The Education Trust

Advocates' Guide to Demanding Racially & Culturally Inclusive State Assessments

By Shayna Levitan, Policy Analyst, and Nicholas Munyan-Penney, Assistant Director, P-12 Policy

Policymakers, communities, advocates, students, and families rely on the comparable data from statewide summative assessments to help identify and address educational inequities in academic opportunities afforded to students, particularly for students of color and those from low-income backgrounds. School and district leaders can use assessment results to target state and local resources to the schools with the greatest need and determine how these resources are affecting student achievement. Statewide summative assessments support state and district decision making because they are the only tests that students take that allow for valid comparisons across student groups, schools, and districts within a state.

However, in recent focus groups, students of color and their families cited strong concerns that statewide assessments are biased because they limit students of color from demonstrating what they know and can do.[1] These concerns — alongside other perspectives heard from educators, and district leaders — informed an EdTrust framework and a set of federal policy recommendations for the future of assessments, which are detailed in the report, Future of Assessments: Centering Equity and the Lived Experiences of Students, Families, and Educators. For the report, we developed four equity pillars that identify key criteria for improving federal assessment policy. Importantly, these pillars focus on improving assessment systems while maintaining existing federal assessment requirements.

This report discusses Equity Pillar 2 of that framework, "Encourage Relevant, Inclusive Assessments" and identifies actions advocates can take at the state level. It articulates a vision for assessment design that engages, affirms, and represents the racial and cultural diversity of our nation's students in statewide summative assessments. Specifically, this brief:

- Informs stakeholders of the role that culture plays in learning and assessment processes
- Identifies the ways in which statewide annual assessments center the dominant culture of Whiteness
- Proposes five elements of racially and culturally inclusive large-scale assessment items
- Outlines current developments that are making statewide summative assessments more inclusive
- Highlights promising strategies and policy recommendations for advocates to use in discussions with state leaders

Now Is the Time

While our public education system has long attempted to whitewash student experiences in school, the contentious opposition to honest and authentic representation has become increasingly embroiled. Dangerous, widespread attacks against cultural representation and the teaching of honest history continue. All of this exacerbates consistent racial and ethnic underrepresentation in school curricula. Research shows that students are more engaged when they see people like themselves in school materials,[2] and that increased engagement from representation leads to:

- Sharpened critical thinking skills and increases in standardized test scores[3]
- Higher rates of course completion, graduation, and school attendance[4]
- Improved self-esteem, social and emotional well-being, empathy, and a greater appreciation for cultural differences[5]

Education advocates must continue to mobilize against the so-called "anti-CRT" movement while also pushing to improve representation in statewide assessments. Ensuring that assessments are authentic and diverse in their representation is just as crucial as complex racial and ethnic representation in curricula to make students feel seen. The same can be said about how inclusive, affirming assessments make students feel more engaged,[6] eliminating barriers to students who can demonstrate their full breadth of knowledge and abilities. Ultimately, if students see their identity and interests reflected, then parents, educators, and policymakers can formulate a more accurate picture of how school systems can serve all students, especially students of color. As it stands, statewide annual assessments currently center White culture and norms.

State leaders are essentially customers to assessment development companies, so if they apply consumer pressure, assessment development companies are more likely to make changes to the assessment. Without outside pressure, most states and assessment developers will not initiate improvements on their own. To be sure, updating content takes additional time, capacity, and resources to de-center Whiteness from current assessment design, as well as update the many processes that undergird the educational assessment industry, which include re-evaluating longstanding psychometric and statistical practices. However, with the majority of students in U.S. schools being students of color, this should be a worthwhile and just effort.

Advocates are well-positioned to take on this activator role, focusing their attention on building grassroots support directed at state-level action. The concluding section of this brief lays out specific state policy recommendations for advocates. In addition to creating diverse curricula, state leaders should aim for annual assessments that are engaging, affirming, and authentic for the benefit of their students.

Ultimately, if students see their identity and interests reflected, then parents, educators, and policymakers can formulate a more accurate picture of how school systems can serve all students, especially students of color

Culture: A Fundamental Part of Learning Processes

Today, most statewide assessments avoid questions (known as "items") and passages that allude to cultural topics or themes on the grounds of fairness and colorblindness: If assessments are supposed to be a neutral and accurate representation of what a student knows and can do, then culture-specific content can prejudice results by introducing extraneous factors. As a result, assessment development companies design items that are stripped clean of any specific cultural references.[7] But by doing so, they are reinforcing the dominant culture of Whiteness.[8]

However, culture plays a fundamental role in how students learn. Scientific evidence indicates that learning processes are intrinsically cultural and significantly influenced by contextual factors, social constructs, and personal perspectives. This is because students build knowledge and understanding by incorporating new information into their existing experiences and awareness.[9] Moreover, not only does current practice rely on outdated learning science,[10] it perpetuates the impossible aim of eliminating culture from assessments altogether.

