OJED
OPEN JOURNALS IN EDUCATION
Volume 10, Issue 1 (2025), pp. 13-26
Journal of School Administration Research and Development
ISSN: 2470-8496 Print/ ISSN: 2470-850X Online
ojed.org/JSARD

RIGOR Walk: Development and Initial Validation of a Framework to Support Rigorous Learning Environments

James Marshall
Douglas Fisher
Nancy Frey
San Diego State University, CA, USA

ABSTRACT

The term *rigor* in education often evokes resistance due to its inconsistent definitions and widespread misconceptions. This study introduces and validates the RIGOR Walk framework, a research- and practitioner-informed tool designed to define, observe, and enhance rigorous learning environments across classrooms. The framework is grounded in five core components—*Relationships, Instruction, Goals, Organization,* and *Relevance*—each supported by observable indicators aimed at promoting deep thinking, academic risk-taking, and meaningful engagement. Following iterative rounds of expert feedback, the framework was refined and validated in three phases: (1) establishing face validity through crowdsourced expert critique, (2) analyzing internal consistency and construct validity via 84 classroom observations, and (3) examining concurrent validity through correlations with student reading growth data. Results indicate high internal consistency (Cronbach's $\alpha \geq .81$) and moderate, statistically significant correlations between framework implementation and the percentage of students demonstrating one year or more of reading growth ($r = .42–.68$, $p < .001$). These findings suggest the RIGOR Walk framework is a reliable and valid tool for identifying and fostering rigorous, equitable, and high-impact instructional practices. Implications for broader content-area applications and longitudinal consistency of implementation are discussed.

***Keywords*:** rigor, teaching and learning, instructional frameworks

In education, *rigor* refers to the level of cognitive challenge and academic demand placed on students in their learning experiences (Hess, 2023). It involves teaching, learning, and assessment processes that encourage students to understand deeply, think critically, and apply knowledge in complex, novel, and meaningful ways. As education leaders, we are charged with creating learning environments that yield predictable results. Rigorous learning environments hold significant promise for contributing to such outcomes. Yet, the term suffers from the lack of a shared definition coupled with a range of misconceptions as to exactly what rigor looks like in practice. Addressing this shortcoming is critical to instructional leaders effectively embracing, facilitating, and overseeing rigorous learning environments and classroom educators creating rigorous learning experiences for their students. It is this type of instructional leadership that can lead to high-performing schools (Sanchez & Watson, 2021).

Rigor in education is characterized by instruction that pushes students to explore complex ideas and solve problems that require more than just memorization or basic understanding. It's about promoting higher-order thinking skills such as analysis, synthesis, and evaluation. For instance, Bloom's Taxonomy—a framework developed by educational psychologist Benjamin Bloom and others—categorizes different levels of cognitive skills that reflect increasing complexity and depth, suggesting that higher levels of this taxonomy reflect more rigorous educational demands (Bloom et al., 1956). Webb's Depth of Knowledge (2002) also frames our understanding of cognitive depth with four levels that range from *recall* and *reproduction* through *strategic thinking* to *extended thinking.*

Rigor also involves applying knowledge in different contexts and integrating various learning experiences to develop skills and insights. This can be seen in project-based learning or solving authentic problems where students apply their knowledge and skills across various disciplines to find solutions. This approach serves to reinforce the material learned while also preparing students for real-life challenges by making learning relevant and engaging (Wiggins & McTighe, 2005).

Contrary to common assumptions that have resulted in the term being passionately disliked by scores of educators, rigor is *not* about making learning harder or more difficult. Rather, it is about setting high expectations while also providing the necessary support to help students meet these challenges. Educational theorists like Lev Vygotsky have emphasized the importance of the Zone of Proximal Development (ZPD), which is the difference between what learners can do without help and what they can achieve with guidance and encouragement from a skilled partner (Vygotsky, 1978). Effective rigorous education practices scaffold learning experiences to stretch the student's capabilities within their ZPD.

Common Misconceptions about Rigor

Misconceptions about rigor in education often stem from misunderstandings about what true rigor involves. Here are some common misconceptions:

1. **Rigor means more homework and harder tests**: A prevalent misconception is that increasing the volume of homework or the difficulty of tests automatically makes an educational experience more rigorous. However, true rigor isn't about sheer quantity of work or the level of difficulty alone, but rather the depth of understanding and the cognitive challenge involved. Rigor should push students to think deeply rather than simply complete more assignments or face more difficult questions without context or purpose (Wiggins & McTighe, 2005).
2. **Rigor is only for "gifted" students**: Another common misunderstanding is that rigor is only appropriate for advanced or gifted students. This misconception can lead to lower expectations for the majority of students. In reality, all students benefit from rigorous educational experiences. Rigorous learning involves appropriately challenging all students, regardless of their starting level, and supporting them to reach higher levels of understanding and skill (Tomlinson, 2014).
3. **Rigor is just about academic content**: While rigor involves challenging academic content, it's also about the application of skills in new and complex ways. Some might think of rigor as limited to traditional academic subjects like math and science. However, true rigor incorporates critical thinking, problem-solving, creativity, and application across all subjects, including the arts and humanities (Wiggins & McTighe, 2005).
4. **Rigor means traditional and strict teaching methods**: There's a belief that rigorous education requires a traditional, lecture-based approach where the teacher is the sole authority and students are passive recipients of knowledge. This view neglects the effectiveness of interactive and student-centered teaching methods in promoting deep learning. Students engage more deeply when they are active participants in their learning process, utilizing discussions, group work, and hands-on projects, which can all be rigorous if well-designed (Metzger & Langley, 2020; Radef, 2021
5. **Rigor eliminates creativity and enjoyment**: Some think that a rigorous curriculum is dry, joyless, and void of creative expression, focusing solely on stern standards and continuous assessment. However, true rigor should engage students' interests and passions, integrating creativity and enjoyment with challenging content to motivate and enhance learning experiences (Csikszentmihalyi, 1990).

