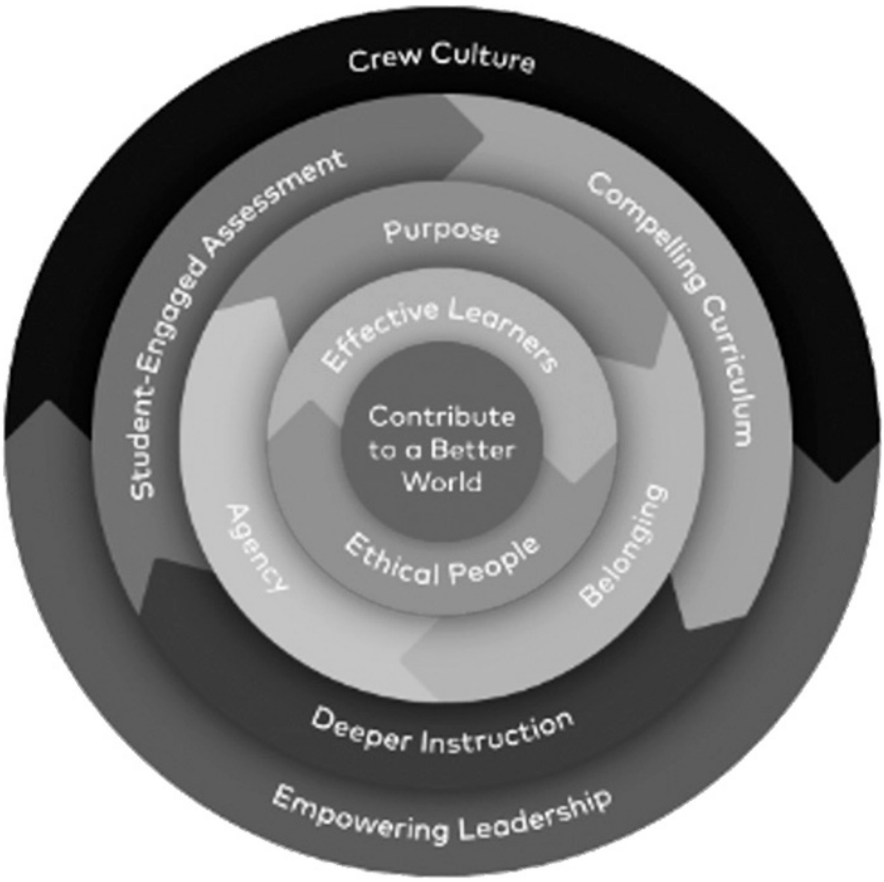
At its core, EL Education supports students to become effective learners and ethical people who contribute to a better, more equitable world. They define effective learners as those who "do more than they think possible, developing the mindsets and skills for success in college, career, and life" and as being comprised of character strengths such as initiative, responsibility, perseverance, and collaboration. Ethical people are described as those who "treat others with respect and compassion, and stand up for what is right" and who personify character strengths such as empathy, integrity, respect, and compassion. Finally, at the center of the framework is the message that EL Education students work to contribute to a better world "putting their learning to use as active citizens, working for social justice, environmental stewardship, and healthy, equitable communities." This central goal is inherently focused on equity and the greater good, including character strengths such as civic responsibility and service, civic empowerment, long-term purpose, social justice, and environmental stewardship.

### *Belonging, Purpose, Agency*

The next layer of EL Education's framework focuses on fostering a *positive student identity* by emphasizing belonging, purpose,

FIGURE 1
EL Education Three-Dimensional Vision of Achievement

FIGURE 2
EL Education Character Framework




and agency to create the necessary *conditions* for students to become effective learners and ethical people who contribute to a better world. Each of these surrounding layers are explicitly focused on ensuring that *all* students are supported in these ways, and, as illustrated below, when measuring these constructs, equitable outcomes for all students is a central goal.

### *Compelling Curriculum, Student Engaged Assessment, Deeper Instruction*

In order to truly prioritize character as an equally important part of school, EL Education embeds character education through all aspects of the day. Their schools implement a compelling curriculum where teachers use protocols to create opportunities for all voices to be heard, and activities and texts that honor the knowledge, languages, beliefs, and skills from the cultures and backgrounds of students and their families. EL Education also leverages student-engaged assessment practices to promote character and equity, tracking progress toward not only standards-based learning, but toward character targets as well. Students set goals, reflect on growth and challenges, and present evidence of their achievement and growth to a range of audiences. Finally, teachers in EL Education

schools engage students in deeper instruction that challenges and empowers students in meaningful work in order to build authentic understanding.

### *Crew Culture and Empowering Leadership*

All of the character framework components described above rest on a foundation of crew culture and empowering leadership, which are inextricably linked to EL Education's character and equity goals. Crew culture is a defining feature of EL schools and is organized around activities designed to facilitate community-building, a sense of belonging, and character development. These activities range from advisory-style check-ins and discussions, to group games designed to build teamwork and foster reflective discussions, to explicit conversations about students' character. Finally, this kind of environment is bolstered by EL Education's approach to empowering leadership where school leaders (including teacher-leaders) work collaboratively with families, students, and staff to make evidence-based decisions that enable all students to achieve and that build a schoolwide community of collaboration and trust.

