**Research Report**

2023-12

# Developing Culturally Relevant Math and Science Items

## Lessons Learned and Student Reactions

JEFFREY T. STEEDLE, CRISTINA ANGUIANO-CARRASCO, NANCY LEWIN, AND JILL MCVEY

## Conclusions

For this project, ACT developed math and science items that attempted to represent unique aspects of cultures, raise awareness of social justice issues, promote cultural learning, be authentic and relatable, be comprehensible, and represent people in positive, nonstereotypical ways. Creating culturally relevant items was a new challenge for ACT test developers, and it involved many rounds of review by diverse content experts. The developers learned that creating culturally relevant items is a more demanding process, but they all reported that it was a positive, enriching experience. During focus groups, high school students viewed the culturally relevant items alongside non–culturally relevant items measuring the same skills. The students expressed interest in seeing culturally relevant items on the ACT, and they reported learning about cultures from the items. However, many students expressed concern about the additional time required to answer longer items during a timed high-stakes assessment like the ACT.

## So What?

Legacy test item development processes—though intended to minimize the possibility of bias—may introduce bias because allegedly neutral item contexts can reflect White cultural norms. Creating culturally relevant assessments is one approach to mitigating that risk by framing differences as assets in a diverse world. Culturally relevant assessment seeks to increase equity for success with test items that elevate different cultures and spread cultural knowledge. This work embodies ACT's commitment to equity, which is summarized in ACT's North Star: "We exist to fight for fairness in education and create a world where everyone can discover and fulfill their potential." Focus group participants reported that seeing themselves represented in test items made them feel included, motivated, empowered, and supported.

## Now What?

With the lessons learned from this project, ACT will continue to develop and evaluate culturally relevant math and science items. In the future, content developers plan to focus on one major goal of culturally relevant assessment per item (e.g., including positive representation that disrupts stereotypes, highlighting a social justice issue, or promoting cultural learning) and to better integrate cultural contexts into the assessment content. Such items would minimize the potential impact of cultural contexts on test difficulty and speededness while still achieving the many goals of culturally relevant assessment. In a follow-up study, we will gather item response data through ACT item field-testing to compare the psychometric properties of culturally relevant and non–culturally relevant math items.

## About the Authors

### Jeffrey T. Steedle, PhD

Jeffrey Steedle is a Principal Solutions Designer in the ACT Measurement Research and Development department, where he conceives of and implements research and development projects for large-scale standardized testing programs. He holds a doctorate in Educational Psychology as well as master's degrees in Statistics and Education. His research interests include testing motivation, item difficulty modeling, score comparability across testing contexts, and differential item functioning.

### Cristina Anguiano-Carrasco, PhD

Cristina Anguiano-Carrasco, PhD, is a Senior Research Scientist in ACT's Center for Social, Emotional, and Academic Learning. She completed her doctoral dissertation in Behavioral Assessment and Measurement at the Universitat Rovira i Virgili, Spain, in 2013. Her research focuses on social and emotional skills, related response biases, and improving the measurement and assessment of constructs such as grit-tenacity, cross-cultural competence, teamwork, and self-efficacy. Dr. Anguiano-Carrasco has served as a member of NCME's CODIT for four years and was the committee's chair in 2020–2021.

### Nancy Lewin, EdD

Dr. Nancy Lewin is Senior Director with ACT's Center for Equity in Learning. She formerly served as a nonprofit executive director at the Association of Latino Administrators and Superintendents in Washington, DC. She has twenty-three years of education experience as a bilingual teacher, school administrator, special education director, chief academic officer, and adjunct professor. She holds an EdD in professional leadership from the University of Houston. As a former migrant worker, she is a champion for equity and believes that access, information and opportunity, and mentoring are essential to the success of all students and leaders.

### Jill McVey, PhD

Jill McVey, PhD, is a Research Scientist in ACT's Center for Social, Emotional, and Academic Learning. She has a background in public K–12 and higher education, where she worked as a teacher, researcher, and program evaluator. She has experience conducting applied research using mixed methods approaches and has studied a range of topics in education, from a preschool social and emotional program to the efficacy of a rural teacher education program.
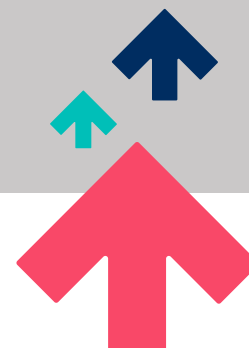
A version of this paper was presented during the virtual portion of the 2023 Annual Meeting of the National Council on Measurement in Education, which occurred March 28–30.

# Executive Summary

Like much of the educational measurement community, ACT acknowledges that legacy test development practices, though intended to minimize the risk of bias by creating "neutral" content, may result in passages and items that advantage examinees with the lived experiences of White people. One possible solution to this problem is developing culturally relevant test content. To investigate this possibility in the constrained context of high-stakes admissions testing, ACT developed math and science items that attempted to represent unique aspects of cultures, raise awareness of social justice issues, promote learning, be authentic and relatable, be comprehensible, and represent people in positive, nonstereotypical ways. The first section of this report describes lessons learned from the item development process. The development phase involved teaching experienced content developers about cultural relevance and implementing an iterative draft/review/revise process with external consultants and a diverse group of ACT team members.

Over the course of content development, we learned several lessons:

- Starting points—We could modify existing items to make them culturally relevant, but this often required updating the math or science content (numbers, expressions, data tables, graphs, etc.) to make it authentic.

- Item length—The developers struggled to avoid making passages and items much longer when adding culturally relevant content, in part because that content required additional explanation to make it comprehensible to all examinees and encourage learning about other cultures.

- Graphics—Added graphics supported representation and learning, but they were unnecessary for responding correctly to the items, and they introduced accessibility challenges for students with visual impairments.

- Representation—Based on feedback from external consultants, items were created to represent broad interest in other cultures. For example, people of many cultures can be depicted as appreciating the interests, arts, and traditions of another culture.

- Interpretation—The content developers reported spending more time deliberating over wording, and the external consultants identified unintended (negative) interpretations of words, phrases, and data.

- Comprehension—To promote comprehension, we removed construct-irrelevant technical language from science passages and rewrote them in the active voice.

- Scalability—The developers also provided several recommendations to support the scalability of developing culturally relevant content: diversifying the pool of item developers and reviewers, investing in professional development, and committing more time and resources to item development.

In the focus groups, students examined culturally relevant and standard versions of items. They provided general observations and reactions, offered their general preferences, and commented on how authentic and relatable the cultural content was, whether it was relatively engaging, and how they thought it might affect test performance. The following themes emerged from the students' comments during the focus groups:

- Item clarity and comprehension—Students expressed a preference for items with clear stems that made it easy for them to understand how the item should be solved. In some cases, students felt that the additional contextual information provided in the culturally relevant items interfered with their ability to answer the question.

- Amount of information—If students perceived the contextual information as unnecessary, they were more likely to say they were overwhelmed by the information or point out differences between the lengths of the two types of items. When the contextual information was perceived as necessary for answering the question, students were more likely to say that the culturally relevant and standard items were comparable.

- Representation of cultures—Students in all focus groups felt that representation of people and cultures was important, but in some cases, they differed on what this should look like on a test and whether it should be included in standardized tests.

- Testing time—Given the high-stakes nature of the ACT® test, students expressed concern about the limited time available to them in each section. They generally felt that shorter items would allow them to save time on the test and therefore felt more comfortable with shorter items.

- Real-world examples—Several students felt that real-world examples were preferable to generic, fictional examples, and in their perception, the culturally relevant items were more likely to feature real-world examples. Students also felt that these were more relatable than generic examples.

# Background

As evidenced by research agendas and ongoing projects, many major assessment organizations are engaged in efforts to evaluate and evolve their item development procedures in order to meet the needs of an increasingly diverse population of test takers (e.g., NCME, 2022b; NCME, 2022c). The driving force behind such efforts is acknowledgment that legacy item development practices—though intended to minimize the possibility of bias—may introduce bias. That is, items developed to be culturally neutral (or "color neutral") may reflect White cultural norms (e.g., ways of communicating, ways of representing and demonstrating knowledge). Consequently, examinees with the lived experiences of White people will be advantaged because they can more easily make sense of relatable item contexts described in familiar language (Randall, 2021).

One approach to addressing this challenge requires developing new assessments that disrupt negative stereotypes, draw attention to inequality and injustice while empowering examinees to address those challenges, present history in complete and accurate ways, and include different ways to demonstrate content mastery (Montenegro & Jankowski, 2020; J. Randall, personal communication, February 4, 2022). This recommendation follows decades of research into and practice of culturally relevant education. The framework proposed by Evans (2021) describes a progression in the ways that classroom education can be sensitive, relevant, and responsive to students' cultural identities while also sustaining them. Evans acknowledges that large-scale standardized tests have the potential to become more culturally sensitive (i.e., be more "aware that cultural differences and similarities between people exist without assigning a value to them"). However, Evans questions whether large-scale standardized tests can become culturally relevant, because that "requires *linkages* between each student's everyday lived cultural experiences and the test items or stimulus materials." Achieving that for every student all the time would be difficult indeed for an assessment like the ACT, which is administered to diverse groups of examinees nationwide (and internationally). But while the ACT cannot be culturally relevant for every examinee all the time, perhaps being culturally relevant for many examinee groups some of the time will prove beneficial.

The challenge of cultural relevance in standardized testing motivated the project described here. Researchers at ACT wondered, "In what ways can we improve cultural relevance in a large-scale, high-stakes standardized testing program?" Answering this question is primarily the domain of test content developers, but there are ample opportunities for measurement research professionals to engage with this challenge. As suggested by J. Randall (personal communication, February 4, 2022), researchers can help by asking, "Where is the resistance?" That is, researchers can identify resistance to culturally relevant assessment and investigate whether that resistance is warranted. For example, some people may resist culturally relevant assessment due to fears that it will be emotionally upsetting to some students. Others may question whether the extra effort (and expense) needed to develop culturally relevant items is worthwhile because examinees may not appreciate seeing themselves reflected in test content. Perhaps the biggest source of resistance is the notion that culturally relevant content, which may include additional text and images, will introduce construct-irrelevant difficulty or consume time during a timed test. Fortunately, these issues can be investigated empirically.

The first goal of the study reported here was to develop culturally relevant test content for the ACT math and science tests. After item development, diverse groups of high school students participated in focus groups in which they viewed the items alongside "traditional" (not culturally relevant) versions. The second goal of this study was to gather students' general observations and reactions and to collect their comments on the authenticity and relatability of the content, their thoughts on whether the cultural content was relatively engaging, their perceptions of potential effects on test performance, and their general preferences. The first part of this report describes content development procedures and lessons learned. The second part presents the methods and results of the focus groups. Results from this study can help advance the practice of culturally relevant item development and increase understanding of the effects of culturally relevant content on students, particularly for organizations that administer large-scale, high-stakes standardized assessments to high school students.

# Part 1: Content Development

## Preparation for Development

Six experienced content developers volunteered to participate in this project—three from ACT's math development team and three from the science team. Prior to participation, they read and discussed recent publications about culturally relevant assessment and its rationale and participated in a cross-functional "community of practice" at ACT focused on diversity, equity, and inclusion. This ongoing series of discussions and interactive presentations was attended by approximately 30–40 team members each month, including content developers, researchers, psychometricians, accommodations experts, and representatives of ACT's Center for Equity in Learning.

The math and science developers were aware of culturally relevant assessment, they recognized its importance, and they expressed eagerness to participate in this project. This project was their first attempt at developing ACT content that was specifically culturally relevant, and they had concerns about how cultural relevance could be incorporated into ACT math and science content given current test specifications and related content development constraints. Moreover, they all identified as White (non-Hispanic) women, primarily with experience living in the Midwestern United States, which they recognized as a limitation. That is, their lived experiences restricted their abilities to generate ideas for culturally relevant items and to later evaluate the authenticity and relatability of the content. This illustrates a serious limitation of the assessment industry's current capacity to develop culturally relevant content. Indeed, several industry leaders have raised awareness of the need to diversify the pool of item writers (e.g., NCME, 2022a). As described in the sections that follow (Consultants and Partnerships, Contexts Survey, and Training), several steps were taken to mitigate this limitation.

Given the limitations of the available content developers, the current study should be considered a best effort by existing staff members to develop culturally relevant content for a large-scale, high-stakes admissions test. This was a small proof-of-concept study, so it was not possible to represent a broad array of cultures in the test content. Each item developed for this project was intended to represent the culture, interests, and concerns of Black or Hispanic/Latinx students.

We acknowledge that there is significant heterogeneity within those groups, so not every item was expected to be relevant to each student. For example, an item about a Mexican tradition would not be relevant to students of Puerto Rican or Dominican heritage.