In his book *Flow,* Csikszentmihalyi describes a state of total openness to learning and performance he terms "flow," which involves complete absorption, focus, and enjoyment in an activity. Rigorous learning experiences, when well-designed and supported, can result in states of flow where there is optimal balance between the challenge of the task and the skill level of the individual performing it. Being in a state of flow is intrinsically rewarding, making the activity itself enjoyable and satisfying (Csikszentmihalyi, 1990).

Understanding rigor correctly is foundational to educators designing and implementing educational experiences that truly enhance learning. This is especially important for effective educational leaders, who Marshall and Fisher (2019) suggest come to the task with a *learnership* thinking, rather than *leadership* thinking, perspective: "A learnership thinker first creates the conditions and opportunities for all adults and students to learn and perform at ambitious, academic levels to achieve in school and life" (p. 77). When rigorous learning environments result, they challenge students constructively and effectively, thus preparing them for the future academic and life challenges they will undoubtedly face.

GENESIS OF THE RIGOR WALK FRAMEWORK

We set out to disrupt the long-standing stigma attached to educational uses of the term "rigor." Our intent was to develop a framework that defined elements of rigorous learning environments and offer indicators by which educators could acknowledge their current rigor-related practices while building upon them to expand their rigor repertoire. Based on the existing research into the positive impact of learning walks, including affirming learning walks (Ross, et al., 2023), we envisioned a rigor-focused set of criteria that could be used to observe and affirm practice where it already exists while encouraging the continued expansion, development, and evolution of rigor-aligned best practice. Our initial review of the literature around effective, rigorous instruction was synthesized into an initial framework comprised of five components on which rigorous learning environments rely: (1) Relationships; (2) Instruction; (3) Goals; (4) Organization; and (5) Relevance. The eventual framework that emerged also included indicators in support of each component that instantiate the application of the framework to practice. The following sections share key research and evidence-based practice elements that contributed to the conceptualization of the five components and the initial set of indicators we tested with education experts and practitioners as the next step of the RIGOR Walk framework's development.

Relationships

Positive relationships between teachers and students, and among students themselves, are crucial for effective learning. Child psychiatrist James Comer highlighted the importance of significant relationships in the educational process, noting that no substantial learning can occur without them (Comer, 1995). Such relationships foster a supportive and trusting atmosphere where students feel safe to engage, inquire, and learn from their errors. This environment not only boosts student engagement and motivation but also nurtures a sense of belonging and community, thereby enhancing both academic and social-emotional development.

The practice of using students' names positively is a key component of building these relationships. Properly recognizing and pronouncing students' names affirms their identity and importance, which can lead to increased engagement and decreased behavioral issues. Studies such as those by Cook et al. (2018) and O'Brien et al. (2014) have shown that greeting students by name at the classroom door can increase engagement significantly and reduce problematic behaviors, thereby contributing to a more inclusive and supportive learning atmosphere.

Furthermore, the physical proximity of teachers to their students during class sessions, referred to simply as "proximity," plays a significant role in enhancing engagement and ensuring effective learning. This strategy involves teachers moving around the classroom, staying close to students to better observe and respond to their needs, particularly those who might disengage. Historically endorsed by educators like Good and Brophy (1987) and based on studies by Etscheidt et al. (1984), this approach helps in managing the classroom dynamics and maintaining high levels of student involvement. Proximity, coupled with the respect and productive interactions among peers, sets a foundation for a safe, engaging, and rigorous learning environment where academic risk-taking is encouraged and valued.

Instruction

The instructional techniques adopted by teachers play a crucial role in shaping the learning outcomes of students. The relationship between teaching decisions and student learning is fundamental, emphasizing the importance of not just focusing on the act of teaching but also on its impact on learning. Some instructional moves are designed to enhance access to rigorous learning experiences, while others might inadvertently reduce the learning intensity by overburdening the teacher and diluting the students' engagement in meaningful tasks. Notably, effective instruction involves selecting from various evidence-based strategies that align with improving student learning outcomes.

A vital aspect of responsive teaching involves using evidence of student learning to inform instructional practices. The concept of formative assessment, introduced by scholars like Cronbach (1963) and Scriven (1967), underlines the importance of using assessment data to refine teaching strategies continuously. Teachers who integrate real-time feedback about student understanding into their lessons can adjust their teaching approaches to better meet the students' needs, either by accelerating the lesson pace or by providing additional clarifications to ensure comprehension. This dynamic adjustment ensures that instruction remains challenging and supportive, maintaining high expectations for all students.

Additionally, the use of scaffolding in instruction is critical to support learning while promoting productive struggle. Originating from the educational adaptations of scaffolding by Wood et al. in 1976, this strategy involves providing appropriate levels of support to students as they develop new skills and strategies. Effective scaffolding ensures that students are challenged within their zone of proximal development, facilitating a balance between too much and too little help. This

approach not only maintains the rigor of the learning experience but also empowers students to take increased responsibility for their learning, fostering independence and confidence in their abilities to meet academic expectations.

Goals

Learning goals serve as essential components of both teaching and learning processes, providing a clear direction for instructional activities and helping teachers ensure that their efforts align with desired educational outcomes. They function as a roadmap for teachers, guiding lesson planning and the delivery of educational content to meet specific objectives. For students, clear learning goals offer a concrete understanding of the expectations set for them, which enhances motivation and engagement by demonstrating the relevance and purpose of their educational activities. Furthermore, these goals are integral to students' ability to conduct self-assessments and seek feedback, allowing them to monitor their progress and identify areas needing improvement, thereby fostering a sense of ownership over their learning journey.