Further, if the goal is to prepare all students to read, write, and meaningfully engage in an increasingly diverse and multicultural society, a test that only assesses whether students can demonstrate knowledge within White norms and parameters severely misses the target. Instead, decision-makers are left with a misaligned understanding of whether schools are doing a sufficient job preparing students for the world they live in and the future we hope to create.

Making assessments more racially and culturally relevant is not an issue of adding culture where none exists. Rather, the aim of racially and culturally inclusive assessments is to diversify the culture that is already present.[11] Molly Faulkner-Bond, a senior research associate at WestEd, sums it up well: "It becomes a question of doing for all children what is currently done for some, whether we mean to or not."[12]

Culture plays a fundamental role in how students learn. Scientific evidence indicates that learning processes are intrinsically cultural and significantly influenced by contextual factors, social constructs, and personal perspectives

5 Elements of Racially and Culturally Inclusive Large-Scale State Assessments

Research and practice on racially and culturally inclusive pedagogy is well established, and to a considerable extent, can be adapted into guidance used for the purposes of large-scale assessment development and design.[13] EdTrust has highlighted five elements based on this body of knowledge, and provides guidance for developing assessment items that meet the criteria of these elements.

**NOTE: While classroom assessment might be the most straightforward application of racially and culturally inclusive assessment practices, it is outside the scope of this brief, which focuses on opportunities in statewide summative assessments.*

Element 1: Authentically reflect students' own cultures and identities

Assessment items must intentionally reflect a range of racial, cultural, and ethnic identities, so that all students see aspects of their identity explicitly uplifted and affirmed in the test material.[14] Assessment design choices — the selected passages, images, and questions — each provide an opportunity to uplift specific cultural elements, so they need to be diversified.

What does engaging, affirming, and authentic representation look like?

Engaging, affirming, and authentic representation comprises a full reflection of an individual's identity, and asks what culture, language, and experiences are being centered. Consider the following guiding questions as they relate to the content of an assessment item:

- What day-to-day activities are the characters doing?
- What food/customs are featured within the story?*
- Which historical figures are featured?
- What perspective is the historical event being told from?
- Who is the author, and how has their experience shaped their worldview?

NOTE: Individual items can portray myriad aspects of cultural activities that reflect unique knowledge, traditions, and beliefs. Given the limited space for each item, assessments should focus on authentically representing elements of a cultural activity, as opposed to attempting to portray the totality of a cultural aspect in an item. At the same time, focusing too heavily on food-related items runs the risk of relying on stereotypical portrayals of the characters featured in the item, and may otherwise severely limit the depiction of more holistic representation of various cultural activities.

Assessment design choices — the selected passages, image art, and questions — each provide an opportunity to uplift specific cultural elements, so they need to be diversified from their current state

Sancocho Is The Perfect Remedy

Example A



Source: Renaissance. Questions based on the passage below are omitted to avoid item exposure.

Sebastian woke up with a sore throat and a runny nose. Abuela Carmen took one look and said, "I know what you need. A bowl of my famous sancocho soup!"

Abuela Carmen didn't work alone. She had two helpers! They were Sebastian's brother and sister, Mateo and Isabel. They went to the kitchen together.

The first step was cleaning all the vegetables. Mateo rinsed the carrots, celery, and corn with water. Isabel was a little older and knew how to use a knife. She cut the vegetables up into small pieces.

The soup has a special ingredient called cassava or yuca. Cassava is dark brown on the outside and white on the inside. It tastes a little like a potato. Abuela Carmen was in charge of cutting up the yuca and chicken into square pieces.

It was time to cook! Mateo dumped everything into a big pot. Abuela Carmen turned on the stove. It wasn't long before the whole apartment smelled great! Isabel added garlic, oregano, and cilantro as a finishing touch.

When the sancocho was ready, Abuela Carmen gave a bowl to Sebastian. He put a spoonful into his mouth. "Wow, Abuela! This is delicious! I feel better already!"

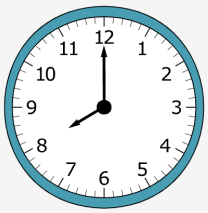
In this example, several intentional choices are worth highlighting:

- *The framing aims to ring true for a particular scenario: sancocho is made for a child that is not feeling well, and the ingredients and process of making sancocho are accurate.*
- *It is suggested that the children live with their grandmother. This is a family structure that Asian and Latino students may be more likely than White Americans to relate to, because it is more common to live in a multi-generational household and/or with extended family. While this is not explicitly addressed, the passage still intentionally includes this aspect of students' experiences.*
- *The writing emphasizes a specific culture yet balances inclusivity appropriately, so that students encountering this item do not need to be familiar with the scenario to draw meaning and demonstrate understanding. Specifically, the grandmother character is introduced as "Abuela Carmen" and then later, just "Abuela." The meal being prepared is introduced as "sancocho soup" and then later, just "sancocho." The passage includes cassava/yuca as a prominent and important special ingredient — yet makes a rough comparison that the root "tastes a little like a potato."*

Kayla's Bedtime

Example B

The clock shows when Kayla goes to bed.
What time does Kayla go to bed?





