Research Report
ETS RR–18-23

Assessing Civic and Intercultural Competency in Higher Education: The ETS *HEIghten*® Approach

Ou Lydia Liu

Katrina Crotts Roohr

Joseph A. Rios

December 2018

ETS Research Report Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

RESEARCH REPORT

Assessing Civic and Intercultural Competency in Higher Education: The ETS *HEIghten*® Approach

Ou Lydia Liu, Katrina Crotts Roohr, & Joseph A. Rios

Educational Testing Service, Princeton, NJ

Economic globalization and interdisciplinary advancements have increased the demand for college graduates to possess transferable skills that would allow them to contribute effectively to the modern workforce. In particular, transferable competencies such as civic competency and intercultural competency are critical for colleges to prepare responsible citizens and productive workers. Despite the recognized importance, the choice and quality of assessment for such competencies have been fairly limited due to the challenges in defining such complex, multidimensional constructs and identifying item types that can adequately assess them. In this report, we describe the principles we followed to operationalize definitions for civic competency and intercultural competency and the process we followed to design assessments for these 2 competencies. Findings from a large-scale pilot test are reported. Results showed that these multidimensional constructs can be adequately assessed and that there is room for students to improve in these areas. Implications for higher education institutions on how to promote these critical competencies are discussed.

Keywords *HEIghten*® Civic Competency and Engagement; HEIghten Intercultural Competency and Diversity; higher education; student learning outcomes; assessment

doi:10.1002/ets2.12212

Promoting Skills That Go Beyond Literacy and Numeracy

With increased national demand for a skilled and competent workforce, college graduates are expected, more than ever, to possess not only disciplinary skills, but also transferable skills that allow them to function effectively in a variety of settings in an increasingly globalized economy (Liu, 2017). Many institutions have made efforts to assess student learning of these transferrable skills to identify gaps and improve practices. According to an administrator survey conducted by the National Institute for Learning Outcomes Assessment (Kuh, Jankowski, Ikenberry, & Kinzie, 2014), over 80% of the surveyed institutions have adopted learning outcomes and currently utilize multiple tools to assess such outcomes. Some of the most frequently cited learning outcomes include critical thinking, written communication, quantitative literacy, oral communication, civic competency, and intercultural competency (Hart Research Associates, 2016). The combination of skills emphasized by higher education reflects a shifting priority for college general education requirements to move beyond numeracy and literacy by focusing on skills that are essential to be successful upon graduating college. Many of these essential skills have been understudied, such as civic and intercultural competency. The need for the development of learners' civic capacity was recognized by education leaders more than half a century ago (Conant, 1945) and has been reinforced by the National Task Force on Civic Learning and Democratic Engagement (2012), which urged measurable standards in evaluating students' learning in the civic domain.

Increased globalization also demands the preparation of college graduates who are competent in cultural awareness and are sensitive to different perspectives, values, and beliefs (Northouse, 2010). Although these two concepts are often studied as independent of each other, civic and intercultural competencies are intertwined and should be considered holistically on American campuses, which are experiencing accelerated internationalization. In fact, many institutions stress the importance of these two competencies in a holistic manner in their mission statements, emphasizing that they inspire students to become civically engaged and promote global perspectives and awareness (Perry et al., 2016). In addition, the Open Doors data released by the Institute of International Education (2016) show that international students at U.S. colleges and universities surpassed 1 million for the first time in 2016 and that more than 313,000 U.S. students received academic credit for study abroad in 2014–2015. The influx of international students and the popularity of study

Corresponding author: O. L. Liu, E-mail: lliu@ets.org

abroad programs further propel the need for institutions to become an effective venue to support sustained and meaningful cultural interactions. Institutions have the ability to shape students' civic and intercultural beliefs and experiences in assisting students to become future leaders who respect fairness, diversity, and inclusion. Students' exposure to civic and intercultural experiences in college could have a wide-ranging impact on their lifelong development in these domains.

The importance of civic and intercultural competency continues beyond the college campus as graduates enter the workforce, which is heavily influenced by multinational perspectives and fully immerses individuals in a society where people from vastly different ethnic, racial, and cultural backgrounds coexist. According to the Hart Research Associates (2015) survey of employers, close to 90% of the surveyed employers believed that all students should understand democratic institutions and values. The 21st-century workforce requires individuals who understand the rapidly changing social, cultural, and ethical issues and can anticipate the impact of such changes on their work (Casner-Lotto & Barrington, 2006; Gould, 2011).

Challenges in Assessing Civic and Intercultural Competencies

Despite the widely recognized importance of civic and intercultural competency, the quality of assessment of these two critical concepts has been subpar, compared to the assessment of more traditional learning outcomes, such as writing and quantitative literacy. One reason underlying the challenge in developing assessment for the civic and intercultural competency is the multidimensional nature of both concepts and therefore a lack of a clear-cut definition.

The reviews by Torney-Purta, Cabrera, Roohr, Liu, and Rios (2015) and Griffith, Wolfeld, Armon, Rios, and Liu (2016) identified many terms that have been used to describe these two competencies, respectively, without a clear understanding of the similarities and differences among such terms. For example, for civic-related concepts, common terms include *civic learning*, *civic capacity*, *civic education*, *civic skills*, and *civic inclinations* (e.g., Markle, Brenneman, Jackson, Burrus, & Robbins, 2013). Similarly, an array of terms has also been used to describe the intercultural domain, such as *intercultural attitudes and worldviews*, *intercultural capabilities*, *cross-cultural competency*, and *intercultural traits* (Leung, Ang, & Tan, 2014). Compared with other terms, the term *intercultural competency* emphasizes that the contexts and settings of intercultural interactions do not necessarily cross national borders and that they can also transpire at a micro-level (e.g., a state, a company, an institution; Griffith et al., 2016). The Torney-Purta et al. (2015) report reviewed more than a dozen influential frameworks and initiatives that were designed to promote civic education and revealed that although some common themes emerged from the various efforts, they each have some unique focus that is particularly incorporated for the purpose of that project. Similarly, Griffith et al. (2016) examined more than 20 models of intercultural competency and found these models differ vastly in their focus, the specificity of their definition, and the number of dimensions and components that are included for the intercultural construct.

Hence, before delving into the assessment design, it is incumbent on test developers to delineate the characteristics, components, and interrelations of the various terms to understand the implications of such factors in assessment design. An operational definition is also needed to guide the test development and provide transparency to users and other stakeholders on how scores can be interpreted and used in relation to the definition.

Another prevalent challenge in the assessment of civic and intercultural competency is the overreliance on self-report Likert scales within existing assessments. Among the close to 30 existing civic assessments and tools reviewed by Torney-Purta et al. (2015), the vast majority rely solely or partly on Likert scales. The issue is even more dominant with intercultural assessments, as only a few among the 32 existing assessments reviewed by Griffith et al. (2016) utilized an item format other than Likert scales. The predominant use of Likert scales limits the scope of assessment of civic and intercultural competency, as Likert scales can be useful for assessing certain affective-behavioral domains but may not be suited for assessing other essential components, such as knowledge and skills. Additional item types are needed to capture the full range of the civic and intercultural constructs.

In addition to the aforementioned issues in ambiguous definitions and limited item types, the psychometric quality of the existing assessments is also mixed and can be improved. One issue is that, given the multidimensionality of the civic and intercultural competency, many assessments report subscores. Although the reliability for the overall instrument is typically satisfactory, some assessments had subscale reliabilities lower than conventionally accepted values (Griffith et al., 2016; Torney-Purta et al., 2015). Another pitfall is the lack of information of the interrelations among the subscales. If some of the subscales are highly correlated, the decision to report scores based on such subscales needs to be revisited, as the separate subscores may fail to provide unique information about test takers.

In response to the challenges in assessing civic and intercultural competencies, Educational Testing Service (ETS) launched a research program, now called the *HEIghten*® Outcomes Assessment Suite, to review current research on these two competencies, among others, and also to design assessments that are less susceptible to the typical challenges facing these two constructs. Note that HEIghten research also focused on critical thinking, quantitative literacy, and written communication, in addition to civic and intercultural competencies. A framework paper was published for each of the HEIghten modules, and the papers mainly consisted of the following four components: (a) a comprehensive overview of current frameworks and research initiatives for the competency of interest in both domestic and international arenas; (b) an overview of existing assessments, including the reliability and validity evidence of these instruments, produced by both the test sponsors and users; (c) an operational definition that provided guidance for new assessment development of the target competency; and (d) a discussion of assessment considerations, including assessment challenges, possible item types and item formats, and other significant issues to be considered in assessment design. Examples of prominent prior work reviewed for civic competency and engagement included the Political Engagement Project at the Carnegie Foundation for the Advancement of Teaching from 2000 to 2007 (Colby, Beaumont, Ehrlich, & Corngold, 2007; Colby, Ehrlich, Beaumont, & Stephens, 2003), the American Democracy Project sponsored by the American Association of State Colleges and Universities (2014), the Civic Engagement Valid Assessment of Learning in Undergraduate Education (VALUE) rubric (Rhodes, 2010) by the Association of American Colleges and Universities (AAC&U), and the global and civic learning component of the Degree Qualifications Profile (Adelman, Ewell, Gaston, & Schneider, 2011, 2014) sponsored by the Lumina Foundation. Similarly, a number of prior definitions and frameworks for intercultural competency were also consulted (Arasaratnam, 2008; Deardorff, 2006; Griffith & Harvey, 2001; King & Baxter Magolda, 2005). For more details of the framework papers, refer to Torney-Purta et al. (2015) for civic competency and engagement and to Griffith et al. (2016) for intercultural competency.