The concept of aligning learning goals with grade-level expectations is crucial for maintaining educational rigor and ensuring that students are meeting appropriate benchmarks for their age and subject area. These goals, articulated through various terms like *learning intentions* or *objectives*, must be accessible and well-communicated to students to keep them focused throughout their educational activities. The historical context provided by Tyler (1949) and Bloom et al. (1956) underscores the longstanding importance of clear, well-structured learning goals in education, supporting the idea that when students understand what is expected of them, their likelihood of successful learning increases significantly.

Moreover, the implementation of success criteria and the practice of regular self-assessment by students are pivotal in translating learning goals into tangible outcomes. Success criteria help students visualize what successful attainment of learning goals looks like, thereby enabling them to assess their progress and adjust their learning strategies accordingly. This practice is supported by educational theories on metacognition and self-regulation, suggesting that students who actively engage in self-assessment are better able to take control of their learning and improve their educational outcomes. Overall, the integration of clear learning goals, success criteria, and self-assessment practices form a foundational framework that supports rigorous, effective, and student-centered education.

Organization

Organization within the classroom setting is crucial for effective learning, as it directly influences students' ability to engage with and benefit from educational activities. A well-organized classroom environment provides students with predictable structures and routines, which can significantly enhance their learning experience by reducing distractions and confusion. Ensuring that students can easily access necessary materials and accommodations further supports this organized learning environment. Moreover, creating a physical space that is accessible to all students, regardless of their abilities, is essential for inclusivity and equity in education, as mandated by legal frameworks like the ADA and Section 504 (U.S. Department of Education, 2020). Embracing principles from Universal Design for Learning (UDL) helps to cater to diverse learning needs by offering multiple means of content representation, engagement, and expression (Meyer et al., 2014).

The physical setup of the classroom also plays a significant role in facilitating learning. An environment that is both content-rich and frequently updated can stimulate students' engagement and curiosity. According to Malaguzzi (1984), the physical environment acts as a "third teacher," alongside adults and peers, influencing learning through its design and the resources it offers. This approach underlines the importance of surrounding students with a variety of learning materials, such as books, visual aids, and interactive technologies, which not only support literacy development but also encourage a deeper interaction with learning content. Ensuring that these materials are current and relevant further enhances their effectiveness as educational tools.

Flexible grouping within classroom organization is another effective strategy for addressing individual learning needs. By forming groups based on students' specific educational requirements rather than standardized assessments, teachers can provide targeted instruction that is more likely to resonate with and benefit each student (Colón et al., 2022). Such flexibility in grouping allows for the adaptation of teaching methods and materials to suit different learning styles and paces, thereby optimizing educational outcomes. Additionally, effective management of student behavior through thoughtful procedures and interventions ensures that the learning environment remains conducive to education, focusing on constructive behavior support rather than punitive measures (Jung & Smith, 2018). This comprehensive approach to organization supports a dynamic and responsive learning environment where all students can thrive.

Relevance

Relevance in education is a crucial component that drives student engagement and learning efficacy. When students perceive their learning as relevant, they are more likely to engage actively and regulate their behavior positively (Stuart, 2023). This concept of relevance extends beyond merely informing students about the future utility of their education. It encompasses the creation of learning experiences that are responsive to students' backgrounds and lived experiences, making the learning process personally significant. Tasks within the educational setting must therefore be meaningful, integrating real-life contexts that resonate with the students' own experiences. This approach aligns with theories of situated learning and cognition, which suggest that knowledge and learning are deeply embedded in the contextual realities of the learner (Brown et al., 1989; Lave & Wegner, 1991). Consequently, when learning tasks are designed to connect with students' lives both inside and outside the classroom, the relevance is heightened, and so is students' motivation to learn (Keller, 2010).

Moreover, relevance is further emphasized when students can articulate the value of their learning experiences. This ability not only supports deeper content understanding but also fosters metacognitive skills, allowing students to reflect on and optimize their learning strategies (Flavell, 1979). According to constructivist theories by Piaget (1952) and Vygotsky (1978), knowledge is constructed through personal experiences and reflections, which underscores the importance of students expressing the value and process of their learning. Such expressions not only aid in formative assessment by providing insights into students' thought processes but also enhance learning autonomy, contributing to more tailored and effective educational approaches (Zimmerman, 2002).

Lastly, the incorporation of students' lived experiences and cultural backgrounds into learning experiences contributes to teaching and learning that is culturally relevant and inclusive. This method fosters a sense of belonging and respect within the learning environment, encouraging students to engage with content more meaningfully (Gay, 2000; Villegas & Lucas, 2002). It also helps students to critically analyze their own perspectives and appreciate diverse viewpoints, enhancing their overall educational experience and preparing them for a more interconnected world (Ladson-Billings, 1995). By linking learning materials to the unique identities and backgrounds of students, educators not only promote inclusivity but also drive deeper engagement and understanding, making learning a more comprehensive and transformative experience.

VALIDATING THE RIGOR WALK FRAMEWORK

Our earliest conceptualization of the framework was organized around the rigor components and between six and ten indicators under each of the five components. We intentionally over-allocated indicators as an early step in the process that established face validity.

Validation Phase 1

With an initial version of the RIGOR Walk framework in hand, Validation Phase 1 employed crowdsourcing as the means for pursuing face validity. As the developers and using our own social networks of experts and practitioners (collectively termed “experts”), we solicited individuals who were willing to review and critique the earliest versions of the framework. Using an iterative review approach, we conducted four rounds of expert review. At the conclusion of each round, the author team received and analyzed the expert feedback, perspective, and revision suggestions. Feedback was typically constructive, often critical, and always thought provoking. Examples of typical comments include the following:

- “When students seek feedback, it’s also important that they act based on that feedback. We want them to use the feedback. Maybe revise accordingly?”
- “I think that students’ names should be used in productive ways not just positive ways, but I think that the whole relationship component is critical. I am pleased to see risk-taking framed as a positive.”
- “The whole relevance aspect is over-looked. Thank you for including that. Would you consider adding unique to identities and having a clearer statement about the way that prior knowledge could be used?”