Source: Curriculum Associates

This example features an image that depicts a child, presumably Kayla, getting tucked into bed by a female adult, presumably her mother. Both women are wearing bonnets, which is a common garment for Black women to wear to sleep. The image also depicts Kayla with a different skin tone as her mother. Image considerations such as these can help items feel more relevant and authentic to students.

Element 2: Authentically represent cultures and identities of others

Racially and culturally inclusive assessments should provide not only mirrors for students to see their own culture and identity reflected, but also windows into other cultures. This combination of windows and mirrors within assessment items also develops a positive attitude toward cultural differences.[15]

In this way, the goal of racially and culturally inclusive assessments is not to have each student resonate with each question. For every one assessment item that provides a mirror for one student, that same item offers a window for another student. Moreover, the goal among assessment items is to balance representation across myriad cultures, interests, and identities, diversifying exposure to multiple identities and interests.

Element 3: Intentionally include important contextual and cultural information

To create assessments with authentic, affirming representations of students' identities (mirrors) as well opportunities to learn about the identities of others (windows), items need to often include additional context in the text.

Accurate and clear context serves dual purposes:

1. For students where the reference resonates with their own identity, contextual details can help an item feel accurate in its representation.
2. For students where the reference does not reflect an aspect of their identity, context can help clarify the scenario/item.

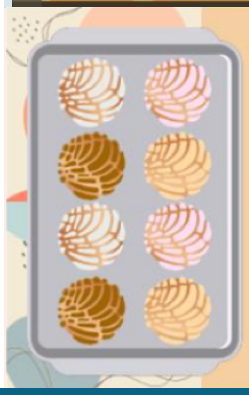
Intentionally including cultural topics that are familiar to only some students (and therefore unfamiliar to others) should be encouraged; however, assessment development companies need to carefully ensure that context is provided. This is most straightforward when cultural references in the assessment items clearly feature an obvious cultural aspect or behavior — an assessment item related to music, literature, traditions, food, fashion, festivals, etc. In this case, assessment developers should be defining terms or providing explanatory text. This is not to suggest that providing explanations is only necessary for a non-White cultural reference; while being cognizant of item length, context across all items can challenge the standardization of Whiteness.

Example C

Carmen's Panaderia



Carmen is a pastry chef who owns her own panaderia, or bakery. From fresh loaves of bread, pies, croissants, and cakes to muffins, donuts, and all kinds of pan dulce, there's nothing she can't confect!



$x ____ = 72$

Source: Renaissance

Let's take a look at her baking plans for the upcoming week. Carmen wants to create conchas, a soft and sweet Mexican bread named for their seashell-like appearance. She wants to create 72 conchas in one batch. If each baking tray holds 8 conchas, how many baking trays will Carmen need?

This example features Carmen planning how to bake her conchas. Carmen is Latina and is a business owner. She is making a range of desserts, including conchas, pan dulces, bread, pies, and croissants. The image is representative of both what conchas look like, who Carmen is, and what her panaderia looks like. Conchas, a traditional Mexican sweet bread, are the primary subject of the assessment item, and there is additional context about the concha, which is described as "soft and sweet Mexican bread named for their seashell-like appearance." This context allows students familiar with conchas to see an item that is reflective of their experiences and those without prior knowledge to learn about the bread.

Element 4. Reflect student interests and intersecting elements of identity

Students — like all individuals — have complex identities, interests, and experiences. Diversifying assessment item design entails accounting for all of the ways items can authentically depict identities, far exceeding content that falls strictly along racial or cultural lines.[16] Moreover, assessment items must include items that showcase a range of interests, religions, abilities, sexual orientations, gender identities, age, national origins, and the intersections across these identities.[17] Recent research from both WestEd and The Center for Measurement Justice indicates that student interests (such as music genres, sports and games, and extracurricular curricular activities) are a major driver of student engagement with assessment items, often transcending ethnic/racial commonalities.[18]

The Rhythm of My Street

Example D

On a hot summer day,
My friends and I sit outside
waiting for the cool of the night.
I hear a faint tap-tapping down the street.
What is that sound?
It's my neighbors at their dominoes table!
I tap my toes to the beat.
My friends hear it, too.
They nod along to the rhythm.
Another beat drops in!
A thud, thud, thud
Of footsteps coming home from a long day at work.
The sweet jingling song of the ice cream truck
joins in.
Then comes the crack of the bat on a ball from the
baseball park.
A car horn blows, a bird chirps, a passing
neighbor calls, "Hello,"
adding new pieces to the street's song.
We tap and nod along.
But soon it's time to say goodnight.
We hear the final sounds of the day:
Goodbyes and–
Promises to listen again tomorrow.