## Consultants and Partnerships

Two educational measurement professors were recruited as consultants for this project. Dr. Kyndra Middleton is a professor in the Department of Human Development & Psychoeducational Studies at Howard University, with research interests in mathematics achievement, comparative studies of issues faced in rural and urban education, increasing the number of underrepresented minorities in STEM (science, technology, engineering, and math) fields, and test score validity. Dr. Joseph Rios is an assistant professor in the Department of Educational Psychology at the University of Minnesota, with research interests in testing motivation, assessing English learners, and increasing inclusion in both assessment design and score reporting. Both consultants had postdoctoral experience in the educational testing industry, and they both serve on technical advisory committees for large-scale assessment programs. In this study, the consultants' main responsibilities were reviewing in-progress items, conducting final item reviews, providing methodological guidance about the focus groups, and reviewing draft reports from this study.

Throughout this project, the ACT Research team partnered with ACT's Center for Equity and Learning. In particular, Dr. Nancy Lewin, Senior Director at the Center for Equity and Learning, took a leadership role on the project. She formerly served as an executive director at the Association of Latino Administrators and Superintendents, and she draws from 23 years of education experience as a teacher, school administrator, special education director, chief academic officer, and adjunct professor. Dr. Lewin worked with the ACT research team on nearly every aspect of this project, including item development and review, focus group recruitment and facilitation, and dissemination of the results.

## Contexts Survey

A survey was delivered to a culturally diverse group of ACT employees to solicit ideas for contexts that could be used in culturally relevant math or science items. Table 1 lists their ideas and categorizes them into the following groups: politics and social change, traditions, food, arts, sports, and STEM. As shown in Table 1, accomplished individuals accounted for many of the ideas. Though the survey respondents were not item developers, they also provided ideas for connecting culturally relevant contexts to math or science content (Table 2). A few of these ideas influenced items eventually developed for this project.

**Table 1.** Item Context Ideas Provided by ACT Team Members

| Category | Examples |
|---|---|
| Politics and social change | César Chávez, Shirley Chisholm, Kamala Harris, Dolores Huerta, Barack Obama, Sylvia Rivera |
| Traditions | Día de los Muertos, quinceañeras, posadas, gatherings with extended family |
| Food | Going to the market, traditional foods (e.g., tamales, pozole, ropa vieja) |
| Arts | Josephine Baker, Sandra Cisneros, Misty Copeland, Michael Ford, Gwen Ifill, Lin-Manuel Miranda, Selena Quintanilla, Flor de Toloache, Harlem Renaissance, regional music, Stax Records |
| Sports | Cynthia Marshall, Serena and Venus Williams |
| STEM | George Washington Carver, Dr. Charles Drew, Jaime Escalante, Victor Glover, Dr. Mae Jemison, Katherine Johnson, Garrett Morgan, Ellen Ochoa |
| Other | National flags, first nations, hairstyling |

**Table 2.** Connections to Math or Science Content Provided by ACT Team Members

| Context | Examples |
|---|---|
| Planning a party/gathering | Math is often helpful for planning a large party or gathering such as a quinceañera (e.g., predicting attendance, planning time for everyone to get dressed and styled). |
| Cooking traditional foods | Food or cooking contexts can be used when assessing measurement and ratios. Math may also be used to plan food quantities (e.g., an appropriate number of tamales). |
| Labor economics | César Chávez and Dolores Huerta can provide context for math items related to labor economics (e.g., wages and produce prices). |
| Entertainment | Lin-Manuel Miranda and the musical *Hamilton* can provide context for math items related to calculating ticket sales. |
| Time | Garrett Morgan's invention of the stoplight can provide context for math items related to time and rates. Mixing chemicals for hairstyling can provide context for math items related to time. |
| Business | Cynthia Marshall and the NBA can provide context for math items related to business (e.g., games per season, revenue from ticket sales, and game attendance). |
| Physics | Katherine Johnson and Ellen Ochoa can provide context for math or science items related to space exploration (e.g., calculating distances). Josephine Baker and Misty Copeland can provide context for science items related to dance or movement. |
| Biology | Dr. George Washington Carver's crop rotation method can provide context for science items related to nutrients in soil. Victor Glover's experience aboard the International Space Station can provide context for biology items related to growing crops in space. |
| Medicine | Dr. Charles Drew's research in blood transfusion and cardiac surgery can provide context for science items related to anatomy. |

### *Training*

The principal investigators of this study planned a kickoff meeting with all the content developers. The meeting began with a recap of the reasons for developing culturally relevant content based on a presentation by Jennifer Randall (NCME, 2022a) to the NCME Large-Scale Assessment Special Interest Group. This included topics such as decentering Whiteness, explicitly disrupting negative stereotypes, recognizing that item context matters, and shifting from fear-oriented to justice-oriented item review guidelines. Next, we introduced the notion that developing culturally relevant content requires *funds of knowledge* (Moll et al., 1992), which are bodies of knowledge, skills, experiences, values, social and family connections, and traditions that serve as strengths and resources for members of a community. For this project, accessing funds of knowledge involved having representatives of various cultures involved in content development and review. At this point, we shared the survey results presented in Tables 1 and 2.

The next portion of the kickoff meeting addressed the topic of asset-framing, an approach developed by Trabian Shorters to guide the ways that people are represented by the organizations that serve them (https://bmecommunity.org/). Specifically, asset-framing involves defining people by their aspirations and their contributions before noting their challenges, then investing in them and their ability to continue benefitting society. Many ACT team members attended one or more presentations about asset-framing in February and October 2022, and ACT is committed to asset-framing's core principles. Asset-framing has parallels to justice-oriented assessment. For example, asset-framing seeks to "undo bias" by telling fuller narratives about people. In assessment development practice, this entails representing people—including their contributions, aspirations, and challenges—in uplifting, nonstereotypical ways and avoiding the use of terms that perpetuate stigmas (e.g., *at-risk*, *gap*, *underserved*, and *disadvantaged*).

Next, we shared numerous examples of cultural relevance in math and science education drawn from internet searches. Most examples reflected cultural relevance in classroom instruction, which has a longer history than culturally relevant standardized assessment. Also, most examples represented elementary and middle school curricula and assessments, which sometimes limited their applicability to an assessment for high school students.

The kickoff meeting concluded with a discussion of proposed criteria for culturally relevant assessment content (e.g., Randall et al., 2022; Stadum, 2020). Those criteria were summarized into a list of desired properties for culturally relevant assessment content:

- represents some unique aspect of a culture

- potentially raises awareness of a social justice issue

- promotes learning about cultures and/or social justice issues

- has authentic and up-to-date cultural content

- is comprehensible to all test takers

- represents people in positive, nonstereotypical ways

- is relatable to the represented culture

- is more engaging than standard assessment content, especially to the represented culture

Each culturally relevant item was paired with a traditional (non–culturally relevant) item that assessed the same content (i.e., used the same numbers, expressions, graphs, experimental designs, etc.) and used parallel wording to the extent possible. Eventually, those item pairs would be presented to focus groups of high school students.

## The Development Process

The iterative content development process began with a period of brainstorming and drafting lasting several weeks. This included reviewing training materials and searching the web for interesting cultural facts or contexts that might be relevant to math or science. During that time, the content teams met regularly with the research team to review and revise the items. Next, the initial drafts of the items were sent to the external consultants for review. The culturally relevant items were shown alongside the traditional items. The content teams received and discussed feedback during conference calls with the external consultants. Feedback and lessons learned are discussed in subsequent sections of this report.

Next, the items were sent to an internal (ACT) panel of reviewers that included many of the individuals who provided ideas for content (Tables 1 and 2). After revisions were made based on both the external and internal reviews, the items were again sent for external review. This time, the consultants responded to the following questions about each item:

- Does the item content match with your understanding of culturally relevant?

- Does the item content match with your understanding of justice oriented?

- Does the item content promote learning about cultures and/or social justice issues?

- Is the cultural content authentic and up-to-date?

- Is the cultural content comprehensible to all test takers?

- Are persons represented in positive, non-stereotypical ways?

- Would you expect students of the represented culture to be more engaged by the cultural content (compared to the standard version)?

- If relevant, how would you suggest improving the representation and/or cultural content?

- If there are multiple versions of the item, which version do you prefer?

- What concerns, if any, do you have about the math/science content?

Based on feedback from the consultants, these evaluation questions will be revised to improve clarity (e.g., by developing operational definitions of terms such as "social justice" and "engaging") so they can be used in subsequent studies. Not all items were expected to be justice-oriented (i.e., to highlight social justice issues and empower students to address those issues), but it was a goal when opportunities arose.

Based on new feedback from the external consultants, the items were again revised. Finally, ACT content experts reevaluated the accuracy of the math and science content and recommended editorial changes. During this process, several graphics were created for the culturally relevant math items.

## Lessons Learned

### Starting Points for Item Development

For the sake of efficiency, the original development plan called for the modification of existing math and science items to make them culturally relevant. The science team followed this approach. An ACT science unit consists of five to seven items related to a passage, which often includes tables, figures, or diagrams. The creation of a science unit is a long process of researching, writing, and revising. Thus, reuse of existing science units helped make this proof-of-concept study feasible, but it also made it difficult to use many of the suggestions presented in Table 1. For example, it would have been difficult to find an existing unit (among those eligible for this project) with science content that matched the specific type of research conducted by a certain scientist. One context idea—Día de los Muertos—was ultimately used in a science unit because of its connection to monarch butterfly migration. Food deserts were discussed as a context idea during training, and this influenced the science team's work on a unit about growing vegetables in a community vegetable garden.

Currently, ACT science units are written without an explicit social context, so revising the science units mainly entailed adding a culturally relevant context. These revisions required few changes to the science content (e.g., experimental results, graphs, item stems, and response options). Three science units were revised as described in Table 3.

**Table 3.** Science Unit Summaries

| Title | Summary | Context target | Shown to focus groups |
|-------|---------|----------------|------------------------|
| Noise Pollution | A unit about the relationship between air temperature, sound frequency, and sound attenuation was modified to include students representing an environmental justice organization studying the differential impacts of noise pollution on different communities. | Urban | No |
| Butterflies | A unit about monarch butterfly migration and lipid (fat) storage was modified to include connections to Día de los Muertos and the possible role of fir tree logging in the decline of monarch butterfly populations. | Latinx | Yes |
| Worm Castings | A unit about the experimental manipulation of worm castings (feces) in potting soil and the effects on crop yields was modified to include information about a communal vegetable garden and growing crops used in traditional foods (tomatoes and chili peppers). | Social justice | Yes |

In several cases, items developed by the math team were influenced by context ideas discussed during training (e.g., planning for a quinceañera, Día de los Muertos traditions, and Lin-Manuel Miranda). Most often, the math team modified existing items by adding or changing the problem context. As described in Table 4, the results included items about calculating the volume of punch needed for a party, determining the number of different orders in which a student could visit historical sites in Chile, determining the rate at which a student should read Colson Whitehead's *The Underground Railroad* for a class, calculating the number of students who chose to visit different exhibits at the Smithsonian National Museum of African American History and Culture, analyzing survey results to gauge support for continued use of a public space for demonstrations, determining the number of gallons of paint left over from painting a mural, and calculating money saved from purchasing a dashiki (a shirt of West Africa) on sale.

Regarding the types of changes made to the items, the math developers expressed their feeling that small, seemingly superficial changes are potentially "meaningful" (e.g., changing *shirt* to *dashiki*, changing *Anna Karenina* to *The Underground Railroad*, changing a generic museum to the Smithsonian National Museum of African American History and Culture). Such changes are "not something to be avoided, but we can do more" according to one math developer. Another developer noted that these items reflect "a move in the direction of diversity, and for some math items measuring some content standards, that may be all there is to do."

The math team quickly recognized that modifying or adding context without changing the mathematical content was sometimes too restrictive. That is, changing the mathematical content (the numbers, expressions, functions, etc.) was necessary to make the culturally relevant context authentic. For example, a proportional reasoning item was modified so that it accurately represented the number of bird species in Ecuador, a trigonometry item was modified to accurately represent the height of the Temple of the Jaguar in Guatemala (rather than a tree), and a geometry item was modified to have realistic dimensions for a Día de los Muertos ofrenda.

Finally, the math team identified several contexts that were best used by creating new items. For example, one item asked students to estimate the number of words in the *Hamilton* soundtrack based on the average rate of singing or rapping. Another item asked students to model the price of a Frida Kahlo painting over time with an exponential function.

Thus, there is potential for modifying existing item content to enhance cultural relevance, but flexibility is often needed to make the mathematical content realistic or authentic to the context. Modified items would need to be field-tested before they could be administered on operational test forms. When interesting contexts are identified—particularly those that come with real-world values (e.g., the number of words in *Hamilton* or the actual sale price of a painting) or involve specific persons (e.g., Dolores Huerta, Garrett Morgan, Cynthia Marshall)—those contexts may not fit well with existing items, so creating new items is likely the best way to use such contexts in authentic, culturally relevant items.