We applied the same deliberation process following each round. We debated modifications based on feedback, reconciled potential revisions with our existing review of the literature for alignment, and then made modifications to the framework that were supported by at least two of the three authors as well as validated by the literature. Another round of crowdsourcing and expert feedback followed each of the first three validation rounds. While some experts persisted to provide input across multiple review rounds, others limited their contributions to a single round. Table 1 summarizes the quantity of expert feedback and resulting evolution of the framework during the face validity phase of the framework validation, including the reduction statistics by framework component.

Table 1

Face Validity Crowdsourcing Effort Statistics across Four Feedback Rounds

Feedback Round	Number of Experts Consulted	Number of Comments Received*	Number of Revisions Made
1	8	58	43
2	10	55	24
3	16	41	17
4	12	27	8

Component	Number of Indicators at Start	Number of Indicators at Conclusion
Relationships	7	5
Instruction	9	5
Goals	6	5
Organization	7	5
Relevance	10	5

Note. *Comment count includes only constructive critiques; count omits expert comments that indicated the framework, or a component of the framework, was successful as presented.

It is important to note that the range of comments, as well as the number of comments, decreased with each of the latter rounds of expert review. This was an important evolution of the validation effort that we interpreted to signal an increasing consensus around the framework's components and indicators. Following the fourth round, we finalized the framework and moved forward to the next phases of validation that involved assessing both reliability (Phase 2) and concurrent validity (Phase 3). The resulting RIGOR Walk framework, which was tested in future validation phases, is presented in Figure 1.

Figure 1

The RIGOR Walk Framework

R	I	G	O	R
Relationships	**Instruction**	**Goals**	**Organization**	**Relevance**
Students' names are used in positive and productive ways.	Evidence of student learning is used to inform instruction.	Learning goals are aligned with grade-level expectations.	The physical environment is accessible for all students.	The learning process incorporates meaningful tasks that embed learning inside and outside the classroom.
Proximity is used to foster connections with students and ensure their learning.	Students interact with peers in meaningful discussions using academic language to complete tasks.	The level of knowledge expected of the learning goal aligns with the standard.	The physical environment is rich and recent.	Students describe the value of what they are learning and how they are learning it.
Students' interactions with peers are respectful and productive.	Scaffolds are strategically used to support learning, invite productive struggle, and ensure productive success.	Students can describe or demonstrate what successful learning looks like.	Grouping patterns are used flexibly to promote learning.	Students' lived experiences, as well as those from backgrounds different from their own, are incorporated into learning experiences, making lessons culturally relevant and inclusive.
Academic risk-taking is encouraged and celebrated.	Lessons include input based on student learning needs.	Students regularly self-assess their learning and revise their actions based on the results.	Student behavior is proactively managed, monitored, and addressed through productive procedures and interventions.	Artifacts and materials reflect the unique identities and interests of students.
Student ideas are valued and explored as bridges to learning.	Students practice and apply what they have learned to familiar and new situations.	Students seek feedback, are provided with actionable ideas, and follow through with next steps.	The flow and pace of the lesson is aligned with the learning goals.	Learning activates students' prior knowledge and experiences and fosters connections to new or more complex content.

Note.

In addition to defining language for each of the 25 indicators, the authors also developed observation criteria for each indicator using four-point scales. Table 2 provides examples of these four-point scales for 3 of the 25 indicators included in the framework. These scales were used in the observations conducted in Phase 2.

One important contextual factor to note is the breadth of the RIGOR Walk framework components. The indicators, at times, necessarily contained multiple, yet related, constructs. As intended, the initial framework was exacted by expert review, reflection, and revision in Phase 1 to influence the reliability of the resulting scales. This was also part of the effort to establish face validity and, as such, Phase 1 reviewers reached agreement on the indicators in terms of their content and observability based on the final language of each.

Table 2

Example Observation Scales

Observation Scale Points	Students' names are used in positive and productive ways.	Students can describe or demonstrate what successful learning looks like.	Artifacts and materials reflect the unique identities and interests of students.
1	Students are not greeted when they arrive. Student names are not used during the lesson. Names are used to address problematic behavior.	There are no success criteria.	Artifacts and materials hold little or no connection to the identities and interests of students.
2	Students are greeted at the beginning of class, but not by name. A small group of students have their names used during the lesson.	Few students can describe the success criteria or demonstrate what successful learning looks like.	Some artifacts or materials appear to connect to an identity or interest of a student. The connections are typically cursory or not fully developed.
3	All students are greeted by name upon arrival. Many students hear their names in positive ways throughout the lesson. Students generally report that the teacher knows their names.	Most students are able to describe the success criteria, and some can explain how these align with their demonstrations of success.	Most artifacts and materials hold some connection to the identity or interests of the students who use them. Connections vary in depth to include examples of both cursory and deep connections.
4	All students are greeted by name upon arrival and most of their names are used throughout the lesson. The reactions of students when they hear their names suggests the teacher has a strong relationship with students.	Most students can describe success criteria, can explain how these align with their demonstrations of success, and can express how they are using them to determine the next steps in their learning.	Artifacts and materials clearly reflect the identity or interests of the students who use them or produce them. Connections are typically deep and meaningful, reflecting well-developed and critical thinking being employed.

Validation Phases 2 and 3

The next validation phases involved the application of the framework in support of classroom observation. Here, we analyzed 84 rigor-focused observation records collected in 84 unique classrooms. These observations included teachers from 14 schools in 10 school districts from 4 different states. The schools were suburban and rural, with a range of students qualifying for free lunch from 37–100% and the range of students qualifying for special education services from 14–18%. In terms of teacher demographics, 62 (74%) were female, 22 were Latinx (26%), 12 were black (14%), 9 were Asian-Pacific Islander (11%), and 41 (49%) were white.