### *EL Education's Approach to Character Measurement*

EL Education's framework serves as a guiding force not only for their approach to teaching, but also for the way they measure success. As noted above, our goal in detailing this particular institution is to highlight two overarching tenets of their measurement approach that we see as critical considerations for the field: their commitments to equity through contextual factors and continuous improvement. Below we detail the ways in which they center these two focal points in thoughtful and impactful ways in the design, development, and implementation of their measurement tools. The organization ultimately uses two forms of measurement (student self-report surveys and school leader administered walkthrough tools); however, we focus primarily on their walkthrough tools as they are an approach that is more directly and intentionally used as a way of addressing contextual considerations and fostering conversations around continuous improvement.

### *Equity in Measurement: Contextual Considerations*

Consistent with their strengths-based, holistic view of character, EL Education wanted to design a measurement tool that could take context into account. Instead of relying on measurement tools that could further entrench an individualistic, immutable vision of character (e.g., where the onus is solely on the student to improve), they pursued a form of measurement that acknowledges character as an ever-changing element of each student's identities evolving within their sociocultural contexts**.**

More specifically, when designing measurement tools, EL Education was particularly sensitive to the ways in which character has been conceptualized and measured historically and how that presented significant equity concerns; as detailed earlier, when character is viewed as innate and unchangeable, and measured accordingly through survey tools, then there is no need to create classroom environments in which students' character can grow and be nurtured. If, on the other hand, character is seen as a set of malleable skills, dispositions, and mindsets that depend in part on contextual factors (e.g., student experiences outside the classroom, their feelings about certain school subjects, their relationships with peers and teachers) as well as developmental opportunities for growth, then educators can take on a strengths-based (vs. deficit-based) approach to teaching and cultivating character.

Aligned with this vision of character, EL Education wanted to look beyond survey tools alone since they can be limited in this regard. As one way to keep contextual influences at the forefront of their vision of character and to design measurement tools that could address

equity concerns, EL Education developed a character observation tool, called a walkthrough. Walkthroughs allow observers to take context into account in a way that survey tools typically do not. For example, if one student is having a challenging moment in class, the observer might record not only that child's behavior, but also the ways in which the rest of the class responded to and supported that student (or, potentially fell short in doing so). In this way, the walkthrough tool attends to the ecosystem of the classroom and school community and allows for the examination of the extent to which students are able to demonstrate their character strengths and support one another's development as well as the extent to which the teacher is providing meaningful support and opportunities for character development.

Walkthroughs were also specifically designed because they provided information about the extent to which character could be observed across the entire school (e.g., the extent to which teachers were promoting and supporting habits of character during all parts of the school day, and the extent to which students were developing these strengths), and because it positioned school leaders to get into classrooms to be able to provide formative feedback and appropriate supports to their staff around character development.

To construct character walkthroughs, schools first solidified their central character goals. That is, before any measurement tools could be considered, a clear definition of core values was decided upon at each participating school. Leaders collaborated with faculty, students, and families to name and define what habits of character best exemplify the values and skills that they aspire to build in their school community, and how those align with the core EL Education principles. For example (as seen in Table 1), they came to understand the constructs of 'empathy and perspective taking' as "the ability to consider the perspectives and share the feelings of others." However, as is also described in the following section, arriving at the understanding of what these terms mean was an iterative, collaborative process that allowed each school to interpret these constructs to fit their own contexts. In other words, the instrument each school uses is the same, but the interpretation of each character strength and how they will be rated may vary across schools.

EL Education then conducted a review of currently existing and robust observation tools and adult reports (e.g., teacher observations). Existing observation tools often require users to observe for long periods of time, require retrospective judgment, or cannot be used across a variety of contexts; in contrast, the EL Education walkthrough is meant to be quick (typically about 7–10 minutes), frequent, and should allow for many observations in many classes over the course of a day across a variety of contexts (e.g., not just in the classrooms, but in the cafeteria, in the hallways at the end of the day, in physical education, art, etc.) and developmental stages. In light of these constraints, EL Education referenced existing observation instruments, but also weighed what behaviors could feasibly and consistently be observed in school contexts within a short time window. Ultimately, they designed walkthrough tools aligned with two of EL Educa-

TABLE 1
Sample Measured Constructs Aligned With One of EL Education's Central Goal of "Becoming Ethical People"

| *Measured Construct* | *Definition* |
|---|---|
| Empathy and perspective taking | The ability to consider the perspectives and share the feelings of others |
| Peer leadership | To engage in courageous, initiative-taking behaviors that compel others toward a common goal |