**Table 4.** Math Item Summaries

| Title | Culturally relevant item summary | Traditional item summary | Development method | Context target | Shown to focus groups |
|---|---|---|---|---|---|
| Gallons of Punch | Identify an expression for the number of gallons of juice needed for a quinceañera party | Identify an expression for the number of gallons of punch needed for a party | Context added/changed | Latinx | Yes |
| Ecuadorean Birds | Use the number of bird species in Ecuador to estimate the number of species worldwide | Use the number of new clients who saw an advertisement to estimate the total number of new clients | Context added/changed, math content modified | Latinx | Yes |
| World Heritage Sites | Determine the number of different orders in which a student can visit five World Heritage Sites in Chile | Determine the number of different orders in which five students can stand in a line | Context added/changed | Latinx | No |
| Hamilton Lyrics | Estimate the number of words in the musical *Hamilton* using the rate of rapping/singing | Calculate the number of calls received at a call center using the rate of calls received | New item | Latinx | Yes |
| Temple of the Jaguar | Identify an expression for the height of the Temple of the Jaguar in Tikal, Guatemala | Identify an expression for the height of a tree | Context added/changed, math content modified | Latinx | Yes |
| Underground Railroad | Determine the rate at which a student should read Colson Whitehead's *The Underground Railroad* for a class | Determine the rate at which a student should read Leo Tolstoy's *Anna Karenina* for a class | Context added/changed | Black | Yes |

| Title | Culturally relevant item summary | Traditional item summary | Development method | Context target | Shown to focus groups |
|---|---|---|---|---|---|
| Smithsonian Museum | Determine how many students chose to visit different exhibits at the Smithsonian National Museum of African American History and Culture | Determine how many students chose to visit different exhibits at a local museum | Context added/changed | Black | Yes |
| Neighborhood Survey | Use survey results to estimate how many people in a neighborhood oppose continued use of public space for demonstrations | Use survey results to estimate how many people in a neighborhood oppose a curfew | Context added/changed | Social justice | Yes |
| Mural | Determine the number of gallons of paint left over after painting a mural in the style of Portuguese street artist Huariu | Determine the number of gallons of paint left over after painting a room in the city hall | Context added/changed | Black | Yes |
| Kahlo Painting | Model the price of a Frida Kahlo painting over time with an exponential function | Model the price of a piece of artwork over time with an exponential function | New item | Latinx | No |
| Ofrenda | Calculate the area of two triangular floor spaces around a Día de los Muertos ofrenda | Calculate the area of two triangles sharing sides with a square | Context added/changed, math content modified | Latinx | Yes |
| Shirt Sale | Calculate money saved when buying a dashiki (a shirt of West Africa) on sale | Calculate money saved when buying a shirt on sale | Context added/changed | Black | Yes |

## *Item Length*

Providing opportunities for cultural learning is one of the potential benefits of culturally relevant assessment. To be clear, this is not learning math or science skills and content knowledge; this is learning about cultures and their traditions, interests, arts, accomplishments, and concerns. Promoting this type of learning was an explicit goal of this project. In the science units, for example, students could learn about the differential impact of noise pollution on predominantly Black and Hispanic/Latinx communities, connections between monarch butterfly migration and Día de los Muertos, and connections between community gardens and food security. Math items included quinceañeras, biological diversity in Ecuador, UNESCO World Heritage Sites in Chile, the Temple of the Jaguar in Guatemala, exhibits at the Smithsonian National Museum of African American History and Culture, the Portuguese street artist Huariu, Frida Kahlo, Día de los Muertos ofrendas, and dashikis.

Fully understanding the culturally relevant context was not essential for responding correctly to the items. However, promoting understanding of the culturally relevant context would support cultural learning, so it was necessary to explain the context for students unfamiliar with it. For example, the Gallons of Punch item was designed to represent a Hispanic/Latina business owner and teach about quinceañeras. The item began, "Camila is the owner of a popular catering company. Camila specializes in quinceañera parties, which celebrate a girl's passage from childhood to adulthood. To plan orders for a party, Camila uses the formula . . ." Contrast this with the non–culturally relevant version, which read, "To plan orders for a party, a caterer uses the formula . . ."

The need for explanation created tension between the desire to minimize item length and the goals of promoting cultural learning and making the items comprehensible to all test takers. The item developers expressed discomfort with longer items. Likewise, the external consultants expressed concerns about trying to teach new ideas during a high-stakes test and about the possible introduction of construct-irrelevant variance caused by increased item length. The consultants noted that the cultural content, which lengthened the items on its own, often required additional explanation to make it comprehensible. It was sometimes helpful to have people unfamiliar with the cultural content review items to check their comprehensibility.

For example, the introduction to the Monarch Butterflies science passage increased from 71 words to 279 words before it was eventually shortened to 180 words. The Sound Attenuation introduction increased from 57 words to 224 words plus a new data table. An early draft of a math item included the acronym UNESCO, which required definition (United Nations Educational, Scientific, and Cultural Organization). Another draft math item stated, "Ecuador is the first country in the world to grant constitutional rights to Nature." The item developers were eager to share this interesting fact with examinees, but explaining its meaning would have made the item rather long. The item was later revised to say, "Ecuador is a world leader in protecting nature." Draft items that mentioned quinceañeras and Día de los Muertos traditions also demanded explanation.

Despite initial hesitation and concerns from external consultants, the content developers moved forward with longer items because this was an expected consequence of developing culturally relevant items. However, significant efforts were made to minimize item length while still meeting other item development goals. The content development teams reported that more time was spent revising than usual due to efforts to control item length. They also wondered whether explanations of cultural content might "otherize" the subject and make it seem out of the ordinary. Based on this concern, they suggested that guidelines for providing supplemental explanations would be helpful.

## *Representation*

Representing people in positive, nonstereotypical ways was a goal of item development. Early in the development process, it was apparent that culturally relevant contexts did not automatically include such representations. That is, an item can be culturally relevant (i.e., about cultural traditions related to family, holidays, food, arts, athletics, etc.), and this can be a positive representation of the culture, but it does not necessarily represent people in positive, nonstereotypical ways. Consider, for example, the math item about planning food for a quinceañera party. The item could simply ask examinees to calculate the gallons of punch needed. However, the final version of the item included mention of a popular female-owned catering company ("Camila is the owner of a popular catering company. Camila specializes in quinceañera parties, which celebrate a girl's passage from childhood to adulthood"). The math item about the Temple of the Jaguar could simply ask for the height of the temple, but the item begins with the text "Ancient Mayans built the Temple of the Jaguar in Tikal, Guatemala." The intention was to attribute agency to the ancient Maya, highlight their advanced technology, and associate them with a Latin American country. There may not always be opportunities to insert this kind of representation, but opportunities to do so should be sought out.

The original science units described experimental procedures, results, and hypotheses without attributing them to anyone. In contrast, the revised science units included students represented as agents of change and as capable of conducting useful science experiments and discussing scientific ideas. However, the constraint of modifying existing science units created challenges for representation. Namely, authentically reflecting real-world science carried out by individuals or organizations might require significant modifications. The external consultants cautioned against mentioning real-world organizations to avoid this difficulty, even though they believed it better to mention specific rather than generic people. For that reason, the science passages describe "a civic advocacy group" and "students in an environmental justice organization in a diverse urban community" and "students near the US-Mexico border." One of the consultants wondered whether this approach was clear and direct enough for students to find relatable.

Based on feedback from the external consultants, the content developers learned that it was generally better to represent broad interest in different cultures. Practically speaking, this meant that the culture associated with people's names should not always be the same as the culture represented in the item. For example, one math item originally said, "For her literature class, Tiana must read Colson Whitehead's *The Underground Railroad* in 8 days." However, as one consultant noted, "Everyone should read *The Underground Railroad*, not just Black students."

For that reason, the name was later changed from Tiana to Tara. Likewise, an item about a street artist whose work often features Black people originally said, "Malik and Aaliyah are fans of Portuguese street artist Huariu," but the names were later changed to Malik and Ayako.

### Interpretation

As noted previously, the content developers spent more time wordsmithing culturally relevant items, but minimizing item length was not the only reason. Indeed, developing culturally relevant items required greater deliberation about the meaning and interpretation of words and phrases. This lesson was learned primarily through discussions with the external consultants, who provided new perspectives on the content added to the math and science items. Mainly, this manifested in interpretations that were unanticipated by the predominantly White content developers.

For example, a data table (with authentic data) was added to one of the science units. This table showed the relationship between the racial/ethnic composition of urban blocks and the median nighttime noise intensity (i.e., level of noise pollution). Where the percentage of Black or Hispanic residents was higher, the level of noise pollution was higher on average. Where the percentage of White residents was higher, the level of noise pollution was lower on average. To the content developers, this information illustrated an environmental justice issue. However, the external consultants noted that this information could reinforce the notion that Black and Hispanic people are louder than White people. To address the issue, the developers further defined noise pollution ("noise generated primarily by mechanized sources such as transportation, industry, and air conditioning equipment"), but the concern persisted, and a wholly satisfactory solution was never achieved.

At one point, the math item about planning food for a quinceañera said, "Camila is the owner of a successful Mexican food catering company." The meaning of the word "successful" seemed obvious to the content developers, but the consultants asked, "What do you mean by successful? Successful by whose standards?" In the end, "successful" was replaced with "popular." Another math item used the selling prices of a Frida Kahlo painting to model exponential growth. The text read, "Her work 'Diego y Yo' sold for $1.4 million in 1990, and was later sold for $34.9 million in 2021, setting a record for the selling price of a piece of Latin American artwork." To the content developers, this text celebrated appreciation for Kahlo's artwork. However, as the consultants pointed out, it could also indicate something about the relatively low value placed on Latin American artwork compared to European artwork. Overall, discussions like these highlighted the importance of including diverse perspectives in the item development and review processes, which helped the developers avoid unintentionally and needlessly alienating examinees.

### Graphics

At the outset of this project, it was expected that some culturally relevant items would include additional graphics to help enhance representation and learning. Depending on the item context, those graphics could help students see themselves reflected in the test, or they could help students learn about something unfamiliar (e.g., a community garden, an ofrenda, a

dashiki). As examples, Figure 1 shows a painting in the style of street artist Huariu and a dashiki. Like some explanatory text added to items, these graphics were nonessential for responding correctly to the item, and the non–culturally relevant versions of these items had no graphics. While the external consultants supported the inclusion of such graphics, the extraneous nature of these graphics made their inclusion controversial among the content developers. Following current ACT style guidelines, the math developers would never include an unnecessary graphic, nor would they have a graphic that is not referenced in the item stem. They also relayed concerns related to the accessibility of the items for students with disabilities. Namely, extraneous graphics could be detrimental to test performance, particularly for students with visual impairments. During testing, students would spend testing time listening to spoken descriptions of the graphics or reading about the graphics in Braille. On high-stakes tests like the ACT, students are likely accustomed to having all provided information be relevant and necessary. In future item development, truly extraneous graphics might be omitted in favor of culturally relevant graphics essential for understanding the item context and responding correctly. This may be easier to achieve when creating items from scratch rather than modifying existing items. The Temple of the Jaguar is shown in Figure 1 as an example of an essential culturally relevant graphic.

**Figure 1.** Mural Painting in the Style of Portuguese Street Artist Huariu, a Dashiki, and the Temple of the Jaguar



### Word Choice for Comprehension

Being comprehensible to all test takers is one desirable property of culturally relevant assessment items. Upon further reflection, the science developers recognized that their existing passages and items sometimes included scientific and technical language that was not necessary for understanding the context and was, therefore, a potential source of construct-irrelevant difficulty for some students. According to the developers, their "goal was to make the items more comprehensible and accessible to more students." Considering that "the language of science has been controlled by White males," they did not want scientific and technical language to "get in the way," especially if teachers might not model the language of science in the classroom.

In the monarch butterflies unit, for example, the phrase "store (accumulate) body lipids" was changed in the culturally relevant version to "store fats," and the terms "overwinter" and "overwintering period" were replaced with "spend the winter" and "winter," respectively. Moreover, the homonym "study" was replaced by "experiment." In the worm castings unit,

"vermicompost (feces from earthworms)" was changed to "worm castings (earthworm feces)." The sound attenuation unit included an additional sentence, in relatively plain language, to explain the scientific context ("This results in the noise intensity level decreasing as the distance from the sound source increases"). The developers also tried to break up long, complex sentences. For example, they described experimental procedures in a numbered list rather than a paragraph. These approaches can be applied when modifying existing content or developing new content. Note, however, that scientific or technical terms may be construct relevant when the item assesses specific science knowledge rather than science reasoning skills.