The data analysis protocol was reviewed by the author's Institutional Review Board and determined not to be subject to review. This is because the observation data analyzed was deidentified prior to being shared with the authors, ensuring there was no way to connect the records back to the 84 individuals who were observed (author institution IRB #IRB-25-0041).

We purposely included classrooms across the grade spectrum, with a minimum of 3 classrooms and a maximum of 11 observed at each grade level from grades K–11. The secondary classrooms were all English language arts classrooms. Using the RIGOR Walk framework as a scaffold, an observer conducted an observation of instruction lasting from 25–40 minutes in each classroom. The observer tabulated the following data for each observation:

- *Grade Level*: Grade level of students observed
- *Classroom Size*: Number of students in classroom observed
- *Indicator Rating*: For each of the 25 indicators (five per component), the observer assigned a rating using the specific, four-point scale-defined criteria developed by the authors and specific to each indicator (see Table 2 for example scales).
- *Classroom Reading Growth*: Percentage of students in each observed classroom who made the equivalent of one year of growth in reading/literacy performance at the time of observation.

Further description of the Classroom Reading Growth diagnostic is merited. Existing iReady reading diagnostic data (Curriculum Associates, 2024) was collected at the beginning, and throughout the school year in classrooms in which rigor observations took place. The iReady reading diagnostic assesses in the areas of foundational skills—phonological awareness, phonics, high frequency words; vocabulary; comprehension of informational text; and comprehension of literary text. The reading performance data in this validation effort represent the aggregated student performance within a classroom and were provided to researchers following each observation to mitigate chances of rater bias. The data is limited to composite class-level performance by teacher, rather than student-by-student performance data.

For purposes of consistency and research scope, each of the 84 observations was conducted by the same individual who was familiar with the original RIGOR Walk framework. We used a single rater for this initial study, but we acknowledge and expect this framework may be used by multiple individuals who undergo training to ensure valid observational interpretations and implementation ratings. A follow-up study employing multiple raters is currently in process.

Tabulated results were entered into a spreadsheet (with data intentionally deidentified to make connections between the records and human subjects impossible), which was then provided to the authors for analysis using IBM SPSS Statistics (Version 29). The intentional diversity of observed grade levels (grades K–11) was based on the rationale that the characteristics of a rigorous learning environment should be consistent across ages; it is the specific content and scaffolding implemented by the classroom teacher that results in the age-appropriate application of the rigor components. However, attending to relationships, instruction, goals, organization, and relevance remain constant regardless of age.

The initial dataset was used to compute scales for the five RIGOR Walk components. Each component (relationships, instruction, goals, organization, relevance) was comprised of 5 indicators rated on a scale of 1–4. Thus, the total rating for each component ranged from 5–20, and a total rigor rating across all 5 components ranged from 25–100.

Validation Phase 2: Assessing Internal Consistency Reliability and Construct Validity

In Phase 2 of the validation effort, we conducted analysis to determine internal consistency reliability and construct validity for each of the five rigor components using the scales described in Figure 1.

Internal Consistency Reliability

Phase 2 analysis involved investigating reliability using Cronbach's alpha for each of the five scales developed to represent the five RIGOR Walk components. Table 3 presents coefficient alpha figures for each element. Overall, results were strong and suggest moderate to strong reliability (internal consistency) for each of the rigor scales.

Construct Validity

Factor analyses were conducted to test construct validity for each of the five rigor components, augmenting Cronbach's alpha analysis. Five separate factor analyses were conducted—one for each rigor component—to assess whether all factors within each component loaded onto a single factor. Each of the five rigor components demonstrated unidimensionality, with a single factor accounting for 60%–85% of the variance and item loadings ranging from .41–.97 Specifically, *Relationships* demonstrated 85.81% of variance with factor loadings from .87–.97; *Instruction* demonstrated 72.11% of variance with factor loadings from .51–.82; *Goals* demonstrated 71.67% of variance with factor loadings from .79–.88; *Organization* demonstrated 60.43% of variance with factor loadings from .76–.95; and *Relevance* demonstrated 61.94% of variance with factor loadings from .41–.93.

Table 3

RIGOR Walk Scale Reliability (Cronbach's α) by Component

Component/Indicators	Cronbach's α	Cronbach's α if Item Deleted
1. Relationships	**.95**	
1.1 Students' names are used in positive and productive ways.		.95
1.2 Proximity is used to foster connections with students and ensure their learning.		.93
1.3 Students' interactions with peers are respectful and productive.		.93
1.4 Academic risk-taking is encouraged and celebrated.		.96
1.5 Student ideas are valued and explored as bridges to learning.		.93
2. Instruction	**.90**	
2.1 Evidence of student learning is used to inform instruction.		.88
2.2 Students interact with peers in meaningful discussions using academic language to complete tasks.		.90
2.3 Scaffolds are strategically used to support learning, invite productive struggle, and ensure productive success.		.88
2.4 Lessons include input based on student learning needs.		.86
2.5 Students practice and apply what they have learned to familiar and new situations.		.86
3. Goals	**.90**	
3.1 Learning goals are aligned with grade-level expectations.		.89
3.2 The level of knowledge expected of the learning goal aligns with the standard.		.87
3.3 Students can describe or demonstrate what successful learning looks like.		.89
3.4 Students regularly self-assess their learning and revise their actions based on the results.		.87
3.5 Students seek feedback, are provided with actionable ideas, and follow through with next steps.		.86
4. Organization	**.81**	
4.1 The physical environment is accessible for all students.		.88
4.2 The physical environment is rich and recent.		.72
4.3 Grouping patterns are used flexibly to promote learning.		.73
4.4 Student behavior is proactively managed, monitored, and addressed through productive procedures and interventions.		.76
4.5 The flow and pace of the lesson is aligned with the learning goals.		.77