FIGURE 3
Sample Walkthrough Tool for Empathy and Perspective Taking and Peer Leadership

| 3. Empathy & Perspective Taking - Teachers | | | Evidence For | Evidence Against |
|---|---|---|---|---|
| 3A. The teacher provides opportunities for students to share ideas and perspectives. | | | | |
| 3B. The teacher provides, models, and reinforces perspective-taking language and behaviors. | | | | |
| **Overall, my teachers effectively support students' empathy and perspective-taking.** | | | | |
| Not Evident | Somewhat Evident | Evident | N/A | |
| **4. Empathy & Perspective Taking - Students** | | | **Evidence For** | **Evidence Against** |
| 4A. Students are willing to listen to and show appreciation for multiple perspectives. | | | | |
| 4B. Students notice and respond empathetically to the emotional states of others. | | | | |
| 4C. (6-12 only) Students engage intellectually and productively with different ideas or points of view. | | | | |
| **Overall, my students demonstrate effective empathy and perspective-taking.** | | | | |
| Not Evident | Somewhat Evident | Evident | N/A | |
| **5. Peer Leadership - Teachers** | | | **Evidence For** | **Evidence Against** |
| 5A. The teacher (or the school) provides opportunities for students to take leadership roles in collaborative work. | | | | |
| **Overall, my teachers effectively support students' peer leadership.** | | | | |
| Not Evident | Somewhat Evident | Evident | N/A | |
| **6. Peer Leadership - Students** | | | **Evidence For** | **Evidence Against** |
| 6A. Students motivate (inspire, appreciate and encourage) their team members to meet their common goals. | | | | |
| 6B. (6-12 only) Students hold their team members accountable for their role to meet their common goals. | | | | |
| **Overall, my students demonstrate peer leadership.** | | | | |
| Not Evident | Somewhat Evident | Evident | N/A | |

tion's central tenets: *becoming effective learners* and *becoming ethical people*. (The third central tenet, *contribute to a better world*, is embodied by character strengths that are more difficult to capture in discrete behaviors in the classroom and so is currently only measured through survey items, and is more holistically on display during their yearly "Better World Day"). Table 1 shows a sample of two of the constructs that are measured through the walkthrough tool that are aligned with becoming ethical people, and Figure 3 shows the corresponding walkthrough tools.

One of the ways that EL Education evaluated the ability of these walkthrough tools to consider contextual factors was through a collaborative process with several EL Education schools: EL Education staff provided feedback on initial drafts for face validity (meaning, did the tool actually measure the character strengths it was aiming to capture), clarity, and the ability to complete the walkthroughs in a short time window. Subsequently the tools were user-tested in two pilot schools with the respective school leadership teams who identified that the indicators needed to be better clarified and that some indicators were difficult to observe consistently or that they varied developmentally. For example, regarding the crew walkthrough (looking at an EL Education school structure focused on weekly meetings between groups of students and an educator), educators noted that terms such as "circling up" and greeting structures, as well as language surrounding other particular rituals, traditions, and crew norms were missing from the initial walkthrough. In addition, when first testing out the effective learners walkthrough, educators at Amana Academy pushed EL Education to continue to clarify their terms. For

instance, one section asked educators to assess whether students are sticking "to a tough problem, even when confronted with mistakes, uncertainty, or challenge"; the Amana Academy educators wondered in their piloting notes "What is a tough problem?" In other words, actually having the opportunity to test this tool allowed teachers to see which character constructs were perhaps more difficult to identify or measure, which then allowed researchers to reconsider how items were presented. After another round of revisions, the walkthrough tool was again piloted in eight additional schools and feedback was received via survey and through school team documents, which informed a final round of revisions. For example, in order to further clarify for educators what they should look for regarding student persistence through difficulty, the language "students vocalizing frustration or difficulty, but staying engaged and on-task" was added to the walkthrough tool.

Another example of calibration and practitioner feedback on contextual factors occurred while piloting the walkthrough tool in an elementary classroom. During the walkthrough, observers noted that the majority of the students were sitting on the carpet analyzing a poem together. After they left the classroom, the group circled up to debrief and calibrate on their ratings for whether the instruction was engaging, challenging, and empowering in inclusive and equitable ways. Of the five observers who engaged in the walkthrough, three noted that equitable participation was not evident, but two noted that it was evident. The team then went around to explain their evidence ratings. The three observers that noted "not evident" explained that they noticed three African American boys sitting in chairs apart from the rest of the crew who were not engaged in any meaningful learning activity or paying attention to what the rest of the group was doing. "Why were they separated, and why weren't they engaged in any kind of learning activity?" one observer asked. The school leaders noted that the three boys were not engaged in the whole group because they were waiting for a specialist to come pick them up for scheduled pull-out time. The ensuing discussion was especially productive. The team raised important questions about how the boys might experience the waiting while class was going on in front of them. Was there a missed opportunity for inclusion in meaningful learning during this wait time? How might this have felt for the children, who were physically separated and disengaged from the rest of the group? And what alternatives might be introduced by the teacher and instructional guide that can ensure that students who are waiting for pull-outs are still supported and included physically, socially, and intellectually as part of the larger group?