Civic Competency and Engagement and Intercultural Competency and Diversity: The HEIghten Definitions

Based on a comprehensive review and alignment with existing literature, researchers proposed the following operational definitions for civic competency and engagement (CCE; Figure 1) and intercultural competency and diversity (ICD; Figure 2). The HEIghten CCE assessment encompasses two competency or cognitive dimensions of civic knowledge and civic skills and two engagement dimensions that include civic attitudes and civic participation. This definition aims to provide a comprehensive depiction of what individuals know and can do when it comes to civic learning and also what they believe and how they participate in civic processes and activities. Civic knowledge includes knowledge of facts, concepts, and principles (e.g., democratic processes, government structures, voting) across various contexts (local, national, international, past or present). Civic skills includes both analytic skills (i.e., the ability to apply political and civic knowledge to systematically analyze civic-related issues/scenarios) and participatory and involvement skills (i.e., the ability to make reasoned judgments about political and civic situations or problem-solving processes, especially in group and/or community contexts). Civic attitudes includes efficacy (i.e., belief that one can understand and influence government and political affairs) and democratic norms and values (i.e., one's belief in basic principles of a democratic and diverse society, with a sense of responsibility to take civic action). Lastly, civic participation includes civic and political behavior and actions in face-to-face or online contexts.

For the HEIghten ICD assessment, we draw on a prior definition that intercultural competency "reflects a person's capability to gather, interpret, and act upon these radically different cues to function effectively across cultural settings or in a multicultural situation" (Earley & Peterson, 2004, p. 105). Intercultural competency is decomposed into three dimensions: *approach*, *analyze*, and *act* (Figure 2). The approach dimension represents the social and emotional characteristics that will likely impact a person's overall effectiveness in an intercultural setting. The approach dimension captures (a) having a positive cultural orientation (i.e., evaluating cross-cultural situations in a positive and favorable way), (b) tolerance for ambiguity (i.e., the ability to maintain composure and well-being in uncertain or volatile situations), and (c) cross-cultural self-efficacy (i.e., the degree to which an individual believes that he or she can achieve a goal).

The analyze dimension captures an individual's ability to become aware of, evaluate, and synthesize culturally related information without activating preconceived judgments or personal biases. The analyze dimension includes (a) self-awareness (i.e., the degree to which an individual understands the impact of his or her culture, values, preferences, and previous experiences on his or her cognitive, emotional, and behavioral responses), (b) social monitoring (i.e., attention
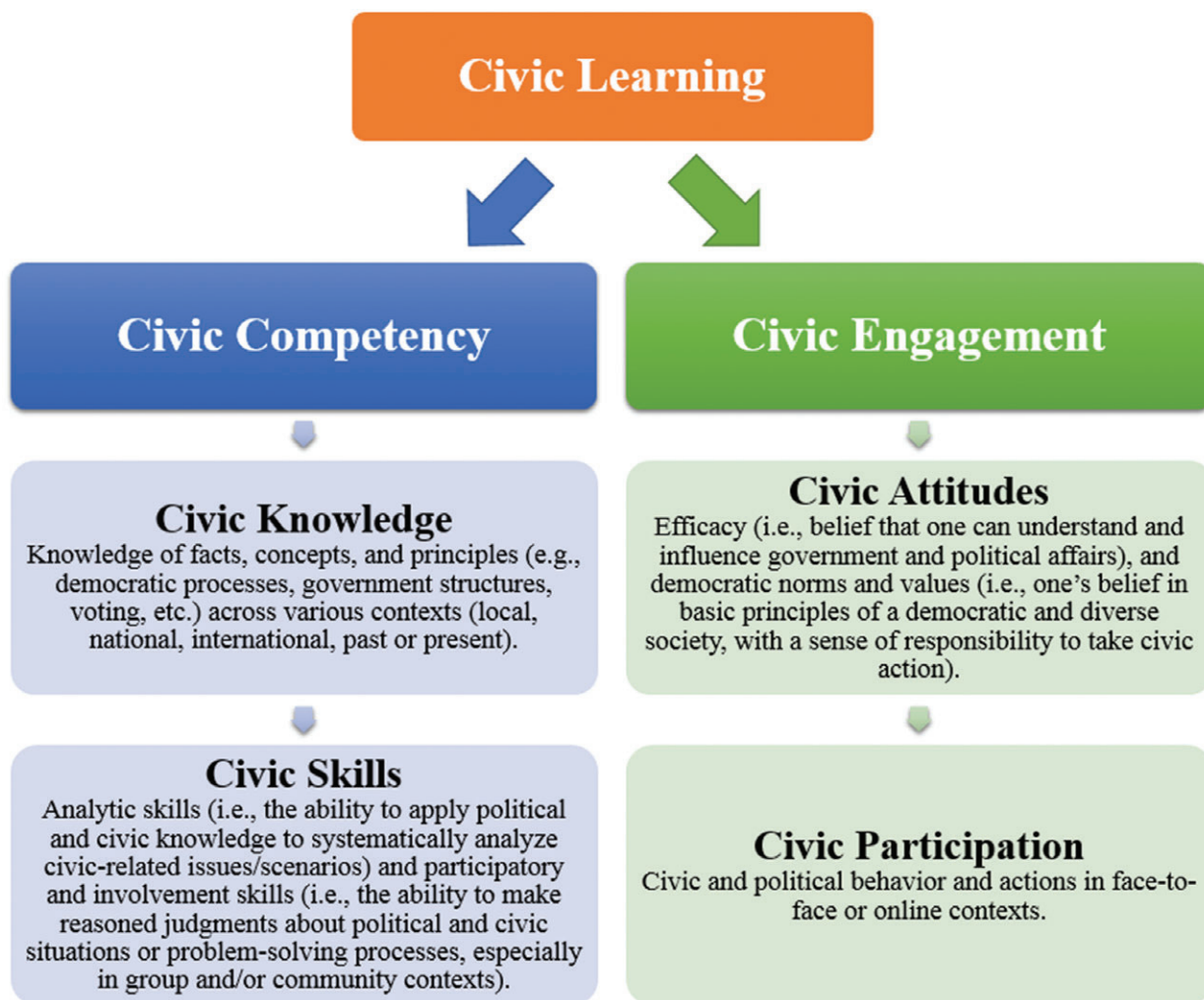
Figure 1 Framework and dimensions for the HEIghten civic competency and engagement assessment.

to the other's physical, verbal, and nonverbal behaviors and cues during a social interaction; attention to others' responses to one's own actions and signals), (c) perspective taking/suspending judgment (i.e., active consideration of others' potential viewpoints; active refrainment from preconceived cultural schema interfering with information processing), and (d) cultural knowledge application (i.e., utilization of relevant declarative cultural knowledge in an interaction).

Last, the act dimension refers to a person's ability to monitor and control his or her emotions and behaviors in intercultural interactions, particularly in difficult and stressful situations. The act dimension includes (a) behavior regulation (i.e., active monitoring and revision of personal behavior to engage in culturally appropriate behavior and avoid engaging in culturally inappropriate behavior) and (b) emotion regulation (i.e., the ability to monitor and revise emotions in an automatic or controlled manner).

Translating Definitions into Assessment: Test Development and a Large-Scale Pilot