Component/Indicators	Cronbach's α	Cronbach's α if Item Deleted
5. Relevance	**.82**	
5.1 The learning process incorporates meaningful tasks that embed learning inside and outside the classroom.		.74
5.2 Students describe the value of what they are learning and how they are learning it.		.75
5.3 Students' lived experiences, as well as those from backgrounds different from their own, are incorporated into learning experiences, making lessons culturally relevant and inclusive.		.74
5.4 Artifacts and materials reflect the unique identities and interests of students.		.88
5.5 Learning activates students' prior knowledge and experiences and fosters connections to new or more complex content.		.80

Validation Phase 3: Concurrent Validity

While the RIGOR Walk framework (and rigorous learning environments more generally) is content agnostic, we acknowledge that the ultimate measure of the framework's efficacy lies in student learning, as indicated by an objective assessment of content knowledge. Therefore, in the final phase of this initial validation effort, we calculated the concurrent validity of the framework by investigating the correlation between each of the five RIGOR Walk framework components and the percentage of the observed teacher's students who were determined to have made at the equivalent of one year's reading/literacy growth (or greater). The underlying hypothesis was framed around a direct relationship between the presence of rigor components and reading/literacy performance. In other words, higher levels of rigor implementation led to greater percentages of students demonstrating one year's growth.

We analyzed the observation data using Pearson's *r* procedure. Table 4 presents a summary of the resulting correlation coefficients and significance levels for each of the five RIGOR Walk components and the overall RIGOR Walk framework (represented as the total number of rating points assigned to the observation, which would range from 25–100). Findings confirmed the underlying hypothesis, with direct and moderate correlations across the RIGOR Walk framework, which suggests that the greater the rigor rating, the greater the percentage of students demonstrating the equivalent of one year's reading growth.

Table 4

RIGOR Walk Components Correlated to Percentage of Students Achieving ≥ One Year Reading Growth

RIGOR Component	df	Correlation (Pearson's *r*): Students (%) w/ ≥ One Year of Reading Growth	*p*
Relationships	82	.64	<.001
Instruction	82	.66	<.001
Goals	82	.61	<.001
Organization	82	.47	<.001
Relevance	82	.42	<.001
Overall RIGOR Framework	82	.68	<.001

Finally, we chose to split the sample into two groups by rigor level, named high and low rigor. Participants in the high rigor group had ratings of three or four (using the four-point rating scale designed for each indicator) for each indicator across four or all five rigor components (n = 60); low rigor group members comprised the remaining group (n = 24) with scores of one or two for each indicator across four or all five rigor components. With the percentage of students achieving

the equivalent of one year's growth as the dependent variable, we compared low (M = 30.79, SD = 13.52) and high (M = 56.35, SD = 21.40) groups using a t-test for independent groups. The difference between groups proved reliable and statistically significant ($t(82) = -5.42$, $p < .001$).

DISCUSSION

The term rigor, when used in education, continues to be an inconsistently defined construct and one to which many educators object, given the lack of definition and the many misconceptions about the components of a rigorous learning environment. We developed the RIGOR Walk framework in support of a shared definition that includes tangible and specific indicators that leaders and educators can use to identify existing rigor-aligned practice while also working to increase rigor-related instructional elements in their schools and classrooms.

RIGOR Walk Framework: Validity

Our initial effort to validate the RIGOR Walk framework benefitted from the perspectives of 46 education expert practitioners. We exacted a framework for latter phase pilot testing through iterative cycles of review and revision. These cycles converged until minimal revisions were made and comments evolved from being considerable (e.g., removal, addition, or significant language revision) to limited (typically, word choice) in the fourth and final round. Throughout this face validation process, we balanced expert insight with the established and reviewed literature base to arrive at a draft framework that was informed by both research and practice.

An initial investigation of the concurrent validity of the RIGOR Walk framework was conducted using existing teacher-level reading performance scores. While the framework is designed as content agnostic, we employed aggregated reading performance data to determine the level of correlation, if any, between the observed educator's rigor ratings and the percentage of their students who had achieved one year or greater reading growth.

Findings suggest a moderate level of concurrent validity with Pearson r correlations ranging from .42–.66 for the 5 individual components and .68 for the full framework, which is based on a total RIGOR Walk framework rating comprised of the sum of all 25 indicators. All correlations proved statistically significant ($p \leq .001$). Of the five components, *Relationships* (.64) and *Instruction* (.66) returned the highest levels of correlation, with *Relevance* possessing the lowest (.42). Additionally, the t-test for independent groups analysis revealed significant differences in the percentage of students demonstrating one year or greater reading growth, when groups based on low- and high-rigor scores were compared.

We concluded that the concurrent validity of the RIGOR Walk framework, specific to the percentage of students making one year's reading growth, was moderate and acceptable. However, we also acknowledge the limited scope of this investigation. With a framework that was created to be broad in its subject matter application, further validation efforts must pursue (1) multiple dimensions of reading/literacy performance and (2) content areas beyond reading/literacy. By examining the correlation between RIGOR Walk ratings and student performance across multiple subject areas, we will more broadly understand the concurrent validity of the framework used universally in support of creating and optimizing rigorous learning environments. We have acknowledged future opportunities for validation using student-level data.

RIGOR Walk Framework: Reliability

The internal consistency analysis of each of the five RIGOR Walk scales (one per component) returned Cronbach alpha levels that we concluded as very good. We applied the guidance of Cronbach (1951) who suggested a value of .7 as the minimally acceptable level in social sciences research, as well as Ursachi et al. (2015), among others, to make these determinations. This includes the specific classification guidelines that suggest the "general accepted rule is that an alpha of 0.6–0.7 indicates an acceptable level of reliability, and 0.8 or greater a very good level" (Ursachi et al., 2015, p. 681).