Source: Renaissance

This example features a poem that is not specific to one culture, race, or ethnicity, but illustrates a multicultural neighborhood in an urban environment. It likens the noise of the neighborhood — across a variety of activities that span hobbies, identities, and cultures — as an infectious rhythm and harmonious song. The image art also reinforces the diversity of the community across race, culture, and ability.

Dora Learning Sign Language

Example E

Dora is learning sign language because she has a friend who is deaf. Dora learned 16 new words yesterday and 29 new words today. How many new words did Dora learn in sign language yesterday and today?

Source: Curriculum Associates



This example features a scenario where Dora is practicing sign language. While Dora is depicted as Latina, she is not engaging in an activity specific to her culture; rather, she is learning a new language to better communicate with her friend who is deaf.

Element 5. Honesty about students' realities — both opportunities and challenges[19]

A hallmark of racially and culturally inclusive pedagogy is building students' sociocultural consciousness, or an "awareness of the systems/structures in place that impact their lives and communities."[20]

How to appropriately integrate sociocultural consciousness into statewide summative assessment items is less established than the other elements, but one promising approach may be to scaffold the integration of the tenets of cultural consciousness. In early grades, the focus could involve items promoting inclusion, comfort, and validation for all students. Subsequently, in later grades, the assessment can include items that foster cultural consciousness by discussing topics, including systems of oppression, in a nuanced way.[21]

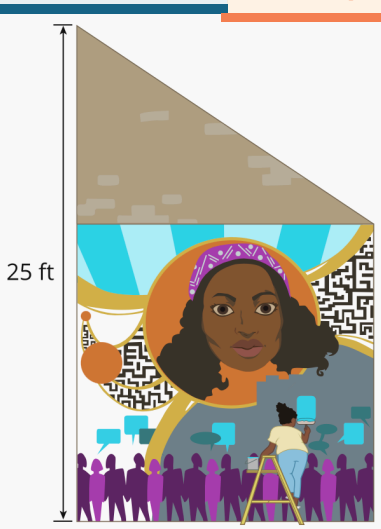
Community Mural

Example F

An artist is painting a mural on the wall of a community center. The artist has already completed a square portion of the mural that has an area of 225 square feet. She has a triangular portion of the wall to complete. The area of the triangular portion of the wall can be represented by the expression:

$$\frac{1}{2}\sqrt{225}\,(25 - \sqrt{225})$$

A diagram of the wall is shown.



Source: Curriculum Associates

This example features a scenario where an artist is painting a mural at a community center, an activity that shows pride and expression through art, activism, and community power.

Community Garden

Example G



Source: Renaissance

Andrea lives in part of the city where it is difficult to buy high-quality, fresh food. To help the people in her neighborhood, she wants to start a local community garden. Recently, Andrea convinced her city to sponsor its first urban gardening program by providing a vacant lot to use. Now that Andrea has land for the garden, she could use some help deciding how to plan the garden beds.

The content connects to structures (grocery stores) and familiar ways of describing food (high-quality, fresh). It also depicts experiences that most students can readily understand (difficulty getting high-quality, fresh food), but is not situated within national politics. However, it still exposes systems of oppression in concrete, local ways.

Creating Enabling Conditions Within Assessment Development Companies

Myriad adjustments need to occur within assessment companies to add capacity and infrastructure to support inclusive assessment design.

Communities should be empowered to lead on what feels authentic in assessment item design, and students should be emboldened to share what makes for relevant, engaging content.

Minimally, assessment development companies need to be engaging in the following five improvement processes:

1. Develop an inclusive item bank

As an initial step, assessment development companies should seek to understand the current state of their item bank (the available question and answer choices that are pulled together to form an assessment). Auditing the items that are currently being used will identify which identities are being explicitly featured. For instance, when the assessment company Renaissance conducted an image audit, tagging the gender and race/ethnicity of each image presented in the assessment, it revealed that more than half (54%) of their images were of White characters.[22]

Conducting an audit that tallies race/ethnicity and other demographic information within the current item bank is not to suggest that assessment developers can simply make swaps in demographics (e.g., switching out White characters for characters of color). Rather, the purpose of conducting an item bank audit is to establish a baseline data point and then write authentic and representative items more comprehensively, which requires re-examination and revision.

For example, Cognia revised the cultural context for a math item, by drafting the following, in which the original item of waffles was swapped out for caldo, a type of traditional Mexican soup that can be made with any meat: "Louisa is making caldo. The table shows the number of cups of water needed for the number of servings. How many servings of caldo can be made using 4 cups of water?" The content specialist later realized this was not an adequate revision, "given that measuring ingredients is not commonplace for meal preparation in [this] culture."[23]

Given the additional considerations needed to authentically develop items, assessment companies need to adjust their internal processes to allow for significantly more time and money.[24] For instance, Renaissance's item writers took 30% more time to write an item to conduct research and fact check to ensure content was inclusive, accurate, and authentic. This additional time translates to higher expenses.[25]

2. Diversify item writers and provide additional support

Assessment companies should prioritize diversifying their staff, particularly roles that are responsible for item creation and review. This is because item writers naturally draft assessment items based on a set of assumptions, often implicit, that represent their own set of experiences. When all item writers share a cultural background, assessment items will reflect their significant blind spots. By hiring individuals with diverse experiences across a range of races, cultures, and identities, assessment items are more likely to be constructed with nuance and authenticity.