### Voice

As noted previously, student agents were added to science units to help make the contexts more relatable and representative. With the addition of these agents, language in the science units was often changed from passive voice to active voice, and the science developers noticed a change in the tone of the passages. The worm castings unit provides an illustrative example. In that unit, the text changed from "Two experiments were performed. . . . Six different combinations of worm castings and potting soil were prepared to put into pots for the plants to grow in" to "The students performed 2 experiments. . . . The students prepared 6 different combinations of worm castings and potting soil." The science developers noted that "passive voice is the language of dominant culture in science." They said that it was sometimes easier to write shorter text in the passive voice because it was not necessary to attribute actions to agents. However, they asserted that "active voice makes it easier to read and adds empowerment and relatability."

### Social Justice

Highlighting social justice issues was a goal of item development, but it was not considered essential for cultural relevance. Thus, the number of items incorporating social justice was limited. Indeed, the math developers reported that it was easier to introduce cultural relevance than social justice issues. One math developer felt uncomfortable raising issues like the use of public spaces for demonstrations, and she wondered whether it was all right to create more opportunities for examinees to take offense by adding social justice content. For this research project, the answer was certainly "yes" because the study sought to elicit examinee reactions to such content. During item writing, the math developers reported internal reactions to draft items—even from non-White members of the development team—such as "You can't say that. It's offensive." However, when asked, "Does it offend you?" these people responded, "Of course not. It's real." At one point, the book-reading item used *Twelve Years a Slave* by Solomon Northup, but it was later changed to Colson Whitehead's *The Underground Railroad* because it was listed on several high school reading lists.

According to the science development team, a challenge of dealing with social justice issues is that you are highlighting inequities and you want to represent affected persons as empowered to address those inequities. The science team were satisfied with their integration of social justice (or environmental justice) into their units (e.g., noise pollution, food insecurity, effects of logging). However, cultural relevance seemed more difficult to achieve. Examples included farming traditional ingredients (tomatoes and chili peppers) and connecting monarch butterflies

to Día de los Muertos ("To many people, the return of the butterflies to Mexico symbolizes the return of their ancestors' souls"). Indeed, the external consultants did not always recognize cultural relevance in the science units. For example, communal vegetable gardens may not be culturally relevant (or generally relevant) to many students. This challenge was possibly exacerbated by the constraints of modifying an existing unit rather than creating a unit from scratch.

### *Scalability*

The content developers were asked to provide recommendations about what would be needed to develop culturally relevant passages and items on a larger scale. They expressed the need for ACT to hire more diverse content specialists who can write items based on what they already know and work with other developers who lack the requisite knowledge and lived experiences. Meanwhile, current content specialists need professional development and feedback like that provided during this research project. Participants in this project described it as "a really educational and fulfilling experience" and "a nice change of pace." The content developers suggested creating a repository of culturally relevant context ideas and tracking representation in developed items to ensure broad representation. Unintentionally, they developed more items with Hispanic/Latinx representation than Black for this project, perhaps because they started with more context ideas relevant to Latinx cultures.

The content developers noted the possible need to revisit style guidelines concerning the use of graphics. They requested guidelines for acceptable item length, and they suggested that editorial teams could help them control word counts. The math developers recognized that it would be more difficult to use automated item generation (AIG) to develop culturally relevant items because the math is uniquely tied to the context and cannot, therefore, be changed so easily. They also acknowledged that some culturally relevant items may be highly memorable because of their unique content. This could be a test security risk and warrant stricter controls on item exposure. Moreover, some culturally relevant content may have a greater risk of becoming "stale" over time (e.g., *Hamilton*), so it may be necessary to reevaluate relevance regularly. Note that many of these changes would require greater expense for testing organizations (e.g., new hiring, professional development, more time spent on item writing, less AIG, and less item exposure). However, should the results of the focus groups and subsequent research support the introduction of culturally relevant items, testing organizations should work to incorporate these types of items—despite cost implications—to help ensure that students without White lived experiences are not disadvantaged due to item content.

### *Summary of Lessons Learned*

This section describes lessons learned during the creation of culturally relevant passages and items.

- It is sometimes possible to modify existing content to make it culturally relevant, but authenticity often requires changing the math or science content or developing items from scratch.

- Word count is a challenge and a concern for those developing culturally relevant items; additional explanation is sometimes necessary to support comprehension and learning about cultures but is not necessary for the student to respond correctly to the item.

- Do not miss opportunities to represent people in positive, nonstereotypical ways, even when the item is already culturally relevant. This includes representing broad interest in other cultures (e.g., people of many cultures appreciating Black interests, arts, etc.).

- Culturally relevant items take longer to develop, in part because they require greater deliberation about the interpretation of words and phrases, but also because more time is needed to research topics and revise text to control word count.

- Graphics are a potentially powerful tool for supporting representation, but they pose accessibility challenges, especially for students with visual impairments.

- The change from passive voice to active voice and the removal of construct-irrelevant scientific and technical language promotes comprehension.

- Writing items that raise potentially sensitive social justice issues can be challenging, at least for White item developers.

- To support the scalability of developing culturally relevant items, testing organizations need to hire more diverse item developers, offer new professional development opportunities for item developers, and commit more time and resources to item development.

## Part 2: Focus Groups

This section reports the results of focus groups in which high school students examined the culturally relevant items alongside traditional (non–culturally relevant) items that contained the same content (i.e., the same required skills, numbers, required operations, experimental results, etc.). The main goals of the focus groups were to

- elicit participants' observations about the test questions,

- evaluate their reactions to culturally relevant content and the authenticity of the content,

- evaluate engagement and learning associated with the culturally relevant content,

- evaluate participants' perceptions of the effect of culturally relevant content on test performance, and

- elicit participants' preferences for culturally relevant or traditional content.

## Method

The study included eight virtual focus groups with White, Black, and Latinx students in 11th and 12th grades, each lasting approximately 60–90 minutes. One researcher facilitated each focus group, and a second researcher took notes. The focus group facilitator stayed on camera, but the supporting researcher stayed off camera to minimize distraction. Audio recordings of the focus groups allowed for detailed transcription and analysis. Each group of students examined five math item pairs and one longer science unit pair, with each culturally relevant item shown next to a matching traditional item. The items were randomly assigned to each group and counterbalanced in their display order, with math items always shown first and the science unit at the end. The items in each pair measured the same knowledge and skills and, in most cases, had identical answer choices. As indicated in Tables 3 and 4, some items were excluded from the focus groups, though they may be used in future projects. Items were selected to represent focus group participants and to best reflect the goals of culturally relevant assessment.

Students were prompted to answer questions about each pair of items, and no mention of culturally relevant content was made at this point. Specifically, the facilitator asked the students to reflect on the following set of questions after viewing each item pair:

- What do you think about these test questions?

- Is there anything specific that you like?

- Is there anything you would like to be different?

- What differences do you notice between the test questions?

After the students saw all the item pairs, the facilitator asked them to reflect on the following questions about all the items:

- How did you feel about the representations of people and cultures in these test questions?

- Were the representations of people and cultures accurate in your experience? Did you relate to any of the scenarios?

- Was there anything in the test questions that elicited strong feelings from you, either positive or negative?

- Do you think representation of people and cultures matters on a standardized test like the ACT? Why or why not?

- Were the questions with cultural representations more or less interesting compared to the traditional questions? Or were they equally interesting?

- Did you learn anything from reading the questions with cultural representations?

- Do you think you would perform better, worse, or the same on the culturally relevant test questions compared to the traditional questions?

- Do you think it would take more, less, or the same time to answer the culturally relevant test questions compared to the traditional questions?

- Now that you have seen the traditional and culturally relevant versions of the test questions, which ones would you prefer to see on the ACT? Why?

When the focus groups were complete, a third researcher created a full transcript of each session from the recording. The transcripts were then deidentified so they could be shared among the research team while preserving students' confidentiality. With finalized transcripts, two researchers independently analyzed the transcript data using both content and thematic analyses. After both researchers independently extracted themes, they reconciled discrepancies and reached a consensus by discussing the students' responses.

### *Participants*

An initial attempt to recruit participants via social media proved unproductive due to volunteers misrepresenting themselves. With help from our contacts in secondary education, however, we were subsequently able to recruit a geographically diverse sample of participants. The final group of 20 participants represented Alabama (5), Florida (1), Georgia (2), Iowa (1), Michigan (3), Mississippi (2), Nevada (4), North Carolina (1), and Texas (1). Eleven participants (55%) were in 11th grade, and nine (45%) were in 12th grade. Seven students identified as Black (35%), seven identified as White Hispanic/Latinx (35%), and six identified as White (30%). Fourteen students (70%) identified as female, five identified as male (25%), and one identified as a transgender male.

We originally planned seven groups of three to four students each (some homogeneous with regard to race/ethnicity, some heterogeneous). However, given the number of commitments, school and otherwise, that high school students have, we occasionally struggled with students committing and joining the meeting at the agreed-upon time. Therefore, the 20 participants were distributed across eight groups as follows:

- Group 1: Three Black 11th-grade students[1]

- Group 2: Four White 12th-grade students

- Group 3: Two Black 12th-grade students

- Group 4: One Latinx and two Black 11th-grade students

---

[1] Two of the students were revealed as residing outside the United States based on IP address. Because extracting themes was not possible with only one participant, this group was removed from thematic analyses, and only the participant from the U.S. was included in the content analyses.

- Group 5: One Black 12th-grade student and one White 11th-grade student

- Group 6: One Black, one Latinx, and one White student, two in 12th grade and one in 11th grade

- Group 7: Four Latinx 11th-grade students

- Group 8: One Latinx 11th-grade student

## Focus Group Themes

Thematic data analysis revealed several themes present in all focus group discussions and a few more themes present in several discussions. The following themes were present in all focus groups:

- item clarity and comprehension

- amount of information

- representation of cultures

The following themes appeared in some but not all focus groups:

- testing time

- real-world examples

In the following sections, student quotes illustrate the themes that emerged from the focus group discussions. Because these items were developed with cultural relevance to Black and Latinx students in mind, the students' racial/ethnic identifications are included with their quotes.

### *Item Clarity and Comprehension*

Across all focus groups, students expressed a preference for items with clear stems that made it easy for them to understand how the item should be solved. As one 12th-grade Black student said, "I prefer Version A [traditional item], just because it was simple, straight to the point, and I knew what I needed to do." In some cases, students felt that the additional contextual information provided in the culturally relevant items interfered with their ability to answer the question. For example, another Black 12th grader said this regarding the ofrenda item:

> Before I even read B [culturally relevant item], I just looked at it and knew it had so much unnecessary wording. It just makes my head hurt almost because I know it could be so much simpler and so much easier than what they made it.

A White 11th grader had this to say about the Gallons of Punch traditional item:

I mean, especially in Version A [traditional item], I like the clarity of the question. Like, it's very obvious what the question is asking about and the unit it's asking about, which I think is really important when you're taking the test. Because if you have a question that's vague or something, it's not as easy to answer. But I feel like on this one, it's really clear what the variables stand for and also what you're solving for, so I don't feel like I would be confused by it in any way.

Students also tended to note when pairs of items had similar levels of clarity. Regarding the Ecuadorean Birds pair of items, an 11th-grade Latinx student commented, "Very good wording. Yeah. And they're both like, straight to the point. You can tell what they're talking about and what you're trying to like, multiply, divide, you know, that kind of stuff."

Comments related to clarity and comprehension varied between items. Among items that showed an obvious difference in length, such as the Gallons of Punch items, where the culturally relevant item was significantly longer, students were much more likely to prefer the shorter, traditional item. However, among items with similar word counts and question formats, such as the Temple of the Jaguar items, preferences tended to be more evenly split between the culturally relevant and traditional items.

### *Amount of Information*

The amount of information theme had to do with how much "extraneous" information was present in the items, as students often preferred succinct items. As one 12th-grade Black student put it, "I like Version B [traditional item] because it flowed a lot better; it was easier to digest, and it gives me all the information I need without having to excessively read." If students perceived the contextual information as unnecessary, they were more likely to say they were overwhelmed by the information or point out differences between the lengths of the items.

One 11th-grade Latinx student put it this way regarding Gallons of Punch:

> I said we all noticed that the difference was that Version B [culturally relevant item] has way more context and background—because they both use the same formula, but they just introduced the problem differently. If I were to see that problem, I would like probably skip it, because I would think it would take me too much time, because it's like too long, and I wouldn't want to spend too much time reading. So I think it, like, it's more intimidating the fact that it's lengthier.

Another 11th-grade Latinx student had this to say about both items in the Underground Railroad pair:

> Because on Version A [traditional item], for example, or like Version B [culturally relevant item], for both of them literally the first half of the sentence, like the first sentence, half of it is completely useless, and it's like, it doesn't have to be there. And like, it throws me off. Like, yeah, I don't know. I just feel like there's a bunch of useless stuff and it's like having to read through, like stuff that doesn't really have anything to do with the problem.