Cronbach's alpha values for each of the five RIGOR Walk components exceeded .8, which indicates a very good level of internal consistency among the five involved indicators which comprise each of the five scales: Relationships = .95; Instruction = .90; Goals = 90; Organization = .81; and Relevance = .82. These findings did include two components with potential opportunities to raise scale reliability: Organization and Relevance.

In the case of Organization, analysis suggested that the Cronbach's alpha coefficient would rise by .07 with the elimination of the indicator "The physical environment is accessible for all students." Interestingly, no such gain was seen for the other item related to physical environment: "The physical environment is rich and recent." In conceptualizing the Framework, we acknowledge the complexity of the five components and the inherent challenges in reducing each to five indicators. In the case of Organization, we concluded that the initial coefficient value of .81 was of sufficient strength to

retain all five items, including the accessibility indicator. We observed that this item may simply diverge in construct, to a degree, from the other four indicators while still contributing to the consistency of the scale and completeness of the Organization component. This was further supported by the fact that the item had the lowest mean (M = 2.83) relative to the other four items, where the means ranged from 3.01–3.36.

A similar finding was observed with Relevance (coefficient α = .82). In this case, the single item that would raise the coefficient level by .06 was as follows: "Artifacts and materials reflect the unique identities and interests of students." Again, we weighed the limited gain in Cronbach's alpha level against both the necessary reduction involved in summarizing the Relevance component into five indicators, as well as the criticality of identity and interest as they relate to relevant learning experiences. The item returned the greatest variance of those on the Relevance scale with a standard deviation of 1.26 relative to the other four items (*SD* ranged .96–1.16). As with Organization, we decided to accept the .82 alpha level and retain the five existing indicators as tested.

Lastly, a confirmatory factor analysis produced intra-component factor loading and variance figures that further illustrated the contributions of each set of five indicators to their related rigor components.

Overall, we concluded that the RIGOR Walk framework scales for each of the five components reflected high levels of internal consistency as measured by Cronbach's alpha. Two opportunities for limited coefficient improvement were identified through our analysis. Upon consideration and given the already established high degree of internal consistency for each scale, we retained the existing items. Future validation efforts, described below, will allow for continued analysis of these initial decisions.

OPPORTUNITIES FOR FURTHER RESEARCH AND VALIDATION

This study has described the development and initial validation of the RIGOR Walk framework, which was created to define elements of rigorous learning environments. With a goal of reclaiming one of the most disliked words in education through an emergent, shared definition of rigor, we engaged in an iterative, research- and expert-informed development process to arrive at the draft RIGOR Walk framework. We then conducted 84 classroom observations and analyzed the resulting data for internal consistency (reliability) and concurrent validity using student reading growth figures. The findings from this initial study suggest high levels of internal consistency and moderate levels of concurrent validity within the context of student reading growth.

While the preliminary validation findings presented in this article are encouraging, we have identified, and already acknowledged, additional opportunities to strengthen the body of evidence in support of framework validation. Some of these opportunities, such as an analysis of student-level data, have already been noted. Additional opportunities are described below.

Validity by Content Area

As noted, our initial analysis of concurrent validity involved the correlation of RIGOR Walk ratings to the percentage of students in a given classroom that had demonstrated one year or greater of reading growth. The correlation results for both the individual RIGOR Walk components and the overall composite RIGOR Walk score suggested a moderate correlation that can be further investigated in multiple ways. First, we recommend a more nuanced investigation of reading/literacy performance. This could include student-level, rather than aggregated, reading performance metrics that include domain-specific performance data (e.g., vocabulary, reading comprehension, etc.).

Beyond measures related to reading/literacy, future validation of the RIGOR Walk framework should incorporate diverse content areas that include mathematics, science, social sciences, and even physical education. Our premise in creating the RIGOR Walk framework is that rigorous learning environments are realized through strategies that can be applied to any content or learning outcome. Studying the relationship between the framework-defined elements of rigorous learning environments and student performance measures across multiple disciplines will determine the extent to which the framework is successful in this design intent. Additionally, findings of such efforts will inform the applicability of the framework specific to instructional focus and content.

Effects of Consistency over Time

An additional validation effort should focus on consistency of instruction. Another hypothesis we made in the framework's construction was that an educator's practice specific to the RIGOR Walk indicators would persist over time. We posited that these practices, when they occur, integrate into an educator's regular performance. As such, we would

expect some reasonable level of consistency in practice over time. The current validation effort prioritized single observations, occurring with 84 educators across a wide range of grade levels. To complement this initial study, we envision a future validation effort that focuses on a limited number of educators with multiple observations taken over time. The data would then be analyzed to understand both (1) consistency in implementing rigorous learning environment elements over time and (2) the relationship between the presence of those elements and various outcome variables (thus also attending to concurrent validity). Such an investigation would compare data based on the level of consistency observed in implementing the RIGOR Walk components, assuming a dataset with varying levels of consistency to support data-based definition of implementation consistency levels.

CONCLUSION

The lack of a shared definition for rigor, coupled with wide-ranging misconceptions, has resulted in a misunderstood and often disliked term for educational leaders and classroom educators alike. The RIGOR Walk framework was developed to address these concerns while also supporting an increase in the number and quality of rigorous learning environments. This study described an initial validation effort for the RIGOR Walk framework, which was designed to investigate the framework's internal consistency and provide some estimate of concurrent validity.

Results of 84 classroom observations suggest the five-component framework has achieved a high degree of internal consistency, suggesting the reliability of the piloted framework's indicator scales. An additional element of this effort investigated the correlation between the RIGOR Walk framework ratings and student growth in reading performance. A moderate, reliable correlation was observed between each of the five rigor components with the composite rigor rating and the percentage of students in each classroom who made the equivalent of at least one year's reading growth.