The designs of the CCE and ICD assessments followed three guiding principles in that they need to (a) closely align with the theoretical framework and also the operational definition, (b) utilize a variety of item formats (other than just Likert items) to elicit a full range of information from students, and (c) situate the items in authentic scenarios that are relevant to higher education. To this end, the CCE assessment uses item formats including single-selection multiple-choice, multiple-selection multiple-choice, drop-down menus, and situational judgment (SJT) items to measure civic competency. There are 30 civic competency items: 15 measuring civic knowledge and 15 measuring civic skills. Note that SJT items are in a multiple-choice format (scored as right or wrong) and are one of the item types used to measure civic skills. SJTs present
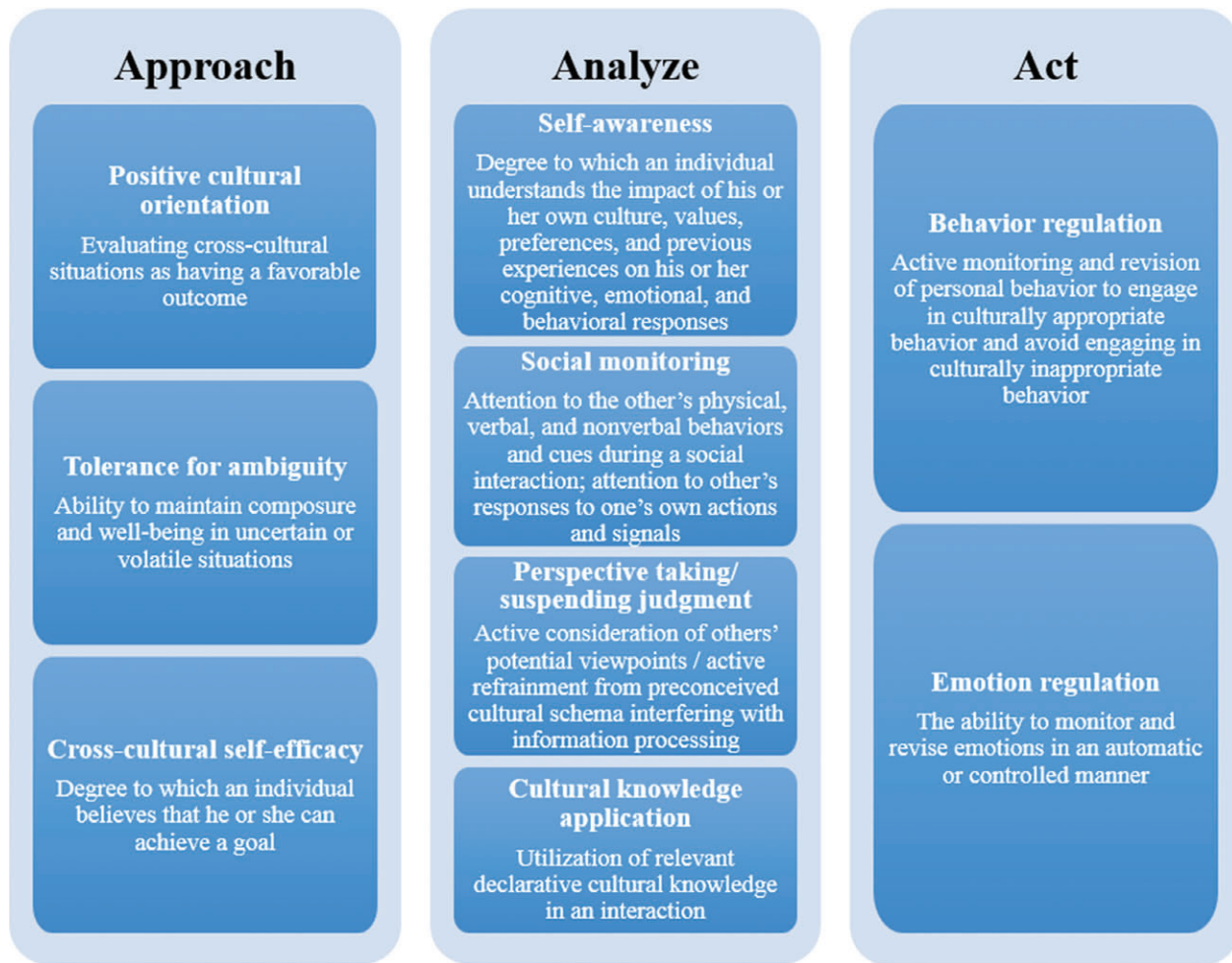
Figure 2 Framework and dimensions for the HEIghten intercultural competency and diversity assessment.

a test taker with a task-related situation where the test taker is asked to choose the appropriate response from a list of alternatives. In this assessment, these questions require students to analyze a set of circumstances and evaluate or choose the correct answer or course of action. Answer keys for the SJTs were determined by content experts and confirmed using empirical data from prototype administrations prior to the pilot test. For civic engagement, item formats include Likert-type items and multiple-selection multiple-choice items. For measuring civic attitudes, hypothetical scenarios about civic-related issues are presented, followed by reactions or quotations from people involved in the scenario. Following the scenarios are Likert-type questions where test takers are asked to rate their agreement with each quotation on a 4-point scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). There are 30 Likert-type questions measuring civic attitudes. The 20 civic participation items include questions about the context and frequency of the participation. All item stimuli across the assessment are embedded in realistic contexts (e.g., campus, local community, state, national, and international levels) and include a variety of maps, texts, graphs, tables, and charts to make the assessment more engaging for students.

The ICD assessment uses both Likert-type items and SJT items to assess students' intercultural attitudes and skills. For the approach domain, Likert-type items present test takers with a set of short statements with which they are asked to rate their agreement on a 4-point scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). There are 34 of these statements, and they are designed to be highly relatable for students. The analyze dimension has 26 items and the act dimension has 14 items, all in SJT format. These two dimensions are later combined for analysis because they are highly correlated (Table 5). Note that it does not necessarily invalidate the theoretical structure of the construct, but the high correlation makes it less practically meaningful to keep the dimensions separate. Test takers are presented with a range of cross-cultural scenarios and a series of questions relating to each scenario. Test takers are asked to indicate the best or most appropriate response

Table 1 Demographic Information of Participants From the Civic Competency and Engagement and Intercultural Competency and Diversity Pilot Study

Demographic information	CCE		ICD	
	N	%	*N*	%
Total sample	1,841		1,901	
Gender	1,617	87	1,706	90
Male	674	37	699	37
Female	926	50	983	52
Race/ethnicity	1,667	91	1,740	92
American Indian/Alaskan Native	9	< 1	10	< 1
Asian/Asian American	111	6	115	6
Black/African American	245	13	184	10
Hispanic/Latino	124	7	104	5
White	1,018	55	1,180	62
Other/multiple responses	130	7	91	5
Class status	1,820	99	1,896	99
Freshmen (<30 credit hours)	494	27	481	25
Sophomores (30–60 credit hours)	559	30	591	31
Juniors (60–90 credit hours)	418	23	453	24
Seniors (90+ credit hours)	349	19	371	20
College major	1,608	88	1,711	90
Business	332	18	418	22
Humanities	68	4	76	4
Social science	500	27	399	21
Natural science	438	24	437	23
Multimajor	270	15	380	20

Note. CCE = civic competency and engagement assessment; ICD = intercultural competency and diversity assessment.

or responses using single- or multiple-select multiple-choice formats, with the aim of creating or maintaining positive interactions and relationships in cross-cultural settings. The scenarios or situations are set in the following contexts: (a) study, teaching, or traveling abroad; (b) international collaboration; (c) guests from other cultures; and (d) subcultures within the United States.

To adequately assess the civic and intercultural constructs, a pilot study took place between January and March 2017. Four pilot forms were designed for the CCE assessment, and three pilot forms were designed for the ICD assessment. Each CCE form had 80 items, and each ICD form had 76 items. Each form was designed to be less than 45 minutes long so the assessment could be administered in one class period, following recommendations from an online survey including more than 200 academic officers at both 2- and 4-year higher education institutions as part of the assessment design process. Currently, both the CCE and ICD assessments have two operational forms in the same format as the pilot forms (i.e., same number of items and administration time).

In recruiting institutions for the pilot study, consideration was given to institutions' selectivity, type (e.g., community college, liberal arts college, research university), size and setting, and geographic location. Each institution may have used different strategies to recruit students (e.g., cash rewards, course credits, free food), which reflects the reality of how institutions typically recruit students to take similar assessments. A total of 1,841 students from 38 U.S. higher education institutions took the CCE assessment, and 1,901 students from 37 U.S. higher education institutions took the ICD assessment. Table 1 shows the demographic information of the participants. Note that the samples for these two assessments were not completely separate, as some institutions administered both assessments to the same group of students.

A host of psychometric analyses were conducted across the pilot forms to examine the quality of the test items, including dimensionality analyses, reliability,[1] item difficulty, item discrimination, differential item functioning analysis with regard to gender (an analysis typically used to detect potential bias),[2] correlations between subdimensions, and relationships of test scores to other related variables. Prior to conducting these analyses, we flagged and removed unmotivated test takers using two sets of rules, one for the multiple-choice sections of each assessment and one for the Likert-type sections. *Test-taking motivation*, which is defined as the willingness to engage as well as investing effort and persistence in working on test items (Baumert & Demmrich, 2001), has long been a concern for low-stakes assessments, as the results have no consequences for test takers. For the multiple-choice sections of the CCE and ICD assessments, test takers were

Table 2 Institution-Level Reliability Estimates Across the Civic Competency and Engagement and Intercultural Competency and Diversity Assessments

Assessment	Institution-level reliability[a]
Civic competency and engagement	
Civic competency	.95
Civic knowledge	.89
Civic skills	.92
Civic engagement	
Civic attitudes	.85
Intercultural competency and diversity	
Approach	.86
Analyze and act	.95
Cultural knowledge application	.81
Suspending judgment	.86
Self-awareness	.77
Social monitoring	.89
Emotional regulation	.91
Behavioral regulation	.82

[a]Institution-level reliability was calculated using a split-sample approach as described in Klein et al. (2009).

Table 3 Student-Level Reliability Estimates Across the Civic Competency and Engagement and Intercultural Competency and Diversity Assessments

Assessment	Student-level reliability[a]
Civic competency and engagement	
Civic competency	.73
Civic attitudes	.71
Intercultural competency and diversity	
Approach	.77
Analyze and act	.82

[a]Student-level reliability was calculated using coefficient alpha. These results are the average estimates across the pilot test forms.

flagged and removed if they (a) failed to answer more than 25% of items and/or (b) responded in less than 10% of the mean item response time for over 20% of items. For the Likert-type sections of the assessments, students were flagged and removed if they (a) possessed 75% or more of nonresponses and/or (b) completed the Likert section in less than 2 minutes. For the purpose of this report, we will not report on the full psychometric results. Instead, in this section, we present results on some noteworthy aspects of students' performance, the relationship between their test performance and relevant criteria, and future directions of research. For instance, Tables 2 and 3 show the institution-level and student-level reliabilities for the two assessments, respectively, for the score reporting categories. Score reporting categories for the CCE assessment include (a) civic competency with two subscores of civic knowledge and civic skills and (b) civic attitudes. The operational CCE assessment will not have a total score for civic participation given the heterogeneity of participation questions. Instead, item-level frequencies will be reported back to the institutions for them to monitor students' participation at various levels and in various contexts. Score reporting categories for the ICD assessment include (a) approach and (b) analyze and act, with six additional subscores for cultural knowledge application, suspending judgment, self-awareness, social monitoring, and emotional regulation.