While this particular conversation did not serve to recalibrate the instrument itself, it did serve as an example of one of the ways that the walkthrough could be used to recognize contextual disparities and then have meaningful discussions on the roots of these inequalities and what changes could be made to address these. What is important about these conversations as an outcome of the walkthrough instrument is that the tool itself acts as a catalyst for recognizing, identifying, reflecting on, and enacting change.

In addition, EL Education school leaders received training on recognizing issues of equity. Educators trained to engage in the walkthroughs were urged to consider questions such as, "who is the teacher calling on?," "If students are doing group work, is every child meaningfully involved?", and to "pay attention to how [students] are given access to the content and curriculum, how their specific language needs are met, and perhaps most importantly, whether they speak at all" (Aguilar, 2015).

These kinds of discussions not only centered equitable support for all students, but they also allowed teachers to further solidify the nuances of inclusion and belonging and how it was envisioned at their school. It is also critically important to note that these walkthrough tools are designed to be ways of starting conversations about character devel-

opment in the classroom, not as high-stakes evaluations for the students or for the teachers themselves. This emphasis on formative measurement for continuous improvement is expanded upon in greater detail below.

### *Equity in Measurement: Formative Assessment and Continuous Improvement*

EL Education's measures were designed not only to demonstrate proof of impact, but also to be used formatively in the classroom to better understand and track character growth in order to drive instructional change. That is, instead of using measurement tools as a high-stakes evaluation of students' character, these tools are meant to be used as a way to support a reflective process where teachers and students alike are considering how they can continuously grow and improve as individuals and as a community. Below we describe the ways in which EL Education ultimately uses these measurement tools in practical, data-driven ways that center equity in their application and interpretation.

First, EL Education uses the walkthrough tool as an opportunity for staff and students to further solidify a shared vision of what character looks like in their schools. For example, similar to the conversations that ensued when providing feedback on the walkthrough tool, when calibrating on a walkthrough indicator (for example, an indicator might ask observers to look for examples of students being supported in being persistent), faculty can unpack "what does this look like and sound like for us?" This vision-setting can be supportive for buy-in to both the tool and the data it produces, help set high expectations for teacher practice and student outcomes, and concretize the expectations for what every classroom should feel like for every student. For example, Cherisse Campbell, principal of Amana Academy in Georgia, described performing a walkthrough in her school community as offering an opportunity to inform her staff of "who we are or where we are, right now, and this is how we can support each other going forward." She noted that walkthroughs are a particularly powerful tool for her educators because they offer a glimpse into the contextual supports in the school community that can be further leveraged to help promote positive student outcomes. She commented:

> We sometimes do [walkthroughs] where we take the entire day and take grade levels into different grade levels. At the end we come back and they do individual noticings and wonderings. But what's so powerful is we take all those [wonderings] and we ask our staff to find those trends. So instead of us, as a leadership team, going and making judgments around the data we're seeing, we ask them to look for, what do you see as a strength, what do you see as a growth area. They can see, this is who we are, and we have the power to improve this together. Also they could look at, you know what, there are some practices that are happening in this building that I can adopt in my classroom. The development of EL Education's Character Surveys and walkthroughs allow EL Education to now engage in a fully data-driven cycle of improvement, driven by user-centered inquiry and collaborative teams. By creating practical measures that have been created in concert with educators and practitioners through the EL Education network, EL Education is focused on capturing the holistic experiences and expertise of the educators and students who comprise their network.

After analyzing data, EL Education educators are asked to engage in inquiry to create, test, and measure the impact of a new idea. Based on their root cause analyses, teachers engage in an inquiry cycle and decide on a new change that they are going to test out in the educational setting in a frequent and structured format. Their inquiry teams then meet frequently in order to discuss how this new change is functioning and whether changes should be made or if the impact is positive and as expected; the change is then continued, revised, or other new practices are implemented. For example, in an inquiry meeting, Cherisse Campbell at Amana Academy and her team of teachers were able to discuss data from a recent walkthrough and discuss possible pedagogical implications:

> Principal: We believe in student engaged instruction, we believe in teacher engaged instruction. For example, today when we were talking through our debrief we were wondering why we weren't seeing as many debriefs as we'd expected.
>
> Teacher: Debrief would be an area of growth for us, when we're bringing students back together at the end of the lesson, to revisit the learning target and revisit the purpose of the learning for the day. We have that at about 25% of the classrooms that we visited today.
>
> Principal: So I kind of have a wondering about that, why do you think that's so low? *Teacher*: I think we need to consider looking at pacing and times of the lessons, and in my conversations with teachers, this has been a theme that keeps coming up, is that they need to focus more on their pacing and timing throughout the lesson. So it's not that they don't want to do it, it's just a matter of sometimes the time is up for that lesson, and then the debrief is the one, is that thing that gets put on the back burner.