When the contextual information was not perceived to be extraneous, students were more likely to say that the items were comparable. For example, this 12th-grade Latinx student noted similarities between the items in the Temple of the Jaguar pair:

> Both to me seem equal, like I don't see a lot of difference. It tells you what you're looking for, what it gives you. And I mean extra information doesn't bother me at this point because it's just like one little sentence at the beginning. Usually, I look for the question first, trying to see all the extra information in there. But I wouldn't mind either Version A or Version B. The diagram is there, so I don't have to spend time trying to draw it out. I know what I'm looking for, so both of them to me seem like equal.

### *Representation of Cultures*

After the students reflected on all the presented items, the facilitator asked them to think about how people and cultures were represented in the questions. In many groups, however, representation came up in discussion before this question, and multiple students noted appreciation at seeing representation in some of the items. While most students felt that the inclusion of the cultural information was positive or neutral, two students had some negative feelings related to at least one of the items; one of these students identified as Black, the other as Latinx. Students in all focus groups felt that representation of people and cultures was important, but in some cases, they differed on what this should look like on a test and whether it should be included in standardized tests.

When students saw their own cultures represented in the questions, many reported that it was a positive experience. One 11th-grade Latinx student put it this way:

> I really do like the representation. Because I feel like most people might not think that state testing is for them. Or just because like, when we're in class, we never learn about our cultures, and like in our books and our lessons, like, it's never talked about. So I feel like for somebody that's not as motivated taking the test going in, if they see those questions, I feel like that will change their whole experience because they feel like seen and they will feel like, like the test is for them and that they could also pass it.

In a similar vein, students felt that ACT was advocating for them and their cultures, as expressed by an 11th-grade Latinx student:

> Yes, because to me, it's just like, for that question to have made it onto the test, like obviously, there's someone up there pushing for my culture and my people. So, you know, in this case, it's you, thank you. So, it's like, it makes me very proud to know that there're like people up there with power that are like, advocating for my community.

Other students pointed out that while they liked seeing their cultural backgrounds represented, they felt that the items were not inclusive enough. Another 11th-grade Latinx student said the following regarding the culturally relevant ofrenda item:

I feel like there's a lot of Mexican representation; like, yeah, I'm Mexican, I appreciate it. But there's also other Hispanic cultures out there that, you know, there's also more Hispanic holidays that are not Mexican but are more like the southern, more below Mexico, like the different Latin countries, and I feel like they don't get represented enough.

A 12th-grade Black student also felt that the representation in these items did not go far enough:

I liked the inclusion of the, you know, different cultures, but I thought they could have done a whole lot better with including more than just, you know, Hispanic culture and all that. Like go as far as, you know, Italian, Arab, you know, just continue the representation of others.

Some students felt that the addition of cultural information could help students become more interested in the test or could provide context to help students answer the question. Regarding whether representation of people and cultures matters on a test, one 11th-grade White student said the following:

With me, my first, like my gut reaction is to say no, I don't really think it matters, because like, at least with me, I'm not really focused on the, like, the fluff of the question; I'm just trying to get through it so I can get the right answer/move on to the next one. But when I think about it more, I do think it kind of matters with representation of people and cultures because it kind of provides us with some kind of connection again like, "Oh, that's something I've seen or something I could see." It kind of highlights the situation that the question is trying to pose, like in either my mind or someone else's mind. That makes it easier to picture what the question is trying to ask and might make it a bit easier to answer the question.

In another focus group, when responding to the same question, an 11th-grade Black student noted the following:

I think it does matter because some of these people may not have known [about the culture]. And also it could interest more people in the test. So, like, they'll actually look forward to answering it because sometimes they come in these tests and they're like, dreading it. But if they read something they're interested in or something new, they're like, "Oh, that's interesting."

In the focus group comprising all White students, all students agreed the representation was important but felt that changing the names of students in the example problems would be enough to help represent culture. One 12th-grade White student said the following:

It's good to see yourself represented, but again, that could be done with just switching up the names. I personally don't think I would have gotten very confused by the extra facts in the questions, but I don't know how relevant or important it is to it.

Finally, another sentiment shared by a few students in different groups was that cultural representation is important in life but not on a test like the ACT. In one group, students were asked where representation would be important if not on the ACT. A 12th-grade Black student said this in response:

> I think me personally, everywhere. I think that people being people, feel like they belong, and being included is the most important feeling. Because I mean, I've grown up and pretty sure you know, go to elementary school, you want to get included and you want people to care for you and to see you as an equal. And then they, you know, obviously the ACT is the ACT. But every day I spent the life with teachers, with friends, with family members, you know, you just want to feel included. So I think pretty much every other aspect of life matters to be included, you know, besides ACT.

### *Testing Time*

The theme of testing time came up in five focus groups. Given the high-stakes nature of the ACT test, students expressed concern about the limited time available to them in each section. They generally felt that having shorter items would allow them to save time on the test and therefore felt more comfortable with shorter items. As one White 12th-grade student noted, "I prefer the shorter ones that get me to the point faster because then I can just move on to the next question quicker." A 12th-grade Black student in another group said, "I just prefer A [the traditional item]; just knowing the environment and the small, limited amount of time, for my mental sanity it's better to see something shorter and know what I have to work with."

Some students felt that a few of the items had a balance between short text and cultural information and said that, given the choice between items of the same length, they would prefer the culturally relevant items. For example, when asked which items they would prefer to see on the ACT, one White 11th-grade student replied as follows:

> Yeah, I mean if I was able to connect to all of them, or again, if they were all a happy medium of both the time, ease of reading, as well as cultural impact like the dashiki one, I would definitely prefer to see the culturally relevant ones. But as it stands right now, I think I would prefer to see just the normal types of questions on there, just from a pure time/test-taking standpoint where I'd want people to focus more on the actual problem with the test than the context of the problem.

Finally, some students suggested including culturally relevant information in other sections of the test where there are longer passages of text, which lends support to ongoing efforts by the English language arts content development teams to enhance cultural representation in English and reading passages. One 11th-grade Latinx student put it this way regarding interest in the culturally relevant items:

> So, [the culturally relevant items] were way more interesting, though, for sure they're more interesting, but like, in the timed math, it's not that it's not interesting, it's just that we don't have time. So I think there would be a little bit of like, wasted knowledge that

we could have got, that a student would have been more likely to read and understand in the English and science portion.

In another group, an 11th-grade White student said the following:

> But when I'm taking the ACT, if it's timed—if it's not timed, then I think that becomes a totally different question—but the ACT, as it is now, with the time, I would not want to see the *longer* culturally relevant versions of the test. I would rather take the standard test, with the math and science sections, especially. I think it becomes different when it's the reading and English sections; I think having culturally relevant information, I don't think that's an overreach, and I think that wouldn't slow people down if you included passages that taught people about different cultures but didn't take any extra time than a normal passage would. I think that that is definitely a perfectly reasonable way to include other people's cultures in the ACT without slowing people down.

### *Real-World Examples*

The theme of using real-world examples came up in four focus groups. Several students felt that real-world examples were preferable to generic examples, and, in their perception, the culturally relevant items were more likely to feature real-world examples. Students also felt that these were more relatable than generic examples. A 12th-grade Black student noted the following about the Temple of the Jaguar items:

> Version A [culturally relevant item] is about a real-world thing, and it—you could relate to it in a way; this could be a real-world problem that you're solving that you could actually do in work or something, and Version B [traditional item], it's more vague and just about a general tree and not anything specific.

In response to the science unit about butterflies, an 11th-grade Latinx student contrasted classroom knowledge with real-world, relatable knowledge:

> Both Version A [traditional item] and B [culturally relevant item] actually taught me more than what they were asking. Because the first one was more scientific, so I kind of learned like, what I would learn in the classroom, whereas like, Version B kind of helped me [pause] I feel like Version B I'll remember more outside of the classroom, and I'll remember it more like a real-world situation. Like real knowledge.

## Content Analysis Results

### *Item-Level Perceptions*

Although the focus group protocol did not specifically call for students to report their preference for one of the items in each pair, students frequently stated a preference during the conversation. Two researchers coded those responses, and the results are summarized in Tables 5–7, which break down students' preferences by race/ethnicity, gender, and grade level, respectively. As noted in the table descriptions, the number of students who examined each item varied.

**Table 5.** Distributions of Reported Preferences by Race/Ethnicity

| Item | N | Prefer CR | | | Prefer Non-CR | | | No Preference | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Black | Latinx | White | Black | Latinx | White | Black | Latinx | White |
| Gallons of Punch | 11 | 0 | 0 | 0 | 4 | 5 | 1 | 0 | 1 | 0 |
| Ecuadorean Birds | 12 | 1 | 2 | 0 | 2 | 3 | 0 | 2 | 1 | 1 |
| Hamilton Lyrics | 12 | 1 | 0 | 0 | 2 | 5 | 1 | 2 | 1 | 0 |
| Temple of the Jaguar | 9 | 1 | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 2 |
| Underground Railroad | 9 | 0 | 3 | 0 | 3 | 2 | 0 | 1 | 0 | 0 |
| Smithsonian Museum | 8 | 1 | 0 | 0 | 1 | 1 | 5 | 0 | 0 | 0 |
| Neighborhood Survey | 8 | 0 | 0 | 0 | 2 | 1 | 2 | 0 | 0 | 3 |
| Mural | 8 | 0 | 0 | 1 | 1 | 1 | 4 | 1 | 0 | 0 |
| Ofrenda | 12 | 2 | 3 | 0 | 3 | 2 | 1 | 0 | 1 | 0 |
| Shirt Sale | 8 | 2 | 0 | 1 | 0 | 1 | 4 | 0 | 0 | 0 |
| Butterflies | 12 | 2 | 5 | 0 | 1 | 0 | 1 | 2 | 1 | 0 |
| Worm Castings | 8 | 0 | 0 | 0 | 2 | 1 | 2 | 0 | 0 | 3 |
| % of Total | — | 8.5% | 12.0% | 3.4% | 17.9% | 19.7% | 19.7% | 6.8% | 4.3% | 7.7% |

**Table 6.** Distributions of Reported Preferences by Gender

| Item | N | Prefer CR | | Prefer Non-CR | | No Preference | |
|---|---|---|---|---|---|---|---|
| | | Female | Male | Female | Male | Female | Male |
| Gallons of Punch | 11 | 0 | 0 | 7 | 3 | 1 | 0 |
| Ecuadorean Birds | 12 | 2 | 1 | 4 | 1 | 3 | 1 |
| Hamilton Lyrics | 12 | 1 | 0 | 6 | 2 | 2 | 1 |
| Temple of the Jaguar | 9 | 1 | 3 | 3 | 0 | 2 | 0 |
| Underground Railroad | 9 | 3 | 0 | 4 | 1 | 0 | 1 |
| Smithsonian Museum | 8 | 0 | 1 | 5 | 2 | 0 | 0 |
| Neighborhood Survey | 8 | 0 | 0 | 3 | 2 | 2 | 1 |
| Mural | 8 | 1 | 0 | 4 | 2 | 0 | 1 |
| Ofrenda | 12 | 4 | 1 | 4 | 2 | 1 | 0 |
| Shirt Sale | 8 | 1 | 2 | 4 | 1 | 0 | 0 |
| Butterflies | 12 | 6 | 1 | 1 | 1 | 2 | 1 |
| Worm Castings | 8 | 0 | 0 | 3 | 2 | 2 | 1 |
| % of Total | — | 16.2% | 7.7% | 41.0% | 16.2% | 12.8% | 6.0% |

**Table 7.** Distributions of Reported Preferences by Grade

| Item | *N* | Prefer CR | | Prefer Non-CR | | No Preference | |
|---|---|---|---|---|---|---|---|
| | | 11 | 12 | 11 | 12 | 11 | 12 |
| Gallons of Punch | 11 | 0 | 0 | 6 | 4 | 1 | 0 |
| Ecuadorean Birds | 12 | 1 | 2 | 3 | 2 | 4 | 0 |
| Hamilton Lyrics | 12 | 0 | 1 | 5 | 3 | 3 | 0 |
| Temple of the Jaguar | 9 | 1 | 3 | 2 | 1 | 0 | 2 |
| Underground Railroad | 9 | 3 | 0 | 3 | 2 | 1 | 0 |
| Smithsonian Museum | 8 | 0 | 1 | 3 | 4 | 0 | 0 |
| Neighborhood Survey | 8 | 0 | 0 | 3 | 2 | 0 | 3 |
| Mural | 8 | 0 | 1 | 3 | 3 | 0 | 1 |
| Ofrenda | 12 | 5 | 0 | 2 | 4 | 1 | 0 |
| Shirt Sale | 8 | 2 | 1 | 1 | 4 | 0 | 0 |
| Butterflies | 12 | 7 | 0 | 1 | 1 | 0 | 3 |
| Worm Castings | 8 | 0 | 0 | 3 | 2 | 0 | 3 |
| % of Total | — | 16.2% | 7.7% | 29.9% | 27.4% | 8.5% | 10.3% |

Nearly 60% of the time, students indicated a preference for traditional (non-CR) items. Students who preferred the traditional items were more likely to be White, male, and in 12th grade. The preference for traditional items was not surprising given that the culturally relevant items tended to have more words, which caused students to perceive that the culturally relevant items would take longer to answer, and this was a concern for students given the stakes associated with a timed college admissions test like the ACT. However, for some items, we detected a tendency for students to prefer the culturally relevant version or to express no preference. For example, the culturally relevant versions of the Temple of the Jaguar math item and the Butterflies science item were more often preferred.