While this is a necessary first step to give assessment companies a better chance of representing diverse identities within assessment items, hiring a more diverse set of item writers is insufficient on its own. Having an item writer or reviewer of color does not necessarily ensure items and review processes adequately recognize and disrupt the ways in which dominant perspectives of White culture pervade the assessment design and review process.[26] Assessment development companies must also commit to dismantling the pervasive centering of White culture in item design, beyond having a more diverse staff. To do so, they must also commit to raising the critical consciousness of the profession by providing training and supports to help item writers incorporate cultural perspectives and recognize blind spots within their work.[27]

3. Update bias and sensitivity guidelines

Each assessment development company writes items in accordance with guidelines, known as bias and sensitivity guidelines, intended to safeguard against biased content and topics that may be particularly sensitive to students. These guidelines, as they stand, present a significant barrier to the inclusion of culture-rich questions.[28] While the intention behind these guidelines is to avoid privileging any one specific culture, they often result in prioritizing a predominantly White perspective.[29] By their very design, these guidelines often strike down assessment items that feature culture-rich content, explicitly preventing these items from being included in assessments.

For instance, one test publisher's guidelines note that "immigration must be treated factually and objectively if the inclusion of the topic is important to measure a state standard" — which may prompt assessments to exclude this as a topic altogether. These guidelines also advise that religion is a topic that should be treated with great care, with the following guidance: "It is acceptable to mention religion. For example, noting that Buddhism is one of the main religions in Singapore is acceptable. Going into detail about the practices of adherents of Buddhism is not acceptable."[30] With this guidance, bias and sensitivity reviews may be likely to disproportionately remove passages by authors of colors, if they are more likely to discuss topics — such as immigration or Buddhist traditions, for example — listed in bias/sensitivity guidelines to strike.

Additionally, these guidelines limit topics that are most engaging and relevant to students.[31] For instance, guidelines from a well-known test publisher dictate that items should "avoid using the names of commercial brands, companies, and other business entities,"[32] despite students' interest in them. Similarly, another test publisher's guidelines advised that "unless important for validity, avoid biographical passages that focus on live celebrities, whose future actions are unpredictable and may result in fairness problems."[33] Other guidelines ask item writers to avoid content related to "climate change caused by human behavior"[34] or "treatment for serious illnesses."[35] Taken together, current bias and sensitivity guidelines contribute to assessment content that presents as being out of touch and contrasted from students' reality, severing the connection between our student's educational experiences and relevant, interesting real-world connections to their learning.[36]

Resistance to revising bias and sensitivity guidelines often stems from concern that a higher threshold of sensitive topics may provoke emotional responses in students, causing harm to their well-being and interfering with their ability to fully demonstrate their knowledge and skills on an assessment.[37] These notions of what will provoke students may often be far-fetched — for instance, an assessment developer removing an item related to migrating animals in order to avoid the notion that this could trigger a student who has immigrated themselves.[38] However, preliminary research from WestEd and the Center for Measurement Justice found no evidence of student discomfort or unease in reviewing and discussing items that showed a range of cultural contexts.[39]

Revisions should still allow assessment developers to exercise reasonable discretion in determining how to present topics in developmentally appropriate ways, and what, if any, themes should be treated with a heightened level of care. But the central question must shift from a place of fear and deficit ("How might this item upset the student?") to one of validation and relevancy ("How might this item engage the student?").[40]

4. Refine item and test bias analyses

In addition to bias and sensitivity reviews, assessment items undergo a range of additional checks and balances. Many of these processes need revisions as well,[41] especially given that assessment development companies consistently hold up these analyses as evidence to suggest their processes reduce bias, rather than reinforce it. However, these safeguarding measures significantly undercut racially and culturally inclusive assessment design and perpetuate a default Whiteness as the standard to judge against.

For example, a standard practice within these statistical analyses involves performing "differential item functioning (DIF)" tests. This statistical analysis aims to ensure some groups of students do not perform differently, on average, from other groups on items. However, this test fails to account for the intersectionality of identities among students because they are conducted separately based on gender, race, socioeconomic status, second language status, and special education status. Assessment development companies must develop a more sophisticated framework to consider how identities interplay with one another.[42]

Additionally, for assessments containing human-scored items, such as essays and other open-ended prompts, scorers rely on example "anchor papers" that provide sample responses at each score. These anchor papers need to showcase a set of responses that are linguistically and dialectally diverse to provide exemplar work that more authentically reflects the racial and ethnic diversity of the entire student population and how a range of students express their knowledge and skills.[43] Diversifying the pool of reviewers responsible for evaluating human-scored items can also contribute to reducing bias and ensuring an inclusive scoring process.