Thus, through looking at the walkthrough data, the principal and her team were able to pinpoint the area of reflections as an area for growth, as well as a strategy to better support their faculty to plan for high-leverage strategies such as debrief conversations. Similarly, Laina Cox, principal at Capital City in Washington, DC, described the process of examining data as, "It is just what do you notice about the data? Where is it working better and how can we then leverage that across other grade levels?"

In the end, change is then spread by other schools and networks adhering to these new ideas and improvements brought about through these plan-do-study-act cycles. As part of their 2017 study on crew belonging, EL Education created a spring convening so that each participating school could share their own poster and Powerpoint presentation regarding what they learned from their data, challenges, and their new plans based on their plan-do-study-act cycle. EL Education has continued to host webinars for their network to help build their educators ability to apply new ideas and practice in the learning context. For example, as part of their work with the Raikes Foundation and their Building Equitable Learning Environments study during the 2018–2019 school year, EL Education hosted webinars such as "using survey data with students," "equitable teaching practices," and "belonging, identity, and community." Moreover, they developed professional learning communities across their network, with each school forming a pair with another participating network school. The aim of such professional learning communities is for each to be able to learn from others' practice, engage in shared problems of practice, and exchange improvement ideas and structures.

## *Equity in Measurement: Other Considerations*

In addition to taking into account practitioner feedback as a part of their commitment to continuous improvement, EL Education also incorporated student participation and feedback as part of their development process. While student voice was more a part of the development of survey measures, we include it here as an example of the ways in which students were included in a meaningful way.

First, in addition to assessing the appropriate age range for each survey by checking lexile levels for each survey, students in Grades 4–7 were also interviewed. The purpose of these interviews was not only to see if the survey levels were developmentally appropriate, but also to see if they accurately captured students' contextualized, holistic character strengths. When these kinds of contextual, socio-cultural considerations are not prioritized and asked about, survey outcome measures can be misinterpreted, often with subgroups scoring lower because they are being compared to White, Eurocentric expectations of behavior and character (Kaler-Jones, 2020; Simmons, 2019; Soutter, 2020; Zaino, 2016).

These interviews consisted of students being read aloud the survey items and asked what each one meant to them, resulting in important considerations for survey dissemination. For example, one of the original *Ethi-*

*cal People* survey items read, "I try to understand rather than judge people who are strangers to me," and only 41.6% of students were accurately able to interpret this item: one fourth grade student incorrectly described this item as "Judging someone means to tell them what to do and you are trying to tell them to do the right thing" whereas a seventh-grade student correctly noted "I try to be considerate and not think bad thoughts if I don't know the person." Similarly, one of the questions in the *Contributing to a Better World* survey, "It is important to fight against social and economic inequality," was not able to be accurately interpreted by any of the students interviewed: one Grade 5 student, for example, did not know the terms economic or inequality and described it as "Like if people are giving you a hard time" and a sixth-grade student similarly struggled with the term economic and said that inequality meant "to try hard." Given such findings, EL Education currently recommends that their surveys are most appropriate for students in the sixth-grade and above, and that school staff should review key terms and concepts with students prior to distributing the survey.

These student interviews reinforce the tension between psychometric and practitioner approaches to measurement. Despite passing psychometric reliability analyses with datasets including younger students, it was clear upon closer inspection that statistical methods alone do not sufficiently address whether measures are developmentally appropriate for linguistically, developmentally, culturally, and academically diverse groups. This round of piloting, coupled with the solicitation of student feedback, signals a true valuing of student comprehension and perspectives, cultural considerations, and an active commitment to equity and student voice.

During piloting, teachers also asked for more demographic breakdowns by schools, grades, classes, and subjects in order to be able to analyze the data more precisely. The purpose of gathering these demographics was of course not to reinforce stereotyped assumptions about subgroup behaviors, but instead to see where the teachers themselves and the school environment could adapt to better support students' needs. For example, if a certain group was reporting a lower sense of belonging, this would be important data for schools to reflect upon to see what steps need to be taken to foster increased inclusion and trust to bolster a true sense of belonging. EL Education ultimately designed a data system that allows them to disaggregate character measurement data by subgroups in order to evaluate whether their approach was serving all students equally well, and, if not, to make adjustments accordingly. This database has been a key piece in allowing for the kind of continuous improvement that is central to its primary goals of character measurement as described above. For example, in a recent investigation into the impact of Crew on equity in EL Education schools, they were able to use disaggregated data to track an increase in sense of belonging for students of color; while there continues to be much work to be done to ensure that all students feel a strong sense of community in schools, these data provided useful information to begin tracking their own progress (EL Education, 2018).

## *IMPLICATIONS FOR CHARACTER MEASUREMENT IN THE FIELD*

The question of character measurement is not just a question of how to do it accurately (though this is certainly an endlessly complex question rife with challenges as outlined above), but how to hold ourselves accountable as educators to ensure that character and equity are centered in our work. Indeed, measuring character is not something that should be taken on lightly or without a firm understanding of the complexity of the ways in which the benefits, drawbacks, and contextual factors intersect. We must enter this work with a clear recognition of the ways in which measurement can be misinterpreted if not conceptualized within a sociocultural, race-conscious,

equity-based lens (Kaler-Jones, 2020; Simmons, 2019; Soutter, 2020; Zaino, 2016).