A First Peek Into Students' Civic and Intercultural Competencies

For civic competency, students on average responded correctly to approximately half of the items on the assessment. The results suggest that there is room for improvement for both students' civic knowledge and their civic skills. Similarly, average civic attitude level was around 60%, meaning that students were likely or very likely to agree with most of the assertions about efficacy and democratic norms and values. With regard to students' civic participation, Figure 3 shows two items on civic participation that demonstrated areas of improvement for students across institutions. Question a asked students to select all the contexts where they had done the following: "Since you have been at this institution, have you collected money
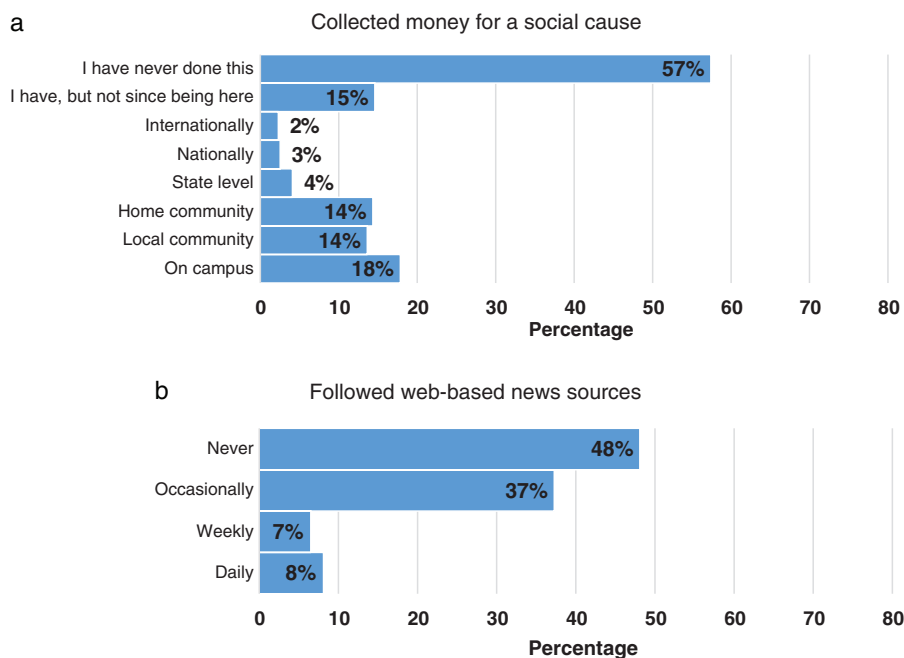
a
Collected money for a social cause

b
Followed web-based news sources

Figure 3 Student responses on civic participation items. Item a was a multiple-selection multiple-choice item, meaning that students could select more than one category. Item b was a Likert-type item where the students selected a single response. Note that the question and answer choices shown in this figure are abbreviated versions of the actual test questions and responses.

for a social cause?" Question b asked students the frequency at which they "followed social/political blogs, podcasts, or other Web-based news sources." Although many students had raised money for a social cause on campus, few students had raised money for state-level, national, or international causes. In addition, only 15% of the students reported following Web-based news sources on a daily or weekly basis, and almost half of students had never done that, which could suggest a need for students to increase their participation in terms of paying more attention to civic- and political-related issues.

For the ICD assessment, on the approach dimension, students on average scored 106.18 ($SD = 11.65$) points out of the possible 144 points, indicating no evidence of a ceiling effect. Overall, the results of the approach section revealed important aspects for students to improve. For example, when responding to the statement "I could adapt to living with a family from another culture," more than half of the students disagreed or strongly disagreed (Figure 4). Social desirable responding (i.e., choosing a particular answer to look good) was also examined via cognitive interviews with students during the prototype stage of the assessment. Specifically, when asked "If this [approach] section had counted for something—for a class grade, for credit—would you have answered differently?" students responded in the following ways: "I thought less about finding the correct answer and more about what I would do in the situation" and "There isn't a right or wrong answer, just what you would do." These results echo the findings of Rios and Anguiano-Carrasco (2016), who found that social desirable responding was negligible for low-stakes assessments comprising Likert-type items, such as the ICD assessment in this case. On the combined analyze and act dimensions involving multiple-choice items, students correctly responded to approximately 56% of the items, suggesting that this section of the assessment was neither too easy nor too difficult for students.

In addition to reporting these various scores back to institutions, proficiency levels with performance-level descriptors will also be made available to institutions after standard setting is conducted.

How Did Freshmen and Seniors Differ on Civic Competency and Engagement and Intercultural Competency and Diversity?

How students' civic and intercultural competencies evolve through college years is another aspect that many institutions care about. For these analyses, we used cross-sectional data with separate cohorts across freshman, sophomore, junior, and senior students. Across college grade level (as indicated by college credit hours completed), senior students outperformed their freshman peers on civic competency. Results show that sophomores, juniors, and seniors statistically significantly
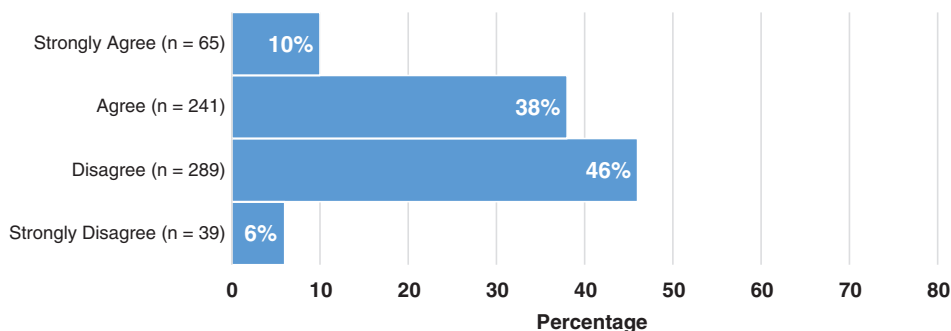
Figure 4 Student responses to "I could adapt to living with a family from another culture" on the approach section of the ICD assessment. This was a single-response Likert-type item.
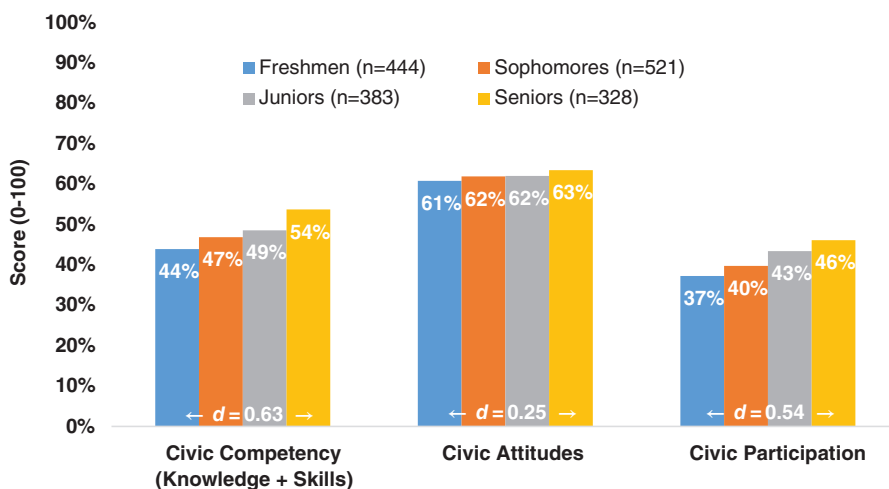

Figure 5 Performance differences on civic competency and engagement scores by college grade level. The effect size represents the standardized mean difference (Cohen's *d*) between freshmen and seniors. Note that civic participation results are only for Forms 1 and 2. The standardized mean difference for civic participation on Forms 3 and 4 was 0.49.

outperformed freshman students on civic competency by 3, 5, and 9 percentage points, respectively (Figure 5), corresponding to effect sizes of 0.19, 0.30, and 0.63 standard deviations, respectively. In terms of civic attitudes, attitudinal levels were similar across the 4 years of college; however, seniors did appear to have attitudinal levels 0.25 standard deviations higher than freshmen. Similar to civic competency, students' civic participation was higher for seniors as compared to freshmen by approximately 0.50 standard deviations. To put these values in context, previous research using the HEIghten critical thinking assessment found that seniors significantly outperformed freshmen by 0.30 standard deviations (Liu, Mao, Frankel, & Xu, 2016). Thus the score differences between freshmen and seniors were of a much larger magnitude on civic competency and civic participation when compared with differences on critical thinking. Note that the analysis did not control for prior cognitive ability and attrition at institutions and therefore any differences observed here could partially be due to a more able senior cohort. That said, there is no evidence that students with better cognitive ability are necessarily more active participants in civic activities and are more adept at intercultural interactions.