5. Prioritize family, student, and community involvement

Students and communities from racially and culturally diverse backgrounds must be front and center within any discussions of racially and culturally inclusive assessment item design and processes. Communities should be empowered to lead on what feels authentic in assessment item design, and students should be emboldened to share what makes for relevant, engaging content. More specifically, assessment developers should engage racially and culturally diverse focus groups of students, like the ongoing initiative by WestEd and the Center for Measurement Justice, to garner insights into what students perceive as interesting, relevant, and authentic.[44] Curriculum Associates also has developed a multiyear research agenda, alongside nationally recognized advisors and partners, to rehumanize assessments with a design thinking approach that investigates the best possible ways to capture students' knowledge by "[focusing] on all students' experiences to create assessments that are fair for all because of inclusive content and test design" and center student voice in assessment design.[45] The 2028 NAEP Science Assessment Framework has also made positive strides to ensure that assessment items allow for students to "see themselves and their communities represented in the items across the assessment as a whole [as well as] the range of assets and 'funds of knowledge' diverse learners bring to the table." The framework goes on to describe that "[f]unds of knowledge are considered to be the historical accumulation of abilities, bodies of knowledge, assets, and cultural ways of interacting that a student might possess.[46]

Additionally, assessment development companies should conduct in-depth "cognitive interviews" with students, which can reveal more nuanced details about how students consume information and perceive various draft test items.[47] These interviews can identify how individual identities and contexts shape their knowledge, worldviews, and learning approaches (sometimes referred to as "ways of knowing"). For instance, the 2028 NAEP Science Assessment Framework provides an example (Exhibit 3.20) of how assessment items can incorporate "the use of non-traditional evidence sources," such as multi-generational/elder accounts, which holds significant cultural validity in non-White cultures, especially within indigenous cultures.[48] This level of insight is pivotal in crafting assessments that resonate with a wide spectrum of learners and encompass an array of unique perspectives and cognitive processes.[49]

The Role of Advocates

Advocates play a role in encouraging the transformation of statewide summative assessments to authentically reflect students' culture, identities, and interests (and ideally, alongside advocating for engaging, affirming, and authentic curricula). It will ultimately be the advocacy communities' pressure that gives way to positive change: The vast majority of states and assessment development companies may not have the capacity or financial incentive to take up this work on their own accord.

Projects to transform large-scale assessments to be racially and culturally inclusive are still largely nascent and uncommon, but state-level advocates can and should exert significant influence now to educate for representative assessment content. Advocates should focus on engaging with state leaders, including policymakers, administrators, and legislators, who can in turn help direct their assessment developers to create racially and culturally inclusive assessment items.

Advocacy Opportunities and State Policy Recommendations

Given that states hold substantial influence as the customers of these assessments, state leaders can harness their consumer power and force changes from their assessment providers. Advocates can take the following actions to put pressure on state actors to do so:

1. **Mobilize students and educators to share their experiences with assessments, underscoring instances where their interests, cultures, and ethnicities are not captured in large scale assessment.**

Advocates can share how the lack of authentic representation in current large-scale assessments affects their community and why diversified item banks are needed. Amplifying voices about this lack of representation is not to suggest that the ask among advocates should be to relax or remove federal requirements on annual statewide assessments. Rather, advocates can mobilize in service of improving these assessments, which can provide more valuable information on how school systems can better serve all students, especially students of color and students from low-income backgrounds.

2. **Call attention to the ways bias and sensitivity guidelines are limiting representation of real life, including culture and interests, in assessments.**

Advocates can also consider highlighting some of the more far-fetched guidelines, such as the examples discussed earlier, to make this point. Advocates may also be able to find their assessment developer's bias and sensitivity guidelines online and review them for additional examples.

3. **Urge state departments of education to require assessment development vendors to engage in practices that will result in more inclusive assessments.**

State leaders who are considering new assessment development vendors can use their Request For Proposal (RFP) application process to include specific prompts that require vendors to be inclusive across a range of races, cultures, identities, and interests. State leaders who decide to continue with the same vendor can use contract renewal negotiations to do the same.

Specifically, advocates can urge their state leaders to prompt responses within their RFP process about how assessment developers plan to promote racially and culturally inclusive assessments throughout the entire design and development process, including developing inclusive items, improving processes, and engaging in focus groups with diverse groups of students and communities. While contract expiration dates are available to look up online via the state's bids and contracts database, there is unlikely to be an open comment period or other formalized opportunities for the advocacy community to voice support or concern when a state is working with a vendor on an RFP or contract extension/renewal. Therefore, advocates should leverage any existing relationships with their state education agency to offer feedback.