Our goal in highlighting this case study of EL Education is not to lift up one example of perfect measurement, but rather to illustrate the ways in which they have approached this feat through an intentional, collaborative process that prioritized equity, contextual considerations, and continuous improvement at every turn. Rather than seeing character measurement as a summative, high stakes way of getting students to conform to benchmarks grounded in White, Eurocentric moral and behavioral expectations, EL Education's approach supports teachers in seeing their individual students' strengths, to examine how their classroom community can shift to better support them, and to truly hold character as a critical, interlinking piece of student achievement. Certainly, their approach does not encompass all equity considerations nor does it answer the many outstanding questions and critiques related to character measurement as a field. Indeed, schools and school leaders must be painfully careful to ensure that character measurement does not become weaponized or used to idealize a singular way to behave, act, or feel.

However, we also must recognize that if we continue to measure academic outcomes alone, then these become our sole indicators of success. If we as educators truly care about centering holistic student well-being, we must critically examine our pedagogy, ask ourselves difficult questions, engage in inquiry cycles, reflect on our successes and failures, and work collectively to hold ourselves accountable to centering character in just, holistic, equitable ways.

## *NOTE*

1. There are innumerable ways to define and understand character and the associated field of SEL. In our work we conceptualize *character* as falling into four overarching areas: moral, performance, intellectual and civic (Baehr, 2015; Lerner & Callina, 2014; Seider, 2012; Shields, 2011) and we align our understanding of SEL with CASEL's (2021) five competencies: self-awareness, self-management, relationship skills, social awareness, and responsible decision-making. EL Education—the focus of this article—uses the terminology "character" and grounds their measures in each of their schools' understandings of this term. However, it is important to note that they also recognize the overlap between character and the related field of SEL (Berkowitz et al., 2017); hence, we draw upon both of these research literatures.

## *REFERENCES*


Allen, J. P., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. C. (n.d.). *Predicting adolescent achievement with the CLASS™-S observation tool* (research brief). University of Virginia. https://education.virginia.edu/sites/default/files/uploads/resourceLibrary/Research_brief_CLASS-S4.pdf

Aguilar, E. (2015, February, 20). *How to coach for equity in schools, Part 1* [Blog post]. https://www.edweek.org/education/opinion-how-to-coach-for-equity-in-schools-part-1/2015/02

The Annie E. Casey Foundation. (2014). *7 steps to advance and embed race equity and inclusion within your organization.* https://www.aecf.org/resources/race-equity-and-inclusion-action-guide/

Arthur, J., Kristjánsson, K., Harrison, T., Sanderse, W., & Wright, D. (2016). *Teaching character and virtue in schools*. Routledge.

Assessment Work Group. (2019). *Student social and emotional competence assessment: The current state of the field and a vision for its future.* Collaborative for Academic, Social, and Emotional Learning

Atwell, M. N., & Bridgeland, J. M. (2019). *Ready to Lead: A 2019 update of principals' perspectives on how social and emotional learning can prepare children and transform schools. A report for CASEL.* https://files.eric.ed.gov/fulltext/ED602977.pdf

Baehr, J. (2015). *Cultivating good minds: A philosophical & practical guide to educating for intellectual virtues.* www.intellectualvirtues.org

Berg, J., Osher, D., Same, M. R., Nolan, E., Benson, D., & Jacobs, N. (2017). *Identifying, defining, and measuring social and emotional*

*competencies* [Final report]. American Institutes for Research. https://www.air.org/sites/default/files/downloads/report/Identifying-Defining-and-Measuring-Social-and-Emotional-Competencies-December-2017-rev.pdf

Berkowitz, M. (2011). What works in values education? *International Journal of Educational Research, 50*(2011), 153–158.

Berkowitz, M. W., Bier, M. C., & McCauley, B. (2017). Toward a science of character education. *Journal of Character Education*, *13*(1), 33–51.

Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *92*(1), 81–90.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Harvard University Press.

Bronfenbrenner, U. (1993). The ecology of cognitive development: Research models and fugitive findings. In R. H. Wozniak & K. W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments* (pp. 3–44). Erlbaum.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2010). *Getting ideas into action: Building networked improvement communities in education. Carnegie perspectives.* Carnegie Foundation for the Advancement of Teaching. https://files.eric.ed.gov/fulltext/ED517575.pdf

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Harvard Education Press.

Campbell, D. E. (2012). Introduction. In D. E. Campbell, M. Levinson, & F. M. Hess (Eds.), *Making civics count: Citizenship education for a new generation* (pp. 1–13). Harvard Education Press.