### *General Questions*

As mentioned previously, after the presentation of each pair of items, students were prompted to answer a few general questions about the accuracy and relatability of the representations of peoples and cultures, their feelings about the items, their interest in the items, and their perceptions of the impact on testing time and performance. The results from those questions are represented in Figures 1–7.

**Figure 1.** Response distributions for the question "Were the representations of people and cultures accurate in your experience? Did you relate to any of the scenarios?"
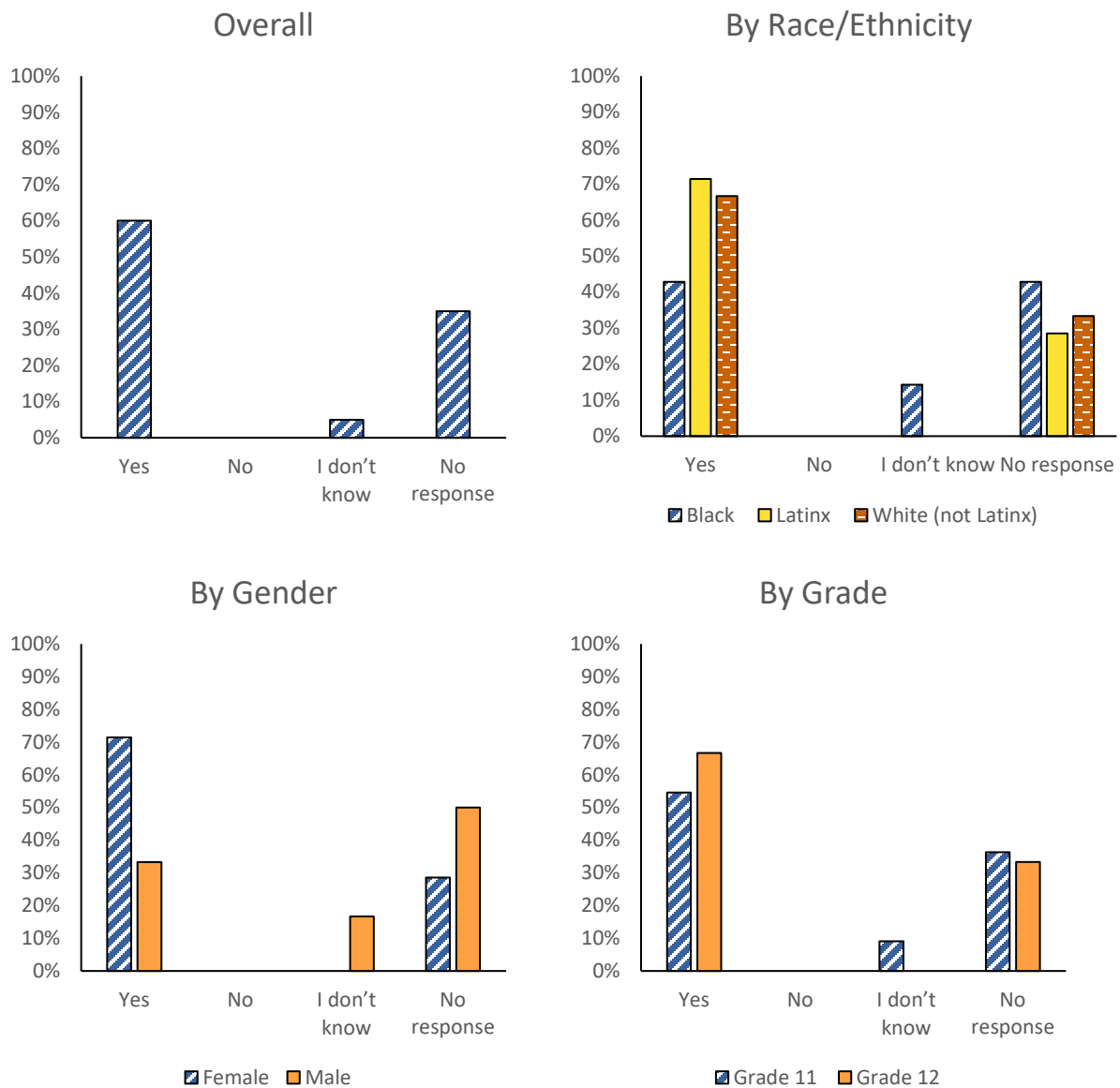


When asked about the accuracy of the representations of people and cultures in the culturally relevant items, a majority of students reported perceiving them to be accurate. Compared to White and Latinx students, Black students were somewhat less likely to report that the items had accurate representation. Female students were more likely than male students to report that the items had accurate representation.

Critics of culturally relevant assessment may be concerned about White students having negative feelings when exposed to item content that is not (allegedly) culture-neutral. Therefore, we asked students if the culturally relevant items elicited any emotions or feelings. A majority of

students either provided no response or reported that no feelings were elicited (11 out of 20), and many indicated that positive feelings were elicited (7 out of 20). Nearly all White students (5 out of 6) reported that no feelings were elicited, and a majority of Latinx students (4 out of 7) reported positive feelings. Negative feelings were mentioned by only 2 students. One Black student had a strong negative reaction to the mention of "students near the US-Mexico border," but she did not explain the reason. One Latinx student reacted negatively to every representation of people and cultures, explaining that when taking the ACT, she prefers to focus on the subject at hand and leave "all that other stuff" aside. Here are the student's own words:

> And another reason that I'm having more problems with Version A [culturally relevant item] is because, well, we are watching the news, right? We've all seen how, I guess, like Black Lives Matter and all of that stuff, and I'm not saying that I'm like against it, because you know, I do feel for the Black people. However, I feel like it's being put out there too much, that even for it to be focused on in these test questions. I don't think that's right either. Like, a test is a test, you know? It's not about what's going on in the world, it's about what they're teaching you to go out in the world.

**Figure 2.** Response distributions for the question "Was there anything in the test questions that elicited strong feelings from you, either positive or negative?"
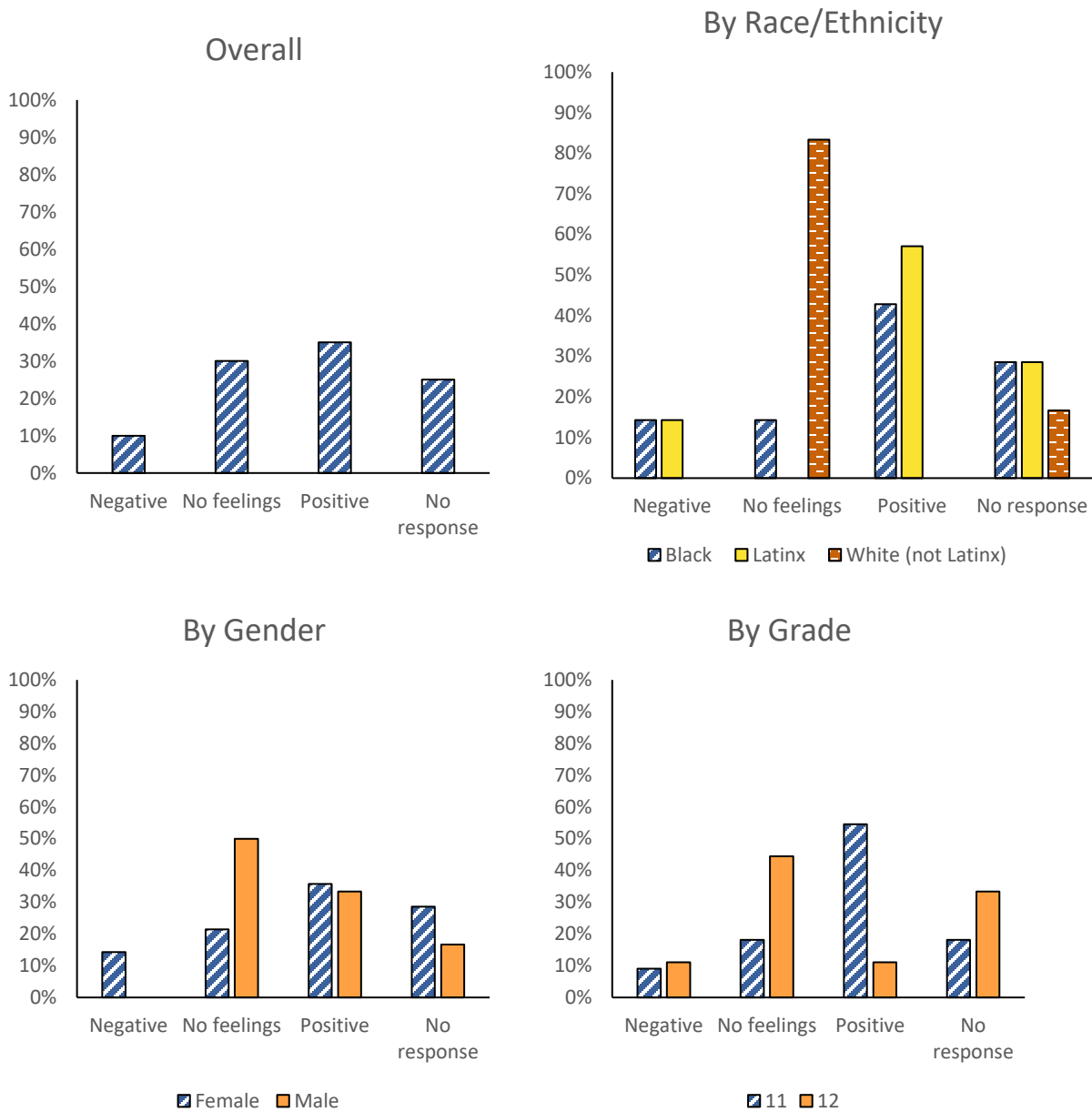


As shown in Figure 3, most students (15 out of 20) reported that the culturally relevant items were more interesting. This was especially true of Black students (6 out of 7), White students (6 out of 6), and 11th-grade students (7 out of 11). A lower percentage of Latinx students said that the culturally relevant items were more interesting (3 out of 7), though two Latinx students did not respond.

**Figure 3.** Response distributions for the question "Were the questions with cultural representations more or less interesting compared to the traditional questions? Or were they equally interesting?"
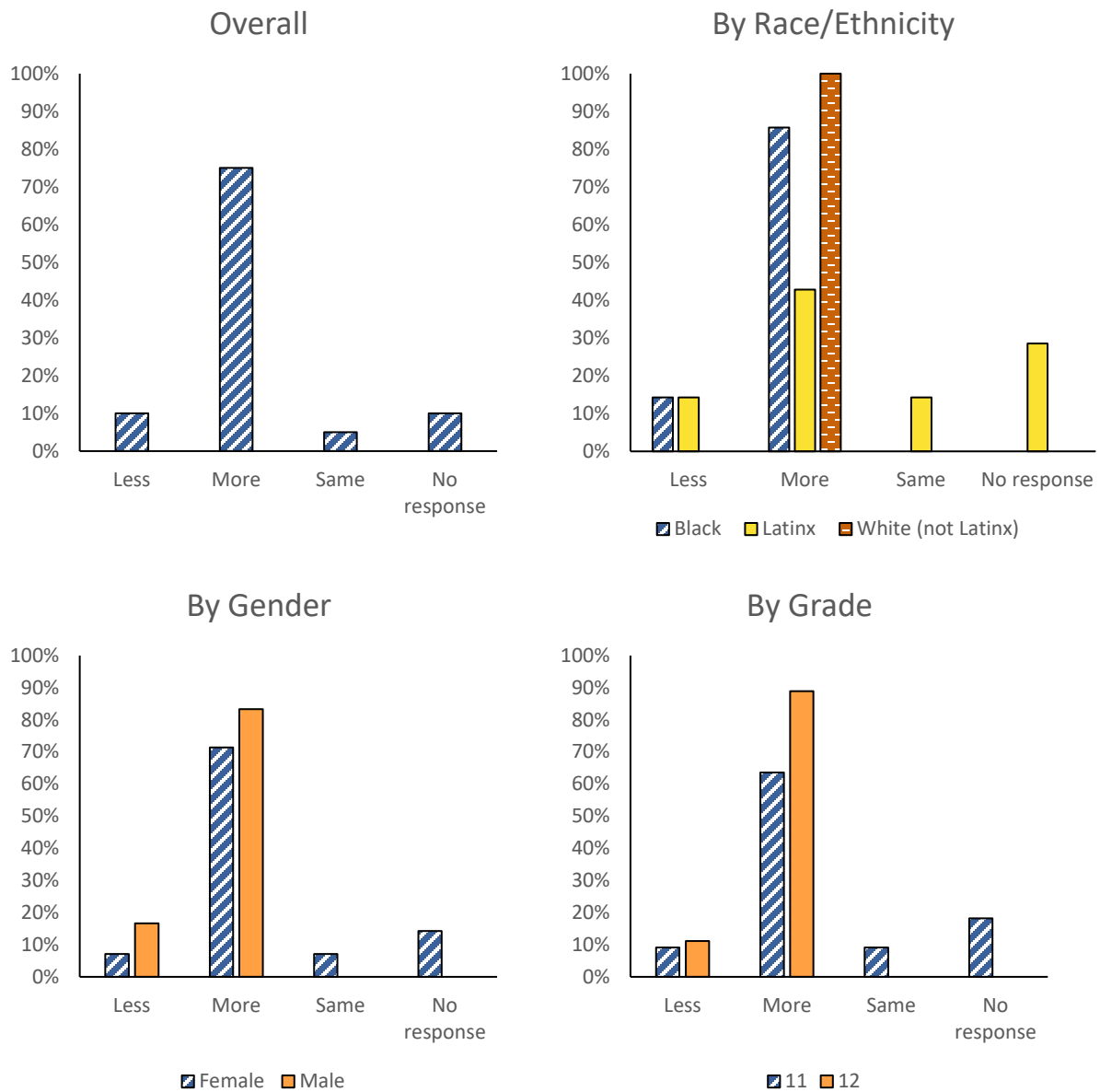