If a state uses the Center for Assessment's proposed RFP Outline, or something similar, these prompts can be integrated into Sec. 5.4,-5.7, Sec. 6.5, and Sec. 10. They can also be added separately as a section inquiring about the commitments and processes the vendor proposes

4. **Push state legislators to allocate additional assessment funding specifically to developing inclusive content and improving processes to allow for inclusive content.**

Advocates can ask their state legislators to allot additional funding toward their assessment development company for the sole use of improving for processes and products that explicitly support racially and culturally inclusive statewide annual assessments. As discussed, a range of systems need significant attention, and there are real costs associated with these changes. State legislative action can lessen this load by allocating funding to help offset the costs, thereby incentivizing assessment developers to take on this work.

5. **Urge state legislators or state department of education to create a task force of experts and community members to ensure inclusive curricula and assessments.**

Advocates can suggest their state develop a task force that is responsible for providing recommendations to assessment developers that meet their state's specific context. This includes allowing researchers, students, and psychometricians to continue to build evidence, test core research questions, develop specific technical solutions, and create channels to engage with diverse communities.

Conclusion

It's imperative that education advocates, practitioners, policymakers, and leaders intentionally center student's race, culture, identities, and interests in to design school experiences for our nation's children. This includes complex, diverse representation of people, groups, and topics in curricular materials; and to extend even further, everyone should work together to improve the representation found in assessments, including statewide annual assessments.

Statewide assessments provide the only comparable data on student academic achievement, enabling states and district leaders to target resources to those with the most academic needs. By acknowledging the profound influence of culture on learning, dismantling outdated assumptions, and confronting the current biases embedded in assessment design, advocates can push their state leaders to forge a path toward modifying statewide assessments that empowers all students, especially students of color and students from low-income backgrounds, to demonstrate what they truly know and can do — improving data that drives decision-making.

The contents of this report were heavily informed by expert voices on this topic. EdTrust would like to acknowledge the following individuals who kindly offered their perspectives across research, policy, and practice: Molly Faulkner Bond (WestEd), Chastity McFarlan (Renaissance Learning), Bob Mislevy (formerly of ETS), Randy Bennett (ETS), Susan Lyons (Lyons Assessment Consulting), Dan Farley (Oregon Department of Education), Kristen Huff (Curriculum Associates), Nirupa Mathew (Curriculum Associates), Juan D'Brot (Center for Assessment), and Erika Landl (Center for Assessment). Many thanks to Chastity McFarlan, Kristen Huff, Nirupa Mathew, and Jennifer Randall (Center for Measurement Justice) for also reviewing earlier versions of this product.

Endnotes

1. Munyan-Penney, N., Mehrotra, S. (2023). Future of Assessments: Centering Equity and the Lived Experiences of Students, Families, and Educators. The Education Trust.

2. Olsen, K. (2021). Promoting inclusive classroom dynamics in higher education: Research-based pedagogical guide for faculty. Stylus Publishing and Stephens, N.M., Brannon, T. N., Markus, H. R., & J. E. Nelson. (2015). Feeling at home in college: Fortifying school-relevant selves to reduce social class disparities in higher education. Social Issues and Policy Review, 9(1), 1-24.

3. Bowman, N. A. (2011). Promoting participation in a diverse democracy: A meta-analysis of college diversity experiences and civic engagement. Review of Educational Research, 81, 29–68 and Cabrera, N. L., Milem, J. F., Jaquette, O., & Marx, R. W. (2014). Missing the (student achievement) forest for all the (political) trees: Empiricism and the Mexican American studies controversy in Tucson. American Educational Research Journal, 51(6).

4. Dee, T., & Penner, E. (2017). The causal effects of cultural relevance. American Educational Research Journal, 54(1), 127-166 and Bonilla, S., Dee, T. S., & Penner, E. K. (2021). Ethnic studies increases longer-run academic engagement and attainment. Proceedings of the National Academy of Sciences, 118(37).

5. Sleeter, C. E. (2011). The academic and social value of ethnic studies: A research review. National Education Association Research Department and Gurin, P., Nagda, B. A., & Zúñiga, X. (2013). Dialogue across difference: Practice, theory, and research on intergroup dialogue. Russell Sage Foundation.

6. Faulkner-Bond, M., Kannan, P. (2023). Student Reactions to Culturally Responsive Mathematics Assessment Items. Equity In Focus. WestEd. Blog.

7. Randall, J. (2021). 'Color-Neutral' Is Not a Thing: Redefining Construct Definition and Representation through a Justice-Oriented Critical Antiracist Lens. Educational Measurement: Issues and Practice. 4040: 82-90. https://doi.org/10.1111/emip.12429.

8. Randall, 'Color-Neutral' Is Not a Thing and Randall, J. (2021, May). The Road to Assessment Hell is Paved with Color-Blind Intentions: Moving Towards Justice-Oriented Antiracist Assessment. Assessment for Learning Conference.

9. Lyons, S., Johnson, M. and Hinds, B.F. (2021) A call to action: Confronting inequity in assessment. Lyons Assessment Consulting.

10. Ibid.

11. Faulkner-Bond, M. (2022). Cultural Relevance Versus Construct Relevance: How do we Create Culturally Responsive Assessments? NCME Webinar.