Carnegie Foundation. (n.d.). *The six core principles of improvement.* https://www.carnegiefoundation.org/our-ideas/six-core-principles-improvement/

Carter, P. L., Skiba, R., Arredondo, M. I., & Pollock, M. (2017). You can't fix what you don't look at: Acknowledging race in addressing racial discipline disparities. *Urban Education*, *52*(2), 207–235.

CASEL. (2019). *Guide to schoolwide SEL. TOOL: Indicators of schoolwide SEL walkthrough protocol.* https://schoolguide.casel.org/uploads/sites/2/2020/10/Indicators-of-Schoolwide-SEL-Observation-Protocol-5.8.20-FINAL.pdf

CASEL. (2021). https://casel.org

Costa, P.T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO- FFI) professional manual*. Psychological Assessment Resources.

Dewey, J. (1938). *Experience and education.* New York, NY: Free Press.

Dewey, J. (1966). *Democracy and education.* New York, NY: Free Press. (Original work published 1916)

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.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432.

EASEL Lab. (2019). *Navigate the complex field of social emotional learning*. Explore SEL. http://exploresel.gse.harvard.edu/

The Education Trust. (2020). *Social, emotional, and academic development through an equity lens.* https://docs.google.com/document/d/1nfSQ6BoKR-byCkkN3rTEth_QcPdVXSo20Pi_vzXX55U/edit

EL Education. (2018). Building equitable learning environments report.

Fiarman, S., Kyles-Smith, K., & Lee, A. (2021). Is your approach to continuous improvement colorblind? *Educational Leadership, 78*(6), 16–21.

Fisher, J. F. (2020, February). *5 ways to ensure SEL is an instrument for equity of opportunity.* Christensen Institute. https://www.christenseninstitute.org/blog/5-ways-to-ensure-sel-is-an-instrument-for-equity-of-opportunity/

Gallagher, H. A., & Cottingham, B. W. (2019). *Learning and practicing continuous improvement: Lessons from the CORE Districts.* Policy Analysis for California Education. https://www.edpolicyinca.org/sites/default/files/R_Gallagher_Oct19_0.pdf

Gallagher, H., Gong, A., Hough, H. J., Kennedy, K., Allbright, T., & Daramola, E. J. (2019, January). *Engaging district and school leaders in continuous improvement: Lessons from the 2nd year of implementing the CORE improvement community.* Policy Analysis for California Education. https://www.edpolicyinca.org/sites/default/files/CORE_Gallagher_Jan-2019.pdf

Hamilton, L. S., Doss, C. J., & Steiner, E. D. (2019). *Handbook of Child Psychology and Developmental Science: Vol. 1. Teacher and principal perspectives on social and emotional learning in America's schools: Findings from the American educator panels.* RAND. https://www.rand.org/pubs/research_reports/RR2991.htm

Holahan, C., & Batey, B. (2019). *Measuring school climate and social and emotional learning and development: A navigation guide for states and districts.* Council of Chief State School Officers. https://files.eric.ed.gov/fulltext/ED597917.pdf

Jagers, R., Rivas-Drake, D., & Borowski, T. (2018). *Toward transformative social and emotional learning: Using an equity lens.* CASEL. https://measuringsel.casel.org/wp-content/uploads/2018/11/Framework_EquitySummary-.pdf

Johanek, M. C. (2012). Preparing pluribus for unum: Historical perspectives on civic education. In D. E. Campbell, M. Levinson, & F. M. Hess (Eds.), *Making civics count: Citizenship education for a new generation* (pp. 57–87). Harvard Education Press.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment.* Little, Brown Spark.

Kaler-Jones, C. (2020). When SEL is used as another form of policing. *Medium.* https://medium.com/@justschools/when-sel-is-used-as-another-form-of-policing-fa53cf85dce4

Kalinec-Craig, C. A. (2017). The rights of the learner: A framework for promoting equity through formative assessment in mathematics education. *Democracy and Education*, *25*(2), 5.

King, G., & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis*, *15*(1), 46–66.

Kirshner, B. (2015). *Youth activism in an era of education inequality.* NYU Press.

Krachman, S. B., & LaRocca, B. (2017). *The scale of our investment in social-emotional learning. A working paper.* Transforming Education. https://files.eric.ed.gov/fulltext/ED605377.pdf

Lee, A., & Riordan, M. (August 17, 2018). Using continuous improvement to disrupt inequity and promote deeper learning. *Education Week.* https://www.edweek.org/leadership/opinion-using-continuous-improvement-to-disrupt-inequity-and-promote-deeper-learning/2018/08

Lerner, R. M., & Schmid Callina, K. (2014). The study of character development: Towards tests of a relational developmental systems model. *Human Development, 57*(6), 322–346.

Lerner, R. M., Lerner, J. V., Bowers, E., & Geldhof, G. J. (2015) Positive youth development and relational developmental systems. In W. F. Overton & P. C. Molenaar (Eds.), *Theory and method* (7th ed., pp. 607–651). Wiley.

Love, B. (2019). *We want to do more than survive: Abolitionist teaching and the pursuit of educational freedom.* Beacon Press.

Mahoney, J. L., Durlak, J. A., & Weissberg, R. P. (2018). An update on social and emotional learning outcome research. *Phi Delta Kappan, 100*(4), 8–13.

Marsh, J. A., McKibben, S., Hough, H. J., Allbright, T. N., Matewos, A. M., & Siqueira, C. (2018). *Enacting social-emotional learning: Practices and supports employed in CORE districts and schools.* Policy Analysis for California Education. https://files.eric.ed.gov/fulltext/ED591088.pdf

McKown, C. (2017a). Promises and perils of assessing character and social and emotional learning. *Journal of Character Education*, *13*(2), 47–52.

McKown, C. (2017b). Social-emotional assessment, performance, and standards. *Future of Children, 27*(1), 157–178.

McKown, C. (2019). *Assessing students' social and emotional Learning: A guide to meaningful measurement (SEL Solutions Series) (Social and Emotional Learning Solutions)*. WW Norton & Company.

McClellan, B. E. (1999). *Moral education in America: Schools and the shaping of character from colonial times to the present.* Teachers College Press.

National Center on Safe Supportive Learning Environments. (n.d.). *School climate survey compendium: K–12 surveys.* https://safesupportivelearning.ed.gov/school-climate-survey-compendium-k-12-surveys

National Forum on Education Statistics. (2016). *Forum guide to collecting and using disaggregated data on racial/ethnic subgroups* (NFES 2017-017). U.S. Department of Education. National Center for Education Statistics.

Osher, D., Cantor, P., Berg, J., Steyer, L., & Rose, T. (2020). Drivers of human development: How relationships and context shape learning and development. *Applied Developmental Science, 24*(1), 6–36.

Park, S., Hironaka, S., Carver, P., & Nordstrum, L. (2013). *Continuous improvement in education.*

*Advancing teaching--improving learning. White paper.* Carnegie Foundation for the Advancement of Teaching. https://files.eric.ed.gov/fulltext/ED560145.pdf

Park, N., & Peterson, C. (2006). Moral competence and character strengths among adolescents: The development and validation of the Values in Action Inventory of Strengths for Youth. *Journal of Adolescence*, *29*(6), 891–909.

Paunesku, D., & Farrington, C.A. (2020). Chapter 4: Measure learning environments, not just students, to support learning and development. *Teachers College Record*, *122*(14), 1–26.

Peterson, E. R., Rubie-Davies, C., Osborne, D., & Sibley, C. (2016). Teachers' explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap. *Learning and Instruction*, *42*, 123–140.

Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, *34*(1), 1–20.

Rose, M. (2013). Being careful about character. *Phi Delta Kappan*, 95 (2), 44–46.

Schlund, J., Jagers, R., & Schlinger, M. (2020) *Emerging insights on advancing social and emotional learning (SEL) as a lever for equity and excellence.* CASEL. https://bit.ly/CASELEquityInsights

Search Institute. (1997). https://www.search-institute.org/our-research/development-assets/developmental-assets-framework/

Seider, S. (2012). *Character compass: How powerful school culture can point students toward success*. Harvard Education Press.

Seider, S., & Graves, D. (2020, January 9). *Making SEL Culturally Competent: Teaching students about oppression and critical consciousness is a vital component of social-emotional learning*. Edutopia. https://www.edutopia.org/article/making-sel-culturally-competent

Seider, S., Tamerat, J., Clark, S., & Soutter, M. (2017). Investigating adolescents' critical consciousness development through a character framework. *Journal of Youth and Adolescence*, *46*(6), 1162–1178.

Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, *37*(1), 21–34.

Shields, D. L. (2011). Character as the aim of education. *Phi Delta Kappan, 92*(8), 48.

Simmons, D. (2019). Why we can't afford whitewashed social-emotional learning. *ASCD Education Update, 61*(4).

Soutter, M. (2020). Measuring joy: A social justice issue. *Phi Delta Kappan, 101*(8), 25–30.

Tyton Partners. (2020, July). *Finding your place: The current state of K–12 social emotional learning.* https://tytonpartners.com/wp-content/uploads/2020/07/Finding-Your-Place_The-Current-State-of-K-12-Social-Emotional-Learning.pdf

Weikart Center for Youth Program Quality. (n.d.) *Youth program quality assessment and school-age program quality assessment.* http://cypq.org/assessment

Yeager, D., Bryk, A., Muhich, J., Hausman, H., & Morales, L. (2013). *Practical measurement.* Carnegie Foundation for the Advancement of Teaching.

Yoder, N. (2014). *Teaching the whole child: Instructional practices that support social-emotional learning in three teacher evaluation frameworks* (Rev. ed.). American Institutes of Research. https://files.eric.ed.gov/fulltext/ED581718.pdf

Zaino, K. (2016). Always on my grind: Passion and perseverance among low-income youth. *The Journal of School & Society, 3*(2), 5–10.