For the ICD assessment, results (Figure 6) show that senior students significantly outperformed their freshman peers in terms of both their intercultural attitudes (approach) and their intercultural abilities underscored by the analyze and act dimensions. The difference was 0.31 standard deviations in their approach scores and 0.45 standard deviations in their analyze and act scores.

Civic Knowledge and Attitudes Do Not Translate into Civic Participation

Although civic learning includes both civic competency and civic engagement, these two constructs are distinct from each other. Correlations adjusted for attenuation (i.e., unreliability) between test scores on the CCE forms (Table 4) showed
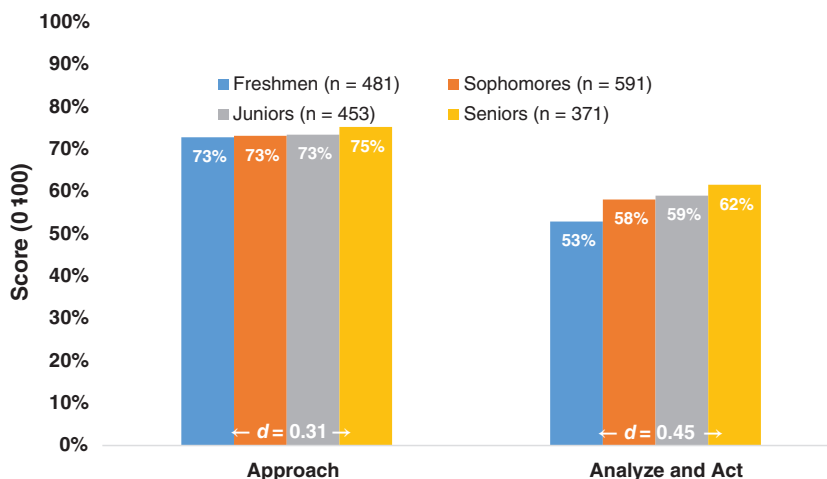
Figure 6 Performance differences on intercultural competency and diversity scores by college grade level. The effect size represents the standardized mean difference (Cohen's *d*) between freshmen and seniors. Note that the results presented are an average across the three pilot forms.

Table 4 Correlations Adjusted for Attenutation Across the Three Civic Competency and Engagement Scores for Pilot Forms

Correlation	Form 1[a]	Form 2[b]	Form 3[c]	Form 4[d]
Competency (knowledge/skills) and attitudes	.32**	.44**	.35**	.26**
Competency and participation	.38**	.24**	.13*	.25**
Attitudes and participation	.41**	.28**	.25*	.56**

**p < .001. *p < .05.
[a]N = 433. [b]N = 421. [c]N = 413. [d]N = 416.

low adjusted correlations between civic participation and civic competency and attitudes. These results suggest that an individual with high performance on civic knowledge and skills or with high attitude levels may not have active civic participation. Similarly, the adjusted correlations were also low between civic knowledge and skills and student attitudes. These low adjusted correlations demonstrate the importance of differentiating these scores in the score report to individuals and institutions so they can clearly identify strengths and weaknesses. In a recent report by the American Council of Trustees and Alumni (2016), the authors argued that although community service programs at colleges are useful, "they give students little insight into how our system of government works and what roles they must fill as citizens of a democratic republic" (p. 1). These preliminary correlation results demonstrate that these three aspects of civic learning are distinct and that each of these aspects needs to be emphasized within the college curriculum.

For dimensions on the ICD assessment, the approach dimension showed low to moderate adjusted correlations with the analyze and act dimensions, respectively (Table 5), which suggests that students' overall attitude toward and efficacy about intercultural interactions do not necessarily reflect their knowledge and skills in this domain. The two cognitive dimensions analyze and act are perfectly correlated after the scores are adjusted for unreliability. Although these are two theoretically distinct dimensions, the empirical data from the pilot study suggest that there is no practical difference between them, and therefore we decided to combine these two dimensions for score reporting purposes. Note that these analyses will be replicated using data from an operational test to empirically evaluate whether these two dimensions should be combined.

Relationship between Assessment Scores and Criterion Variables

To provide preliminary validity evidence to support the use of the CCE assessment scores, we evaluated the relationship between CCE scores (proportion correct) and self-reported scores on the *Advanced Placement® (AP®)* U.S. Government examination, self-reported AP U.S. History scores, self-rated civic knowledge/civic skills, and self-rated civic engagement level. Results showed that (a) students with AP U.S. Government or U.S. History scores of 3 or higher scored statistically

Table 5 Correlations Adjusted for Attenutation Across the Three Intercultural Competency and Diversity Scores for Pilot Forms

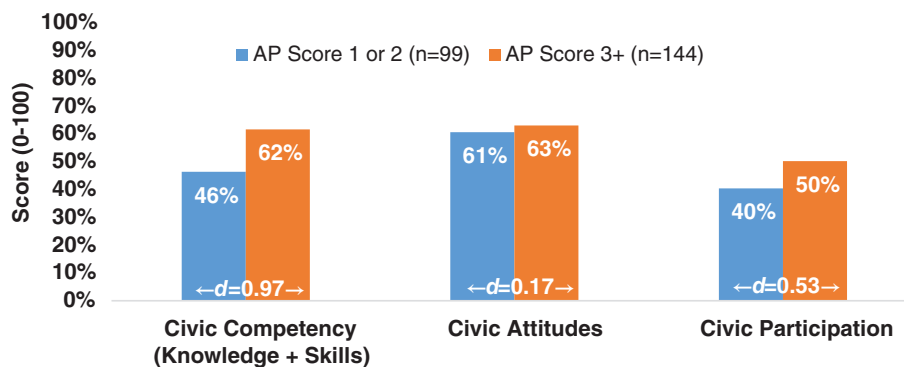| Correlation	Form 1[a]	Form 2[b]	Form 3[c]
Approach and analyze	.33*	.32*	.35*
Approach and act	.30*	.22*	.27*
Analyze and act	1.00	1.00	1.00

Note. *p < .05.
[a]N = 635. [b]N = 634. [c]N = 632.

Figure 7 The relationship of civic competency and engagement scores with self-reported advanced placement U.S. government scores. The effect size represents the standardized mean difference (Cohen's *d*) between scores. Note that civic participation results are only for Forms 1 and 2. The standardized mean difference for civic participation on Forms 3 and 4 was 0.82 for AP U.S. government.
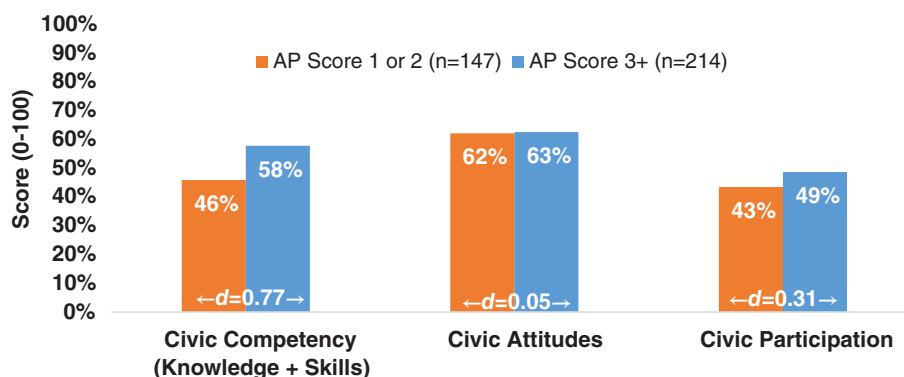

Figure 8 The relationship of civic competency and engagement scores with self-reported advanced placement U.S. history scores. The effect size represents the standardized mean difference (Cohen's *d*) between scores. Note that civic participation results are only for Forms 1 and 2. The standardized mean difference for civic participation on Forms 3 and 4 was 0.58 for AP U.S. history.

significantly higher on the civic competency and civic participation portions of the assessment as compared to students with AP scores of 1 or 2 (Figures 7 and 8, respectively); (b) students who indicated they had good or excellent civic knowledge and skills had statistically significantly higher civic competency, civic attitudes, and civic participation scores than those who rated as average or poor (Figure 9); and (c) students who indicated they had above average or high levels of civic engagement had statistically significantly higher civic competency, civic attitudes, and civic participation scores as compared to students who reported low levels of civic engagement (Figure 10).