12. Ibid.

13. McFarlan, C. (2022). Cultural Relevance Versus Construct Relevance: How do we Create Culturally Responsive Assessments? NCME Webinar.

14. McFarlan, C. (2022). Cultural Relevance Versus Construct Relevance: How do we Create Culturally Responsive Assessments? NCME Webinar; Lyons et al., A call to action; Bennett, R. (2023) Toward a Theory of Socioculturally Responsive Assessment, Educational Assessment, 28:2, 83-104.; and Randall, J. (2023). Seeking Justice-Oriented Assessment: Culturally Responsive and Antiracist Item Development and Review. Presented at National Conference on Student Assessment.

15. Huff, K. (2022). DEI in Assessment Practices. NCME Webinar.

16. Faulkner-Bond, M. (2023). Cultural Content in Assessment Items: Student Choice, Self-Report, and Affect. Presented at National Conference on Student Assessment.

17. Randall, Seeking Justice-Oriented Assessment.

18. Ibid.

19. Randall, Seeking Justice-Oriented Assessment; Randall, 'Color-Neutral' Is Not a Thing; McFarlan, Cultural Relevance Versus Construct Relevance.

20. Ladson Billings, G. (1995). Toward a Theory of Culturally Relevant Pedagogy. American Educational Research Journal, Vol. 32, No. 3. 465-491.

21. Nwadike, A. (2022). Cultural Relevance Versus Construct Relevance: How do we Create Culturally Responsive Assessments? NCME Webinar.

22. Lyons et al., A call to action.

23. Ibid.

24. McClarty, K. (2022). Cultural Relevance Versus Construct Relevance: How do we Create Culturally Responsive Assessments? NCME Webinar.

25. Lyons, S., Sanford, S., Hinds, F., Dirksen, D. (2022). Engaging Communities to Develop Culturally Responsive Assessment. Cognia and Lyons Assessment Consulting.

26. Huff, DEI in Assessment Practices, and Mathew, N. (2023). Rehumanizing Assessments. Presented at National Conference on Student Assessment.

27. Ibid.
28. Randall, 'Color-Neutral' Is Not a Thing and Randall, The Road to Assessment Hell.
29. Ibid.
30. Randall, J. (2023): It Ain't Near 'Bout Fair: Re-Envisioning the Bias and Sensitivity Review Process from a Justice-Oriented Antiracist Perspective, Educational Assessment, DOI: 10.1080/10627197.2023.2223924; Randall, 'Color-Neutral' Is Not a Thing; and Lyons et al., A call to action.
31. Ibid.
32. Smarter Balanced Assessment Consortium: Bias and Sensitivity Guidelines. (2022).
33. Faulkner-Bond, M. (2023). Cultural Content in Assessment Items: Student Choice, Self-Report, and Affect. Presented at National Conference on Student Assessment and Lyons et al., A call to action.
34. Smarter Balanced, Bias and Sensitivity Guidelines.
35. ETS Guidelines for Developing Fair Tests and Communications. (2022).
36. Smarter Balanced, Bias and Sensitivity Guidelines.
37. Ibid.
38. Note that Smarter Balanced has begun the process of updating some of its bias and sensitivity guidelines.
39. Smarter Balanced, Bias and Sensitivity Guidelines.; Randall, J. (2023). See also Maine's "Passage Quality Checklist" as part of their NWEA Through-Year Assessment Blueprint: https://www.maine.gov/doe/sites/maine.gov.doe/files/inline-files/Through%20Year%20Assessment%20Blueprint.pdf; Lyons et al., A call to action.
40. Lyons et al., A call to action.
41. Randall, J. (2023). Seeking Justice-Oriented Assessment: Culturally Responsive and Antiracist Item Development and Review and Faulkner-Bond, M. (2023). Cultural Content in Assessment Items: Student Choice, Self-Report, and Affect.
42. Ibid.
43. Lyons et al., A call to action.
44. Russell, M., Kaplan, L. (2021) "An Intersectional Approach to Differential Item Functioning:Reflecting Configurations of Inequality," Practical Assessment, Research, and Evaluation: Vol. 26, Article 21
45. Lyons et al., A call to action.
46. Ibid.
47. Randall, Seeking Justice-Oriented Assessment and Faulkner-Bond, Cultural Content in Assessment Items.
48. Mathew, N. (2023). Rehumanizing Assessments. Presented at National Conference on Student Assessment.
49. 2028 NAEP Science Assessment Framework. (2023). Available at: https://www.nagb.gov/content/dam/nagb/en/documents/publications/frameworks/science/2028-naep-science-framework.pdf.
50. Bennett, Toward a Theory and Biernacki, P. J. (2022). Assessing the options: Considerations for provision of choice in assessment. Smarter Balanced Assessment Consortium.
51. Nelson-Barber, S., & Trumbull, E. (2007). Making assessment practices valid for Indigenous American students. Journal of American Indian Education, 46(3), 132-147.
52. Biernacki, Assessing the options: Considerations for provision of choice in assessment.