When considered together, and in combination with face validity that was iteratively established through multiple rounds of expert review and revision, the RIGOR Walk framework provides an initially validated framework that can be used to envision, identify, and strengthen rigorous learning environments that support student learning.

REFERENCES

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. David McKay Company.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

Colón, G., Zgliczynski, T., & Maheady, L. (2022). Using flexible grouping. In J. McLeskey, L. Maheady, B. Billingsley, M. T. Brownell, T. J. Lewis (Eds.), *High leverage practices for inclusive classrooms* (pp. 265-281). Routledge.

Comer, J. (1995). Lecture given at Education Service Center, Region IV. Houston, TX.

Cook, C. R., Fiat, A., Larson, M., Daikos, C., Slemrod, T., Holland, E. A., Thayer, A. J., & Renshaw, T. (2018). Positive greetings at the door: Evaluation of a low-cost, high-yield proactive classroom management strategy. *Journal of Positive Behavior Interventions*, *20*(3), 149–159.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334. https://doi.org/10.1007/BF02310555

Cronbach, L. J. (1963). Course improvement through evaluation. *Teacher's College Record, 64*(8), 672–683.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.

Curriculum Associates. (2024). i-Ready Diagnostic for Reading. Author.

Etscheidt, S., Stainback, S., & Stainback, W. (1984). The effectiveness of teacher proximity as an initial technique of helping pupils control their behavior. *Behavioral Disorders, 18*, 33-41.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. A new area of cognitive-development inquiry. *American Psychologist, 34*(10), 906–911.

Gay, G. (2000). *Culturally Responsive Teaching: Theory, Research, and Practice*. New York, NY: Teachers College Press.

Good, T. L., & Brophy, J. E. (1987). *Looking in classrooms* (5th ed.). Harper & Row.

Hess, K. (2023). *Rigor by design, not chance: Deeper thinking through actionable instruction and assessment.* ASCD.

Jung, L.A., & Smith, D. (2018). Tear down your behavior chart. *Educational Leadership, 76*(1), 12-18.

Keller, M. (2010). *Motivational design for learning and performance: The ARCS model.* Springer.

Ladson-Billings, G. (1995). Toward a Theory of Culturally Relevant Pedagogy. *American Educational Research Journal, 32*(3), 465-491.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Malaguzzi, L. (1984). *When the eye jumps over the wall: Narratives of the possible*. Regione Emilia Romagna, Comune di Reggio Emilia.

Marshall, J., & Fisher, D. (2018). Making preparation practical: Reducing aspiring administrator time to competence through five types of leaderly thinking. *Journal of School Administration Research and Development, 3*(1), 74–80.

Metzger, K., & Langley, M. (2020). Active learning classroom design and student engagement: An exploratory study. *Journal of Learning Spaces, 9*(1), 1–12.

Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal design for learning: Theory and practice*. CAST.

O'Brien, M.T., Leiman, T., & Duffy, J. (2014). The power of naming: The multifaceted value of learning students' names. *QUT Law Review*, 14(1). https://openresearch-repository.anu.edu.au/bitstream/1885/31373/2/01_O'Brien_The_Power_of_Naming:_The_2014.pdf

Piaget, J. (1952). *The origins of intelligence in children* (M. Cook, Trans.). International Universities Press.

Radef, A. M. (2021). The effectiveness of interactive strategies in teaching and learning. *ScienceDirect*, 67–72.

Ross, D., Lamb, L., & Johnson, J. (2023). Using affirming learning walks to build capacity. *Journal of School Administration Research and Development, 8*(1), 47–54.

Sanchez, J., & Watson, J.M. (2021). Effective instructional leadership practices in high performing elementary schools. *Journal of School Administration Research and Development, 6*(2), 61–70.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–85). Rand McNally.

Stuart, D. (2023). *The will to learn: Cultivating student motivation without losing your own.* Corwin.

Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). Association for Supervision and Curriculum Development.

Tyler, R. (1949). *Basic principles of curriculum and instruction.* University of Chicago Press.

Ursachi, G., Horodnic, I. A., & Zait, A. (2015). How reliable are measurement scales? External factors with indirect influence on reliability estimators. *Procedia Economics and Finance, 20,* 679-686.

U.S. Department of Education. (2020). *Office for Civil Rights: Protecting students with disabilities*. Retrieved from https://www2.ed.gov/about/offices/list/ocr/504faq.html

Villegas, A.M., & Lucas, T. (2002), Preparing culturally responsive teachers: rethinking the curriculum. *Journal of Teacher Education, 53*(1), 20-32.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Webb, N. L. (2002). Depth-of-knowledge levels for four content areas. Language Arts, 28, http://ossucurr.pbworks.com/w/file/fetch/49691156/Norm%20web%20dok%20by%20subject%20area.pdf

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Association for Supervision and Curriculum Development.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89-100.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

JAMES MARSHALL, Ph.D., is Professor and Senior Director of the Doctoral Program in Educational Leadership at San Diego State University in the United States. His major research interests lie in the area of optimizing human and organizational performance through the design and leadership of learning programs that yield predictable results. Email: marshall@sdsu.edu

DOUGLAS FISHER, Ph.D., is Professor and Chair in the Department of Educational Leadership at San Diego State University in the United States. His research interests lie in the areas of reading and literacy, differentiated instruction, school leadership, and curriculum design. Email: dfisher@sdsu.edu

NANCY FREY, Ph.D., is a Professor in the Department of Educational Leadership at San Diego State University in the United States. Her research interests lie in the areas of reading and literacy, special education, and school leadership. Email: nfrey@sdsu.edu

Manuscript submitted: ***February 2, 2025***
Manuscript revised: ***April 14, 2025***
Accepted for publication: ***April 17, 2025***