When administering the ICD assessment, we also gathered information on background variables, such as the number of countries students traveled to, whether they had opportunities to learn about different cultures, and how often they interact with ethnically or nationally different people. Students' experience traveling abroad was associated with their ICD performance. For example, students who traveled to more than three countries outperformed students who had never traveled abroad on both the approach (*d* = 0.57; Figure 11) and the analyze and act dimensions (*d* = 0.25). Students who self-reported having more opportunity to learn and to interact with people from ethnically diverse backgrounds
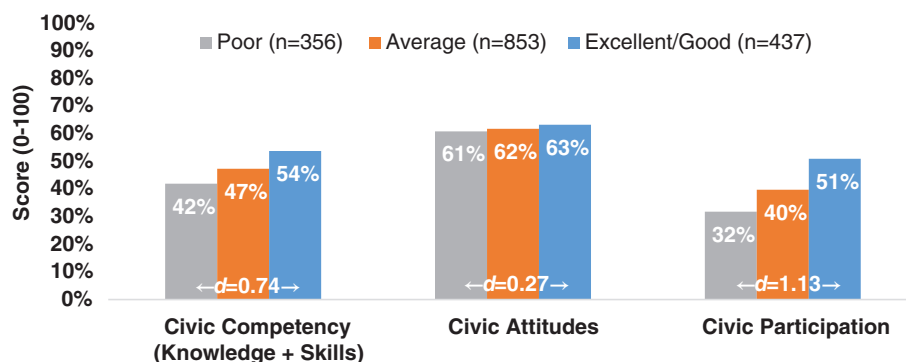
Figure 9 The relationship of civic competency and engagement scores with self-rated civic knowledge/skills. The effect size represents the standardized mean difference (Cohen's *d*) between scores excellent and poor and above average/high and low groups, respectively. Note that civic participation results are only for Forms 1 and 2. The standardized mean difference for civic participation on Forms 3 and 4 was 0.94 for self-rated knowledge/skills.
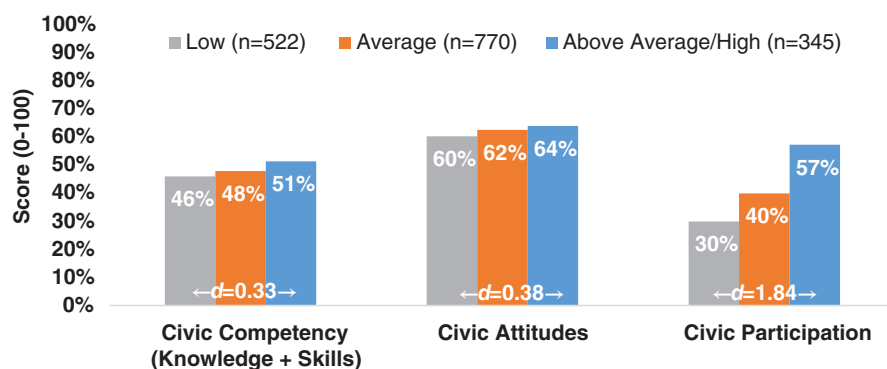

Figure 10 The relationship of civic competency and engagement scores with level of civic engagement. The effect size represents the standardized mean difference (Cohen's *d*) between scores excellent and poor and above average/high and low groups, respectively. Note that civic participation results are only for Forms 1 and 2. The standardized mean difference for civic participation on Forms 3 and 4 was 1.69 for self-rated level of civic engagement.

also performed significantly better than students who did not have such opportunities, and the effect sizes ranged from moderate to large (Figures 12 and 13), which points to the importance for institutions to provide opportunities on campus to increase students' intercultural experience. It is important to note, however, that opportunity to travel abroad may also be directly related to socioeconomic status.

Do Males and Females Differ on Civic and Intercultural Competency?

We also evaluated differences in performance across males and females. Results show that males statistically significantly outperformed females on civic competency by 2 percentage points or 0.26 standard deviations. For the two aspects of civic engagement, however, females performed significantly higher (Figure 14). Previous research has found mixed results in terms of gender differences in civic learning assessments; however, much of the literature has shown females with higher overall scores (Reason & Hemer, 2015). For instance, Hu (2008) found that females are more likely to engage in civic behavior, and other studies have shown that females report higher scores on civic values (Cabrera et al., 2002; Hurtado, Engberg, Ponjuan, & Landreman, 2002; Pascarella, Edison, Nora, Hagedorn, & Terenzini, 1996). That said, Dolan (2011) found that males outperformed females on topics related to political knowledge. Similarly, Delli Carpini and Keeter (1996) found that the gender performance difference depended on the type of political and civic knowledge. For instance, adult males performed higher on topics related to war and political power, whereas females performed higher on topics related to social welfare policy or education.

   On the ICD assessment, females outperformed males on both the approach and the analyze and act dimensions (Figure 15). According to the Institute of International Education (IIE), among the U.S. students who study abroad, 67%
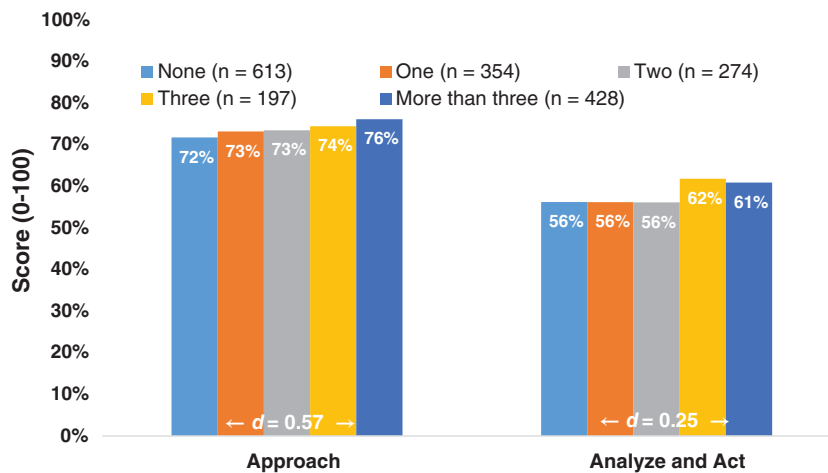
Figure 11 Performance differences on intercultural competency and diversity scores by travel abroad experiences. The effect size represents the standardized mean difference (Cohen's *d*) between individuals traveling to more than three and no foreign countries. Note that performances are based on an average across three pilot forms.
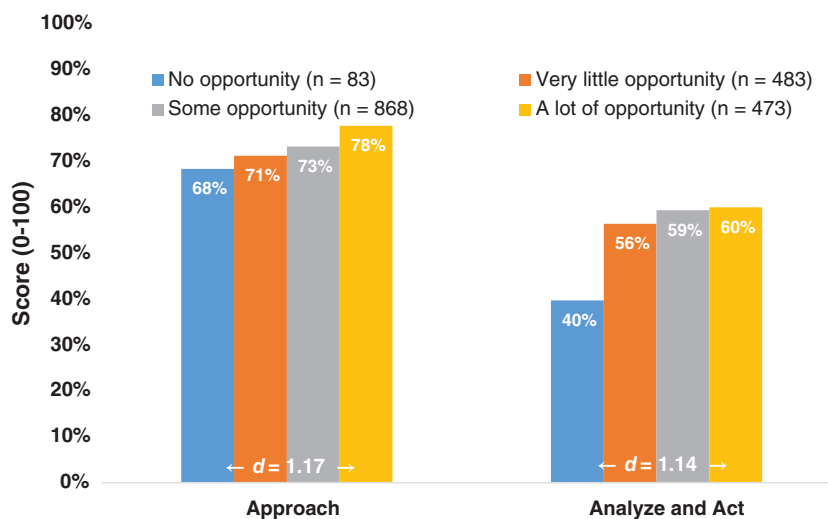

Figure 12 Performance differences on intercultural competency and diversity scores by opportunity to learn about other cultures. The effect size represents the standardized mean difference (Cohen's *d*) between individuals have a lot of and no opportunity to learn about other cultures. Note that performances are based on an average across three pilot forms.

are female (Witherell, 2016). Previous research has also shown that gender is a strong predictor of intent to study abroad (Stroud, 2010). There is evidence that intercultural competence is a malleable skill and that it could be improved with proper training that could be provided through study abroad programs (e.g., Engle & Crowne, 2014). It is important for higher education institutions to reach out to all students when promoting intercultural competency.

Design of the Score Reports

For the CCE and ICD scores to be useful, it is critical to determine the most effective way to report the scores to institutional and student users. Having clear and useful score reports can support users in making appropriate score-based inferences (Hambleton & Zenisky, 2013). If a score report is developed in a clear and meaningful way, then the score report can provide actionable data to test users and guidance for interpreting scores appropriately.

As there is no "one-size-fits-all" method for developing score reports, one approach for score report development is to conduct focus groups and interviews with test users to better understand their needs and preferences (Goodman & Hambleton, 2004; Hattie, 2009; Zenisky & Hambleton, 2012). For the CCE and ICD assessments, we have conducted
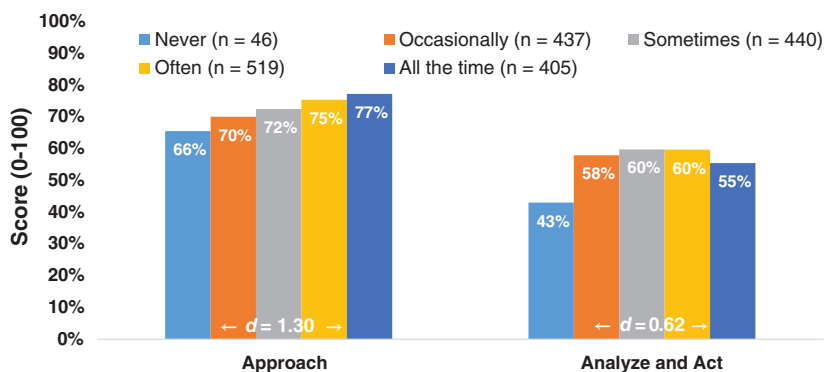
Figure 13 Performance differences on intercultural competency and diversity scores by opportunity to interact with students of a different ethnicity or nationality. The effect size represents the standardized mean difference (Cohen's *d*) between individuals who interact all the time and never with students who are ethnically or nationally different from themselves. Note that performances are based on an average across three pilot forms.
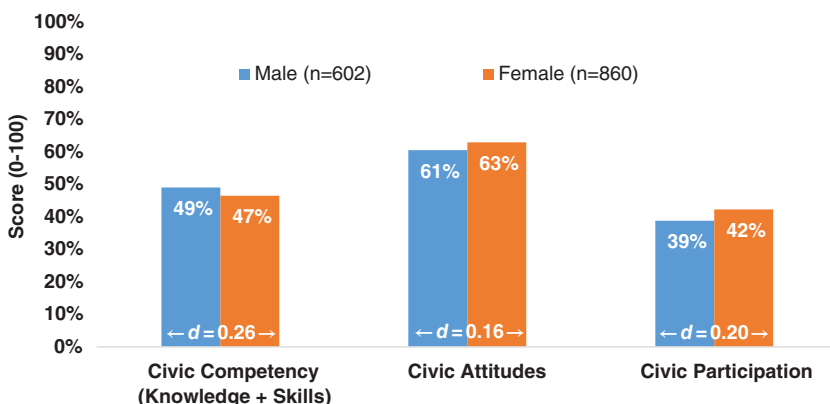