Another objective of culturally relevant assessment is to support learning about cultures. For that reason, we asked students if they learned anything from the culturally relevant items. As shown in Figure 4, most students (11 out of 20) reported having learned something. In contrast to our expectations, White students were least likely to state that they learned something (2 out of 6). Latinx students, on the other hand, mostly reported having learned something (5 out of 7). One possible explanation is that White students paid less attention to the problem context, as

indicated by comments like "I'm not really focused on the, like, the fluff of the question; I'm just trying to get through it so I can get the right answer/move on to the next one."

**Figure 4.** Response distributions for the question "Did you learn anything from reading the questions with cultural representations?"
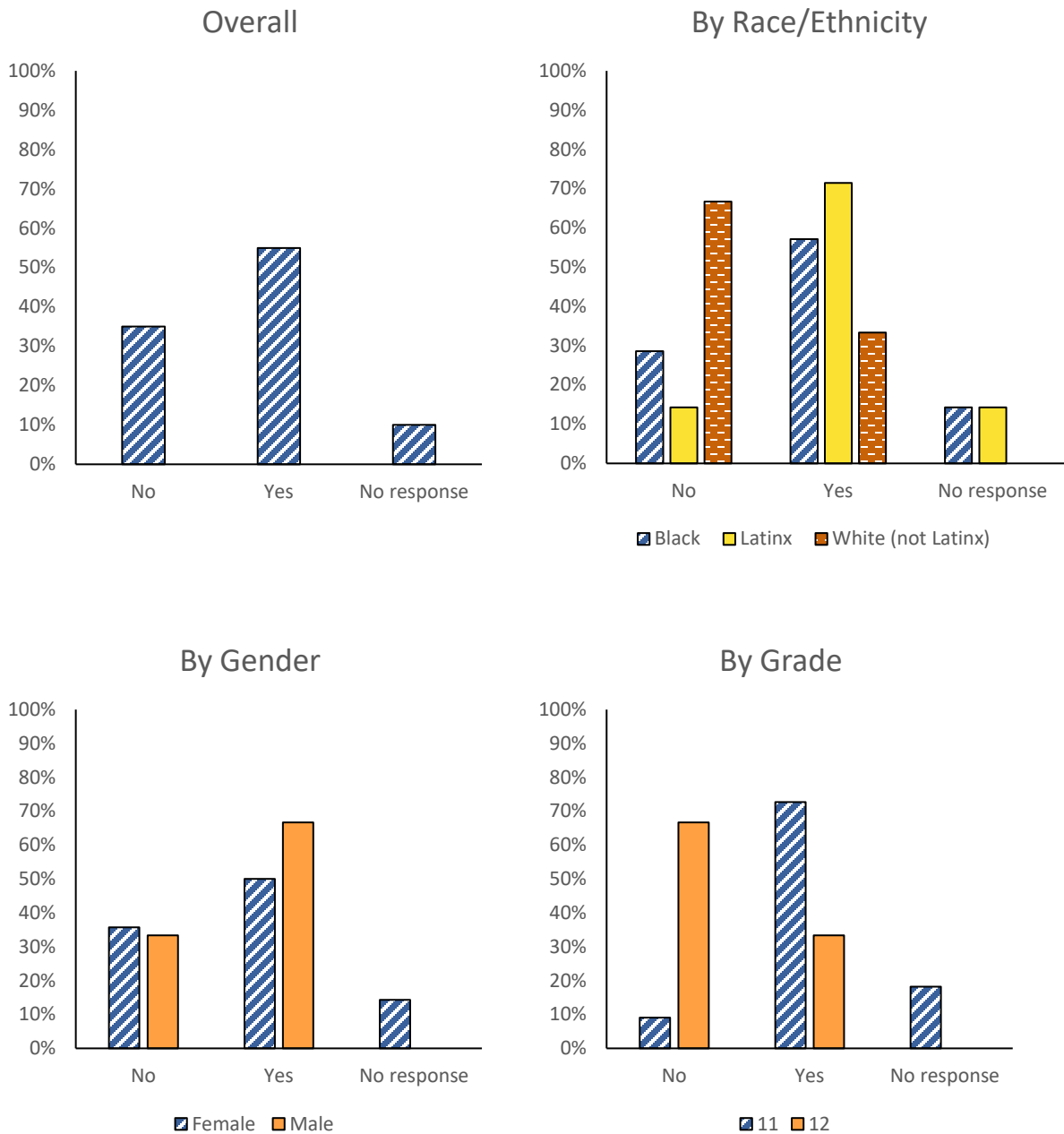


Most students (12 out of 20) said that they thought they would perform the same or better on culturally relevant items compared to non–culturally relevant items (Figure 5). When we consider racial/ethnic groups individually, none of the White participants reported that they

would perform better on culturally relevant items; half reported they would perform the same (3 out of 6). In contrast, none of the Latinx participants indicated that they would perform worse—most reported they would perform the same (2 out of 7) or better (3 out of 7). The greatest number of 12th graders thought they would perform the same on culturally relevant items (6 out of 9), while the greatest number of 11th graders thought they would perform better on culturally relevant items (5 out of 11). A possible explanation for this finding is that 12th graders are generally more engaged with the college admissions process, have more experience taking college admissions tests, and have only traditional items for practice. In contrast, few students in the fall of 11th grade have much experience with college admissions testing, and they have not yet experienced any of the related pressures to perform well.

**Figure 5.** Response distributions for the question "Do you think you would perform better, worse, or the same on the culturally relevant test questions compared to the traditional questions?"
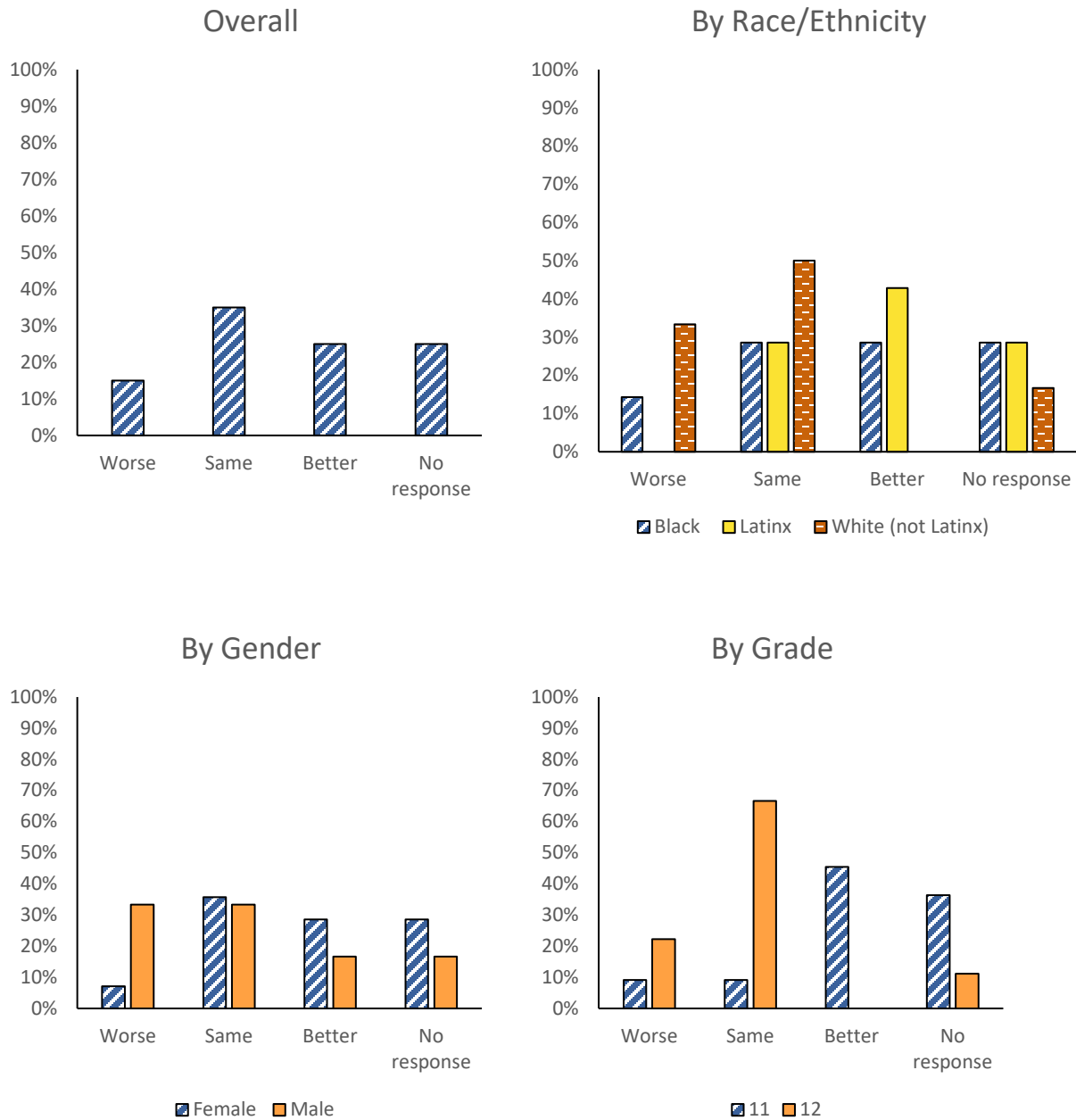


When asked if it would take more, less, or the same amount of time to answer the culturally relevant items (Figure 6), most students said they thought it would take longer (12 out of 20)—a perception shared by all six participating White students. Black students were almost evenly distributed between more time (4 out of 7) and the same time (3 out of 7). One Latinx student reported that it would take less time. Most 12th graders reported that it would take more time (8

out of 9), while 11th graders were mostly distributed between more time (4 out of 11) and the same time (5 out of 11). This finding is consistent with beliefs about performing worse, the same, or better on culturally relevant items (Figure 5). Again, this could be related to student proximity to the college admissions process and experience taking college admissions tests.

**Figure 6.** Response distributions for the question "Do you think it would take more, less, or the same time to answer the culturally relevant test questions compared to the traditional questions?"
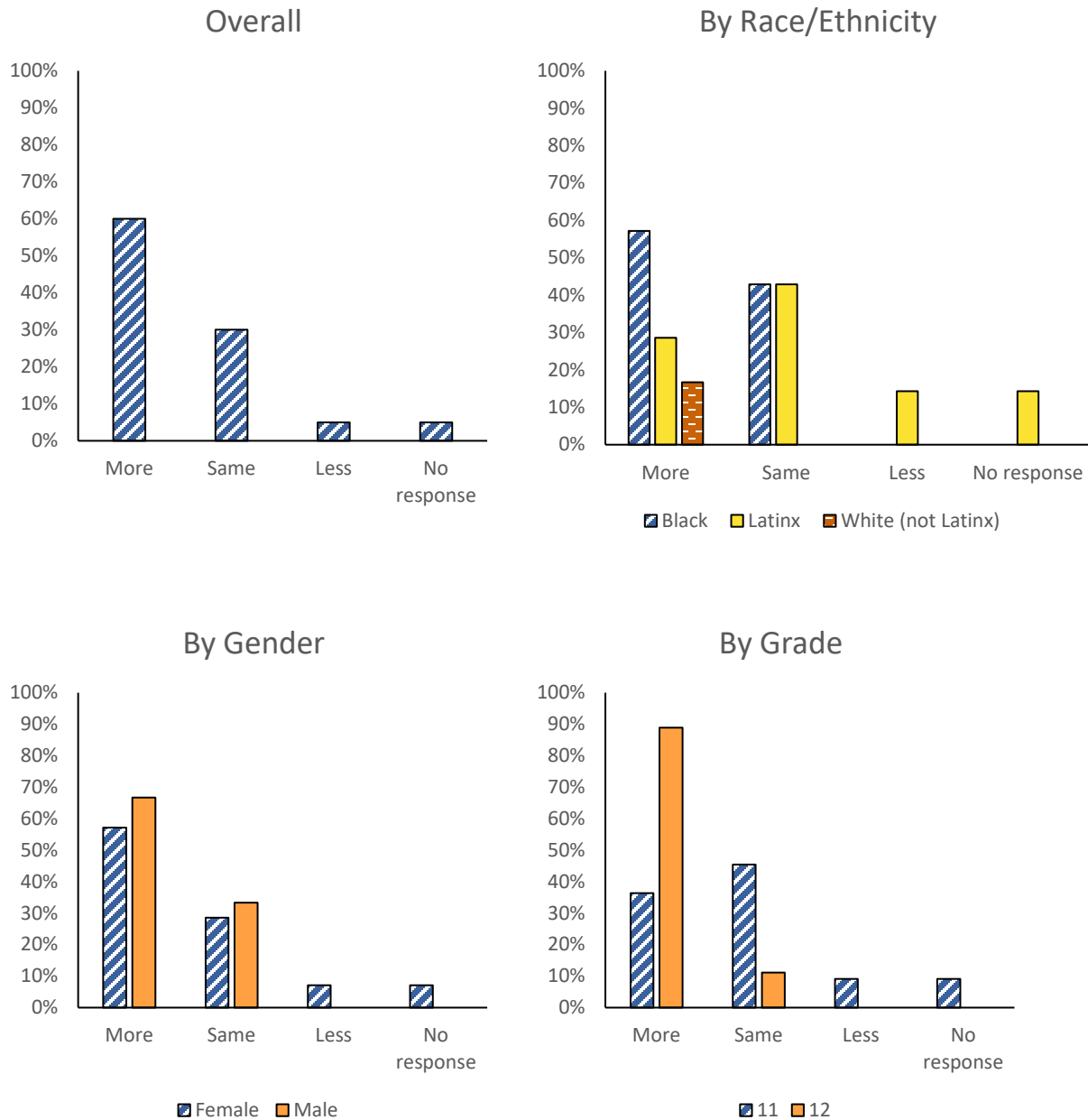
When asked to state their preference on item content type, half of the students reported preferring culturally relevant content (Figure 7). Half of the participating White students preferred traditional content (3 out of 6), and one third preferred a mix of both (2 out of 6). Most Black students (4 out of 7) and Latinx students (5 out of 7) preferred culturally relevant content. Participating 11th graders overwhelmingly preferred culturally relevant items (9 out of 11), but participating 12th graders indicated a stronger preference for traditional content (5 out of 9).

**Figure 7.** Response distributions for the question "Now that you have seen the traditional and culturally relevant versions of the test questions, which ones would you prefer to see on the ACT? Why?"
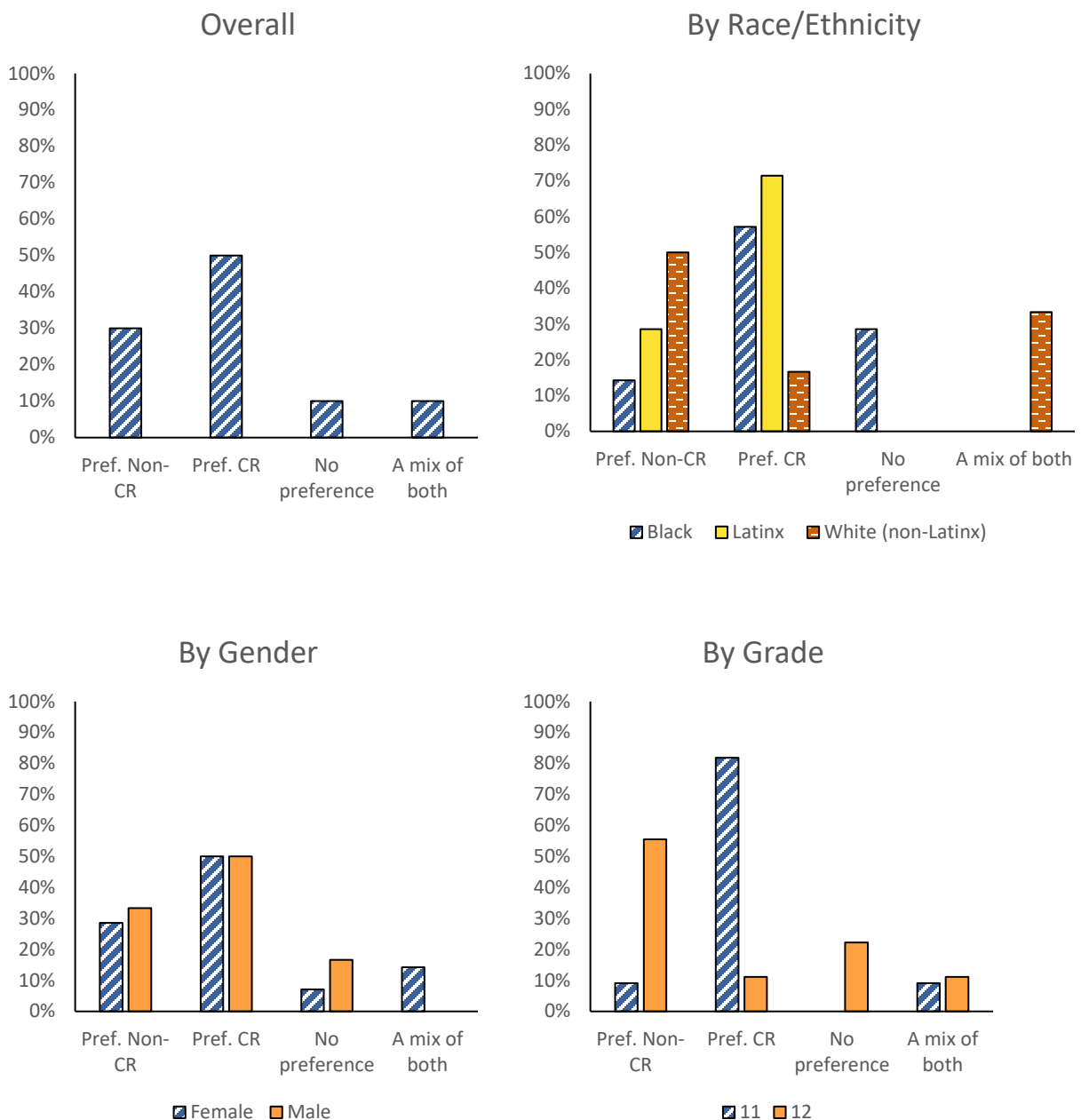
## Focus Groups Summary

Many focus group participants expressed interest in seeing culturally relevant content on the ACT. They reported finding this type of content more engaging than traditional content, and in many cases, students reported learning new things about other cultures. Moreover, most students (12 out of 20) felt that their performance on the ACT would improve or stay the same on the culturally relevant items. However, focus group comments indicated the potential to improve some of the item prototypes. The students found many of the prototypes too long, and they had concerns about the extra time needed to read longer questions. Additionally, in some items, students felt that the culturally relevant content was not well integrated with the math or science and that, as a result, it felt extraneous. Therefore, working with editorial specialists who may help content specialists better integrate the culturally relevant context while reducing word count seems like the logical next step in improving our development of culturally relevant items for the ACT.

## Commentary and Next Steps

Zaretta Hammond's (2015) framework for culturally responsive teaching lists "inclusive curriculum and assessment" as one of the four grounding principles. These principles focus on using differences as assets in a world of cultural and linguistic diversity. Hammond calls for a commitment to providing "resources written and developed by racially, culturally and linguistically diverse perspectives," in line with the efforts represented by this research. As Rudine Sims Bishop (1990) explained, it is important for students to see themselves in the materials that serve as mirrors and windows. With the windows, students gain understanding and appreciation of people with experiences that differ from their own, and with the mirrors, they can learn more about people who look like them and have similar cultural practices, beliefs, and values. This line of work supports efforts that are aimed at systemic change and serve to increase justice, equity, and equality.

Culturally relevant assessment seeks to increase equity for success using test items that elevate different cultures and spread cultural knowledge. Based on the students' feedback, it is apparent that students are open to culturally relevant test items if the additional information does not detract from their ability to demonstrate their knowledge and skills during a timed test. When they viewed the culturally relevant assessment items, students who did not typically see themselves reflected in tests were able to engage with inclusive content. For example, a student who did not initially engage in discussion exhibited a striking change in affect and engagement after seeing items that connected to their cultural background. Another student shared the following with us after participating in a focus group:

> I would like to say that this experience was very empowering. And it makes me really happy to think about—that people are thinking about these questions and that these questions are being brought up and that there's going to be change happening, because even though we might be graduating and won't be taking the ACT anymore, I feel better knowing that students will have a way better experience taking the ACT than we did, and

> hopefully we can increase those scores. Yes, so even more of us can achieve a great score. So thank you for allowing us to have this empowering experience.

Content developers were challenged to keep item length under control while meeting the expectations of culturally relevant assessment content, and students expressed concerns about item length and test speededness during focus groups. Yet students expressed an interest in seeing this content on the ACT, though they would prefer that their concerns be addressed before such items are included on the test. To balance competing demands, item developers in the future could focus on achieving one major goal in each item, such as including positive representation that disrupts stereotypes, highlighting social justice issues, or promoting cultural learning. With careful item development and form-construction planning, all major objectives could be achieved across items in a test form, though rarely in a single item. The content developers will take these lessons into account in a follow-up study in which 16 culturally relevant math items and 16 paired traditional items will be administered to randomly equivalent groups through operational ACT field-testing. This will allow direct comparisons between the psychometric properties of the culturally relevant and traditional items (e.g., difficulty, discrimination, and differential item functioning).

# References

Bishop, R. S. (1990). Mirrors, windows, and sliding glass doors. *Perspectives: Choosing and Using Books for the Classroom*, *6*(3), ix–xi.

Evans, C. M. (2021, November 3). Are there limits to making large-scale standardized testing culturally responsive? *Center for Assessment.* https://www.nciea.org/blog/culturally-sensitive-relevant-responsive-and-sustaining-assessment/

Hammond, Z. (2015). *Culturally responsive teaching and the brain: Promoting authentic engagement and rigor among culturally and linguistically diverse students*. Corwin Press.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice*, *31*(2), 132–141. https://doi.org/10.1080/00405849209543534

Montenegro, E., & Jankowski, N. A. (2020). *A new decade for assessment: Embedding equity into assessment praxis* (Occasional Paper No. 42). University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA). https://www.learningoutcomesassessment.org/wp-content/uploads/2020/01/A-New-Decade-for-Assessment.pdf

NCME. (2022a, February 17). *DEI in assessment practices- 2.17.22- All assessments* [Video]. YouTube. https://www.youtube.com/watch?v=oyaVWY6lWcM

NCME. (2022b, February 24). *DEI in assessments webinar- 2.24.22- Admission and selection assessments* [Video]. YouTube. https://www.youtube.com/watch?v=iIG4NGlkkOM

NCME. (2022c, March 10). *DEI in assessment practices- 3.10.22- K-12 assessments* [Video]. YouTube. https://www.youtube.com/watch?v=Bf80j0eUUOw&t=3286s

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*, *40*(4), 82–90. https://doi.org/10.1111/emip.12429

Randall, J., Slomp, D., Poe, M., & Oliveri, M. E. (2022). Disrupting white supremacy in assessment: Toward a justice-oriented, antiracist validity framework. *Educational Assessment*, *27*(2), 170–178. https://doi.org/10.1080/10627197.2022.2042682

Stadum, J. (2020, January). *Why cultural relevancy in science assessment matters* [Paper presentation]. 2020 Montana Office of Public Instruction Data and Assessment Conference.

# Acknowledgments

**ACT**®

## ABOUT ACT

ACT is a mission-driven, nonprofit organization dedicated to helping people achieve education and workplace success. Grounded in more than 60 years of research, ACT is a trusted leader in college and career readiness solutions. Each year, ACT serves millions of students, job seekers, schools, government agencies, and employers in the U.S. and around the world with learning resources, assessments, research, and credentials designed to help them succeed from elementary school through career.

For more information, visit act.org