Figure 14 Performance differences on civic competency and civic attitudes by gender. The effect size represents the standardized mean difference (Cohen's *d*) between males and females. Note that civic participation results are only for Forms 1 and 2. The standardized mean difference for civic participation on Forms 3 and 4 was 0.11.

online and in-person focus groups with institutional stakeholders to better understand how they intend to use the results from these assessments (e.g., for accreditation or curricular improvement) and what types of information would be useful in supporting these intended uses.

For CCE, we conducted one in-person and three online focus groups with a total of 10 higher education institutions that were involved in the pilot study. Institutional users indicated that having a high-level summary of the results is important for showing these findings to other individuals throughout the institution, including faculty, and is also useful for accreditation purposes. Institutions also stressed the importance of having benchmarking information so that they could compare their performance to other institutions. Subscores, performance-level descriptors, and item-level results were most useful for diving more deeply into the results and for determining how the results could relate back to curriculum. These additional types of scores could also be useful for evaluating the impact of changes made in the curriculum or for comparing different groups of students (e.g., transfer students). A few institutions also indicated that they intended to use the results to support their mission statements regarding citizenship.

For ICD, one in-person and three online focus groups were conducted with 14 higher education institutions from the pilot study. Results from these focus groups are similar to those for CCE, with institutions planning to use the assessment for accreditation, improvement of the curriculum, benchmarking, and evaluating learning over time. Institutions expressed a particular interest in the evaluation of campus programs, such as a course on diversity or a program on cultural issues. Group-level subscores were of special interest for ICD to pinpoint particular strengths and weaknesses of various subgroups of students. Specific to ICD, there was also interest in linking scores to the study abroad programs and to residential life on campus (especially for campuses with large international student populations).
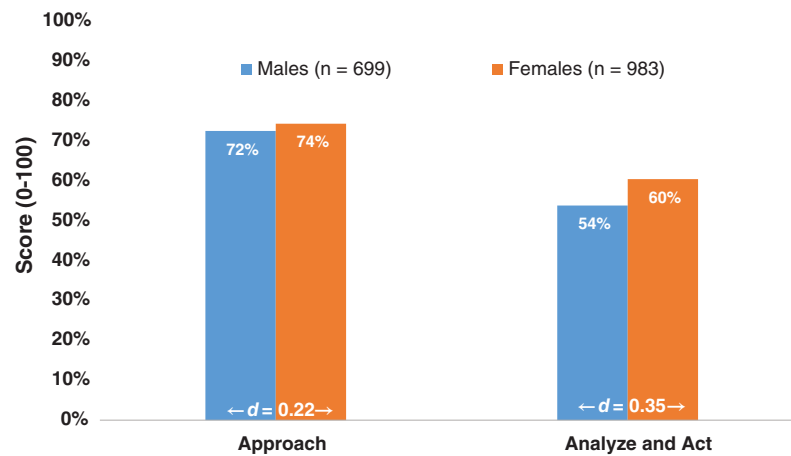
Figure 15 Performance differences on approach and analyze and act scores by gender. The effect size represents the standardized mean difference (Cohen's *d*) between males and females. Note that these results are based on the average across all three forms.

Both the CCE and ICD focus groups stressed the importance of including various elements on the score report to optimize the information provided to users. Feedback gathered from the focus groups revealed a strong preference for elements such as total score; subscores that represent students' performance on the subscales as defined by the assessment; proficiency levels with performance-level descriptors that provide information on what students may be expected to know or do at each level; score distributions; benchmarking information in relevance to both the national sample and a sample of institutions of similar profile (i.e., information presented from a comparison group); and item-level information, such as the percentage of students who endorsed a particular statement. The final score reports will be designed relying on the feedback from these focus groups along with psychometric and statistical analysis results from the pilot and field tests.

Conclusion

Both educators and employers have asked for well-rounded college graduates for an effective workforce and high-functioning society at large. Skills such as civic and intercultural competencies are critically important for college graduates to make fulfilling contributions as citizens and employees. To gain a competitive edge in today's global economy, employers now seek college graduates with sufficient knowledge of civic engagement and intercultural competency (Casner-Lotto & Barrington, 2006; Gould, 2011; Hart Research Associates, 2015), as satisfactory performance of many jobs requires effective global citizens. College graduates are expected to be able to work with colleagues from all parts of the world with diverse opinions, cultures, and values. They should recognize the increasing global challenges that they must jointly address as citizens (Adelman et al., 2014). As indicated in the Degree Qualifications Profile, "by exploring global, intercultural, scientific and economic topics, students pursue questions that both prepare them for civic participation and create a larger context for their specialized interests" (Adelman et al., 2014, p. 12).

Despite the increased importance of civic and intercultural competence, limitations with prior measurement tools have constrained the assessment of and also the opportunity for improving students' civic and intercultural competencies. One major limitation has been the myriad terminology and definitions for both of these competencies. Torney-Purta et al. (2015) and Griffith et al. (2016) both reviewed and synthesized existing definitions and frameworks to create operational definitions for civic competency and intercultural competency, respectively. These two frameworks provide clear definitions of terms and recommendations for assessing these competencies. The HEIghten assessment is one example of how the frameworks can be used to guide assessment design. The frameworks can also be used by institutions when they develop their missions and student learning outcome statements and when they make decisions about assessment adoption or develop their own instruments.

Early Results Suggest Room for Student Improvement in These Competencies

Results from this large-scale pilot show that the multidimensional constructs can be adequately assessed. The results also show that there is clearly room for improvement in terms of students' civic and intercultural knowledge and skills that

institutions can focus on in their efforts to further promote these essential competencies. Both civic and intercultural competencies are malleable traits in that they can be improved with proper training and development (e.g., Engle & Crowne, 2014). For example, the findings from this study show that students who reported having more opportunities to learn about other cultures and to interact with others from ethnically diverse backgrounds performed substantially better on both the attitudinal and knowledge and skills dimensions of the HEIghten ICD assessment. These results suggest that more exposure to intercultural experiences is associated with greater awareness and enhanced performance, providing early evidence that the ICD assessment measures what it was designed to measure in that the relationship between ICD assessment scores and relevant variables confirmed expectations.

It is important for institutions to offer programs and opportunities to help students increase their awareness of and gain experience in civic and intercultural domains. Institutions need clear and actionable data from these assessments to evaluate and improve their curriculum. Focus groups with institutional stakeholders confirmed that subscore information can be useful data for these purposes and can also be used to evaluate institutional missions. As institutions around the globe constantly strive to achieve academic excellence, it is imperative that they promote critical skills, such as civic and intercultural competencies, so their graduates are likely to gain a competitive advantage in the global workplace. On a related note, the two constructs of interest in this study are related to social and emotional skills that have been heavily discussed and promoted in schools. A recent report under the aegis of a 28-member Council of Distinguished Scientists affirmed the importance and even economic value of social and emotional skills (Jones & Kahn, 2017) — that for every $1 investment in social and emotional skills interventions, there are long-term benefits averaging $11 (Belfield et al., 2015). The promotion of skills such as civic and intercultural competency should also extend from higher education to K–12 for a greater continuity in building these skills for students. For both K–12 and higher education, the concept of the "whole student" matters.

Notes

1 Student-level reliability was calculated separately for each test form using coefficient alpha. Institution-level reliability was calculated using a split-sample approach with institutional means as described in Klein et al. (2009).
2 DIF analyses were not conducted with regard to race/ethnicity because of small sample sizes.

References

Adelman, C., Ewell, P., Gaston, P., & Schneider, C. G. (2011). *The Degree Qualifications Profile: Defining degrees — A new direction for American higher education to be tested and developed in partnership with faculty, students, leaders, and stakeholders*. Indianapolis, IN: Lumina Foundation for Education. Retrieved from ERIC database (ED515302).

Adelman, C., Ewell, P., Gaston, P., & Schneider, C. G. (2014). *The degree qualifications profile: A learning-centered framework for what college students should know and be able to do to earn the associate* (bachelor's or master's degree). Lumina Foundation for Education, Indianapolis, IN.

American Association of State Colleges and Universities. (2014). *eCitizenship: New tools, new strategies, new spaces*. Retrieved from http://www.aascu.org/programs/adp/eCitizenship/

American Council of Trustees and Alumni. (2016). *A crisis in civic education*. Washington, DC: Author.

Arasaratnam, L. A. (2008). Acculturation process of Sri Lankan Tamil immigrants in Sydney: An ethnographic analysis using the bidirectional model (BDM). *Australian Journal of Communication, 35*, 57–68.

Baumert, J., & Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, *16*, 441–462. https://doi.org/10.1007/BF03173192

Belfield, C., Bowden, A. B., Klapp, A., Levin, H., Shand, R., & Zander, S. (2015). The economic value of social and emotional learning. *Journal of Benefit-Cost Analysis, 6*, 508–544. https://doi.org/10.1017/bca.2015.55

Cabrera, A. F., Nora, A., Crissman, J. L., Terenzini, P. T., Bernal, E. M., & Pascarella, E. T. (2002). Collaborative learning: Its impact on college students' development and diversity. *Journal of College Student Development*, *43*, 20–34.

Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century U.S. workforce*. Retrieved from the P21: Partnership for 21st Century Learning website: http://www.p21.org/storage/documents/FINAL_REPORT_PDF09-29-06.pdf

Colby, A., Beaumont, E., Ehrlich, T., & Corngold, J. (2007). *Educating for democracy: Preparing undergraduates for responsible political engagement*. San Francisco, CA: Jossey-Bass.

Colby, A., Ehrlich, T., Beaumont, E., & Stephens, J. (2003). *Educating citizens: Preparing America's undergraduates for lives of moral and civic responsibility*. San Francisco, CA: Jossey-Bass.

Conant, B. J. (1945). *General education in a free society: Report of the Harvard Committee*. Cambridge, MA: Harvard University Press.

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education, 10*, 241–266. https://doi.org/10.1177/1028315306287002

Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. New Haven, CT: Yale University Press.

Dolan, K. (2011). Do women and men know different things? Measuring gender differences in political knowledge. *Journal of Politics, 73*, 97–107. https://doi.org/10.1017/S0022381610000897

Earley, P. C., & Peterson, R. S. (2004). The elusive cultural chameleon: Cultural intelligence as a new approach to intercultural training for the global manager. *Academy of Management Learning & Education*, *3*, 100–115. https://doi.org/10.5465/AMLE.2004.12436826

Engle, R. L., & Crowne, K. A. (2014). The impact of international experience on cultural intelligence: An application of contact theory in a structured short-term programme. *Human Resource Development International, 17*, 30–46. https://doi.org/10.1080/13678868.2013.856206

Goodman, D. P., & Hambleton, R. K. (2004). Student test score reports and interpretive guides: Review of current practices and suggestions for future research. *Applied Measurement in Education*, *17*, 145–220. https://doi.org/10.1207/s15324818ame1702_3

Gould, J. (Ed.). (2011). *Guardian of democracy: The civic mission of schools*. Washington, DC: National Conference on Citizenship.

Griffith, D. A., & Harvey, M. G. (2001). An intercultural communication model for use in global interorganizational networks. *Journal of International Marketing, 9*(3), 87–103. https://doi.org/10.1509/jimk.9.3.87.19924

Griffith, R. L., Wolfeld, L., Armon, B. K., Rios, J., & Liu, O. L. (2016). *Assessing intercultural competence in higher education: Existing research and future directions* (Research Report No. RR-16-25). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12112

Hambleton, R. K., & Zenisky, A. L. (2013). Reporting test scores in more meaningful ways: Q research-based approach to score report design. In K. F. Geisinger (Ed.), *American Psychological Association handbook of testing and assessment in psychology* (pp. 479–494). Washington, DC: American Psychological Association. https://doi.org/10.1037/14049-023

Hart Research Associates. (2015). *Falling short? College learning and career success*. Washington, DC: Association of American Colleges and Universities.

Hart Research Associates. (2016). *Recent trends in general education design, learning outcomes, and teaching approaches: Key findings from a survey among administrators at AAC&U member institutions*. Washington, DC: Association of American Colleges and Universities.

Hattie, J. (2009, April). *Visibly learning from reports: The validity of score reports*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Hu, S. (2008). Do financial aid awards in college affect graduates' democratic values and civic engagement? *Journal of College and Character*, *10*(1), 1–16. https://doi.org/10.2202/1940-1639.1062

Hurtado, S., Engberg, M. E., Ponjuan, L., & Landreman, L. (2002). Students' precollege preparation for participation in a diverse democracy. *Research in Higher Education, 43*, 163–186. https://doi.org/10.1023/A:1014467607253

Institute of International Education. (2016). *Open doors 2016 "fast facts"*. Retrieved from https://www.iie.org/-/media/Files/Corporate/Open-Doors/Fast-Facts/Fast-Facts-2016.ashx?la=en&hash=9E918FD139768E1631E06A3C280D8A9F2F22BBE1

Jones, S. M., & Kahn, J. (2017). *The evidence base for how we learn: Supporting students' social, emotional, and academic development*. Washington, DC: Aspen Institute.

King, P. M., & Baxter Magolda, M. B. (2005). A developmental model of intercultural maturity. *Journal of College Student Development, 46*, 571–592. https://doi.org/10.1023/A:1014467607253

Klein, S., Liu, O. L., Sconing, J., Bolus, R., Bridgeman, B., Kugelmass, H., … Steedle, J. (2009). *Test validity study (TVS) report*. Retrieved from the Council for Aid in Education website: http://cae.org/images/uploads/pdf/13_Test_Validity_Study_Report.pdf

Kuh, G. D., Jankowski, N., Ikenberry, S. O., & Kinzie, J. (2014). *Knowing what students know and can do: The current state of student learning outcomes assessment in U.S. colleges and universities*. Champaign, IL: National Institute for Learning Outcomes Assessment.

Leung, K., Ang, S., & Tan, M. L. (2014). Intercultural competence. *Annual Review of Organizational Psychology and Organization Behavior, 1*, 489–519. https://doi.org/10.1146/annurev-orgpsych-031413-091229

Liu, O. L. (2017). Ten years after the spellings commission: From accountability to internal improvement. *Educational Measurement: Issues and Practice, 36*(20), 34–41. https://doi.org/10.1111/emip.12139

Liu, O. L., Mao, L., Frankel, L., & Xu, J. (2016). Assessing critical thinking in higher education: The HEIghten approach and preliminary validity evidence. *Assessment & Evaluation in Higher Education, 41*, 677–694. https://doi.org/10.1080/02602938.2016.1168358

Markle, R., Brenneman, M., Jackson, T., Burrus, J., & Robbins, S. B. (2013). *Synthesizing frameworks of higher education student learning outcomes* (Research Report No. RR-13-22). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2013.tb02329.x

National Task Force on Civic Learning and Democratic Engagement. (2012). *A crucible moment: College learning and democracy's future*. Washington, DC: Association of American Colleges and Universities.

Northouse, P. G. (2010). *Leadership: Theory and practice* (5th ed.). Thousand Oaks, CA: Sage.

Pascarella, E. T., Edison, M., Nora, A., Hagedorn, L. S., & Terenzini, P. T. (1996). Influences on students' openness to diversity and challenges in the first year of college. *The Journal of Higher Education, 67*, 174–195. https://doi.org/10.1080/00221546.1996.11780255

Perry, L., Stoner, K. R., Stoner, L., Wadsworth, D., Page, R., & Tarrant, M. A. (2016). The importance of global citizenship to higher education: The role of short-term study abroad. *Cambridge Journal of Education and Science, 1*, 754–769.

Reason, R. D., & Hemer, K. M. (2015). *Civic learning and engagement: A review of the literature on civic learning, assessment and instruments*. Retrieved from https://www.aacu.org/sites/default/files/files/qc/Civic%20Learning%20and%20Engagement%20A%20Review%20of%20the%20Research%20Literature%20on%20Civic%20Learning%20Assessment%20and%20Instruments.pdf

Rhodes, T. L. (2010). *Assessing outcomes and improving achievement: Tips and tools for using rubrics*. Washington, DC: Association of American Colleges and Universities.

Rios, J. A., & Anguiano-Carrasco, C. (2016, April). *Socially desirable responding in noncognitive low-stakes assessments: Should we be concerned?* Paper presented at the annual conference of the American Educational Research Association, Washington, DC.

Stroud, A. H. (2010). Who plans (not) to study abroad? An examination of U.S. student intent. *Journal of Studies in International Education, 14*, 491–507. https://doi.org/10.1177/1028315309357942

Torney-Purta, J., Cabrera, J. C., Roohr, K. C., Liu, O. L., & Rios, J. A. (2015). *Assessing civic competency and engagement in higher education: Research background, frameworks, and directions for next-generation assessment* (Research Report No. RR-15-34). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12081

Witherell, S. (2016). *IIE releases open doors 2016 data* [Press release]. Retrieved from https://www.iie.org/Why-IIE/Announcements/2016-11-14-Open-Doors-Data

Zenisky, A. L., & Hambleton, R. K. (2012). Developing test score reports that work: The process and best practices for effective communication. *Educational Measurement: Issues and Practice, 31*(2), 21–26. https://doi.org/10.1111/j.1745-3992.2012.00231.x

Suggested citation: