Received: 7 February 2022 | Revised: 5 February 2023 | Accepted: 26 July 2023

DOI: 10.1002/jee.20552

RESEARCH ARTICLE

JEE Journal of Engineering Education

Exploring scenario-based assessment of students' global engineering competency: Building evidence of validity of a China-based situational judgment test

Kirsten A. Davis[1] | **Brent K. Jesiek**[1] | **David B. Knight**[2]

[1]School of Engineering Education, Purdue University, Neil Armstrong Hall of Engineering, West Lafayette, Indiana, USA

[2]Department of Engineering Education, Virginia Tech, Goodwin Hall, Blacksburg, Virgina, USA

Correspondence
Kirsten A. Davis, School of Engineering Education, Purdue University, Neil Armstrong Hall of Engineering, 701 West Stadium Avenue, West Lafayette, IN 47907, USA.
Email: kad@purdue.edu

Funding information
National Science Foundation, Grant/Award Numbers: 1160455, 1254323; Virginia Tech Graduate School

Abstract

Background: Engineers operate in an increasingly global environment, making it important that engineering students develop global engineering competency to prepare them for success in the workplace. To understand this learning, we need assessment approaches that go beyond traditional self-report surveys. A previous study (Jesiek et al., *Journal of Engineering Education* 2020; 109(3):1–21) began this process by developing a situational judgment test (SJT) to assess global engineering competency based in the Chinese context and administering it to practicing engineers.

Purpose: We built on this previous study by administering the SJT to engineering students to explore what prior experiences related to their SJT performance and how their SJT performance compared with practicing engineers' performance on the SJT.

Method: Engineering students completed a survey including the SJT and related self-report survey instruments. We collected data from three groups of students: those who had studied abroad in China; those who had studied abroad elsewhere; and those who had not studied abroad.

Results: We found that students' SJT performance did not relate to their scores on the self-report instruments, but did relate to their participation in study abroad programs. The students also performed better on the SJT when compared to the practicing engineers.

Conclusions: Our results highlight the need to use multiple forms of assessment for global engineering competence. Although building evidence for the validity of the Global Engineering Competency China SJT is an ongoing process, this data collection technique may provide new insights on global engineering competency compared to traditionally used assessments.

KEYWORDS

assessment, intercultural competence, international programs, study abroad

1 | INTRODUCTION

Engineering work responds to broader changes in technology and society. As globalization continues to advance and evolve, people across the globe are more connected through technology and overseas travel than ever before (Friedman, 2005; Friedman, 2017). Globalization has led to the expansion of corporate networks around the world, as most companies now have connections with suppliers, customers, contractors, or their own employees in different countries. This development in corporate global connectedness means that engineers are likely to experience situations in their day-to-day engineering work where they need to collaborate with or design for people in other countries (Johri & Jesiek, 2014). Such situations demand unique skills and behaviors including intercultural communication, cultural sensitivity, and consideration of contextual differences and requirements. These skills that are necessary for global engineering work are collectively known as *global engineering competency* (GEC; Jesiek, Zhu, et al., 2014).

Just as engineering work evolves in response to the changing world, engineering education receives pressures to adapt to changes in the workplace. To align educational environments with professional practice, it is important to understand what competencies engineers need to be successful in the increasingly complex and global workforce. Previous works have suggested that there is a disconnect between university and industry environments (Stevens et al., 2014; Walther & Radcliffe, 2007), resulting in engineers who are not adequately prepared to successfully collaborate in developing solutions to complex problems. Research focused on characterizing engineering work has explored the types of problems faced in the engineering workforce, how working with people is a part of engineering practice, and the boundaries that engineers must negotiate in their work (Jesiek et al., 2016; Stevens et al., 2014; Trevelyan, 2010). Specific to GEC, case studies and interviews have been analyzed to understand the types of global situations faced by working engineers (Jesiek et al., 2015; Jesiek, Zhu, et al., 2014). A model of GEC has been developed based on this information, but more work is needed to gather evidence of validity for the model and related assessment tools (Johri & Jesiek, 2014).

To address this need, our study builds on prior work in which country-specific global engineering scenarios have been developed into a situational judgement test (SJT) to assess GEC (Jesiek et al., 2015; Jesiek, Zhu, et al., 2014). These scenarios evaluate participants' awareness of appropriate behavior in specific situations by having participants rate multiple potential approaches to each situation that vary in terms of appropriateness within specific contexts. In developing a new assessment instrument, it is important that researchers present a collection of evidence for validity that supports the use of the instrument for its intended purpose (American Educational Research Association, 2014; Educational Testing Service, 2015). Our study seeks to contribute evidence of validity for one of these scenario-based assessments focused on the Chinese context (China SJT). Initial evidence of validity for the China SJT was presented in an earlier study where the instrument was administered to practicing engineers (Jesiek et al., 2020), and our work contributes additional evidence by addressing the following research questions:

1. How does students' SJT performance relate to their responses on similar instruments?
2. How does students' SJT performance relate to their
   a. group membership based on having participated in global programs in China during college, global programs elsewhere, or no global programs?
   b. prior international experiences, professional experiences, and demographic characteristics?
3. How does students' SJT performance compare with that of practicing engineers?

2 | LITERATURE REVIEW

The theoretical underpinning for this study is GEC, that is, "the attributes uniquely or especially relevant for cross-national/cultural engineering practice" (Johri & Jesiek, 2014, p. 660). There have been several attempts to define GEC more specifically and to outline the characteristics associated with this competency. Different methods have been used to explore this topic, and researchers have suggested a variety of theoretical frameworks based on this research. Because there is limited evidence of validity in existing models for this construct, we provide a summary of the prior studies that have explored this idea and consider the similarities and differences between them. We also highlight some of the assessments in the engineering education literature that have been used as proxy indicators to measure student attitudes, awareness, and beliefs related to this competency.

2.1 | Prior studies of global engineering competency

Although many studies have explored how to teach global skills to engineering students, fewer studies have identified which global skills should be taught. We focus on six studies that presented either a list of skills or a complete framework of GEC, as summarized in Table 1. One of the key differences between these GEC frameworks is whether and how they demonstrate integration between the global and engineering aspects of GEC. In the last column of Table 1, we characterize each framework based on the level of integration: either high, medium, or low, with some comments to explain the rating. A "high" characterization indicates that all aspects or dimensions of a framework demonstrate integration; a "medium" characterization indicates that some dimensions are integrated, but others focus only on engineering or global skills; and a "low" characterization indicates that the engineering and global skills are represented by different dimensions. These characteristics will be discussed in more detail in the remainder of this section.

Despite the various methods used to develop these frameworks, there are several points of agreement. First, these frameworks all discuss the need for engineers to collaborate and communicate with people across cultural differences. This commonality reflects the arguments of Downey et al. (2006) that cross-cultural collaboration tends to be the focus in conversations about global competency in engineering. However, most of the frameworks move beyond simply addressing collaboration, considering also the influence of professional skills, engineering skills, or both. Allert et al. (2007) and Levonisova et al. (2014), for example, both include a *professional competency* component in their models, referring to characteristics such as leadership, work ethic, interdisciplinary skills, innovative thinking, and self-learning. Parkinson's (2009) list of characteristics highlights similar professional skills and additionally includes engineering-focused skills such as understanding product design or supply chain management. These skills can also be found in the *technical competency* dimension of the Allert et al. (2007) model and are alluded to in Ragusa's (2014) adaptation of the global citizenship model. Similarly, the categories suggested by Jesiek, Zhu, et al. (2014) contain aspects of professional and technical skills, although perhaps not as explicitly as the other models. Thus, in general there is agreement that GEC requires skills beyond simply communicating across cultural difference, including both technical and professional competencies.

The main difference across these frameworks is how they describe the relationships between global, technical, and professional competencies. Both the Allert et al. (2007) and Levonisova et al. (2014) frameworks divide these competencies into *separate* dimensions, although the latter lists some integrated skills in their description of the *Global*

TABLE 1 Summaries of global engineering competency (GEC) frameworks.

Framework	Citation	Framework basis	Structure	Integration?
The Globally Competent Engineer	Downey et al. (2006)	Historical cultural analysis	Definition: Work with people who define problems differently	High—Implied
Attributes of the Global Engineering Professional	Allert et al. (2007)	Experience running global programs	Three dimensions: Professional, global, technical	Low—Separate dimensions
Global Competence for Engineers	Parkinson (2009)	Literature review and experience running global programs	Thirteen attributes of global competence, some engineering-specific	Medium—Some attributes
Engineering Global Preparedness	Ragusa (2014)	Existing model of Global Citizenship	Four dimensions, added engineering to each existing dimension	Medium—Some dimensions
Conceptual Model of Engineering Global Preparedness	Levonisova et al. (2014)	Delphi study with 18 subject-matter experts from industry and academia	Four categories: Cross-cultural, professional, contextual, global engineering	Low—Part of one dimension
Global Engineering Competency	Jesiek, Zhu, et al. (2014)	Analysis of global case studies, interviews, and focus groups with global engineers	Three dimensions: Technical coordination, engineering cultures, ethics/standards	High—Explicit integration

Engineering dimension (alongside technical skills). On the other hand, the Parkinson (2009) and Ragusa (2014) frameworks combine these topics *within* individual dimensions or characteristics. For example, one of the characteristics listed by Parkinson (2009) is "Understand implications of cultural differences on how engineering tasks might be approached" (p. 11). Similarly, part of the Ragusa (2014) definition of the *Engineering Global-centrism* dimension in the framework is "Making judgements based on global needs for engineering and associated technologies, while not focusing on ethnocentric standards" (p. 405). This integration suggests a view that GEC is not simply the possession of global skills and technical skills, but is the ability to unify them.

Jesiek, Zhu, et al. (2014) present a framework that explicitly argues for this unification and recognizes GEC as a unique competency in itself. The authors differentiate GEC from global, technical, and professional competencies, describing it as "Capabilities and job requirements that are uniquely or especially relevant for effective engineering practice in a global context" (p. 3). This framework identifies three dimensions, each of which integrates engineering and global components. *Technical coordination* describes the informal management tasks engineers perform in working on teams to complete their projects. *Understanding and negotiating engineering cultures* addresses how engineers navigate cultural and national differences in technical work practices (e.g., taking theoretical vs. practical approaches to problem solving). *Navigating ethics, standards, and regulations* refers to differences in ethical and regulatory perspectives across cultures and national boundaries and how these differences influence engineering work. These various perspectives on the relationships that form or connect GEC and related competencies suggest the need to further explore how GEC manifests in global engineering work (Johri & Jesiek, 2014).

A second and slightly subtler difference between frameworks is how they characterize GEC specifically. As previously described, two of the frameworks do not seem to consider GEC as a unique competency on its own (Allert et al., 2007; Levonisova et al., 2014). Although the framework proposed by Ragusa (2014) presents an integrated version of GEC, it focuses heavily on global citizenship perspectives and their relationship to engineering, making it most applicable in service learning or development environments (Johri & Jesiek, 2014). The characteristics from Parkinson's (2009) list that represent GEC (rather than simply global competence) describe how engineering tasks and ethics/standards may vary across cultures. These are similar to the categories *Understanding and Negotiating Engineering Cultures* and *Navigating Ethics, Standards, and Regulations* in Jesiek, Zhu, et al.'s (2014) framework. Moreover, although all of the frameworks discuss global collaboration in a general sense, only Jesiek, Zhu, et al. (2014) consider it from an engineering-specific perspective, making it an integrated component of GEC. Their category *Technical Coordination* builds on work by Trevelyan (2010) suggesting that the social and technical components of engineering work cannot be viewed distinctly. Similarly, Jesiek, Zhu, et al. (2014) present global collaboration, communication, leadership, and management by engineers as deeply connected to their technical competency. Thus, this final framework by Jesiek, Zhu, et al. (2014) presents the most integrated perspective on the types of skills and behaviors that may be included within the scope of GEC. Our study builds on this framework to explore the awareness and beliefs about GEC of both engineering professionals and engineering students.

2.2 | Assessing global competency and global engineering competency

The most common quantitative assessments for global competency are self-report surveys designed to measure different aspects of global competency (Johri & Jesiek, 2014). For example, the Intercultural Development Inventory (IDI) claims to measure *intercultural sensitivity*, defined as the complexity with which an individual experiences cultural difference (Hammer et al., 2003). On the other hand, the Global Perspectives Inventory (GPI) tries to capture student development, aligning with the model of *intercultural maturity* which views growth as a process of making meaning of experiences (Braskamp et al., 2009; King & Baxter Magolda, 2005). Global engineering programs have made use of both of these instruments (Levonisova et al., 2015; Lohmann et al., 2006), as well as others, such as the Miville-Guzman Universality–Diversity Scale (MGUDS; used by Jesiek, Haller, & Thompson, 2014) and the Cultural Intelligence Scale (used by Knight et al., 2017). Choosing an instrument seems to depend on which aspect of global competency the assessors are interested in measuring and their financial resources (i.e., some instruments follow a fee-for-use model), but in sum, most of the existing instruments take the form of a self-assessment rather than a behavioral assessment.

Few assessments focus on GEC-related constructs (Johri & Jesiek, 2014). One instrument that builds directly on a theory of global competency (global citizenry theory) is the Engineering Global Preparedness Index (EGPI; Ragusa, 2014). This self-assessment survey is similar to the typical global competency assessments as it was adapted from an existing survey for an engineering context. Although the instrument may help assess knowledge associated with GEC, some researchers have called for more direct assessment of skills and behaviors related to GEC (DeBoer et al., 2013; Jesiek, Zhu, et al., 2014). Along these lines, Lohmann et al. (2006) propose using analyses of senior design projects and employer evaluations to assess GEC, and Jesiek, Haller, and Thompson (2014) asked questions about GEC concepts during post-trip interviews. Similarly, DeBoer et al. (2013) advocate the use of performance- or scenario-based tasks to assess students' behavioral responses and provide examples of rubrics that could be used toward that end. However, these practices are program-specific in nature and do not connect to a concrete definition of GEC, as such assessments have only recently begun to appear in the literature.

In contrast, Jesiek et al. (2020, 2015) describe the development process of country-specific situational judgment tests (SJTs) based on their model of GEC. Framed using the dimensions of technical coordination, engineering cultures, and ethics, standards, and regulations, the researchers developed SJT questions for each dimension for multiple different countries. These questions involve a short scenario description based on a real experience in global engineering followed by four or five possible actions the reader could take in the situation. The reader is asked to rate each possible response on a scale from 1 ("not at all effective") to 10 ("very effective"), and reader responses are then compared to responses from subject-matter experts (SMEs; Jesiek et al., 2015; Jesiek et al., 2020). The SMEs used for the development of both the scenarios and responses (two separate groups) were professional engineers working primarily in large corporations and who had multiple years of experience working in a specific cultural context. Although the development of this method is still under way, the team has developed SJT questions for China, Japan, India, Germany, France, and Mexico. The creators argue that this method more directly connects to dimensions of global competency known to be relevant in engineering work (Jesiek, Zhu, et al., 2014) and that the SJT format focuses on behavioral tendencies, making it hard to fake or guess the "right" answers (Jesiek et al., 2015, 2020).

Because evidence of validity is still being collected for the SJT instruments, there are few examples of their use in assessing global engineering programs, although one global program has used an early version of a scenario in their program assessment plan (Knight et al., 2017; Ogilvie et al., 2015). Jesiek et al. (2020) collected initial evidence of validity for the China SJT when they administered the China SJT to practicing engineers along with related instruments intended to serve as proxy measures of similar competencies and obtained preliminary evidence of criterion validity. This study found positive relationships between SJT performance and Chinese cultural knowledge, age, and years of work experience. Jesiek et al. (2020) call for future work that collects further evidence of validity for the GEC China SJT, specifically highlighting the need to administer the SJT to engineering students and compare their performance with that of the practicing engineers. Building on this prior work presents an important opportunity to expand our understanding of both GEC and related assessments and is the focus of the present study.

3 | METHODS

This study contributes to the ongoing work by Jesiek et al. (2020) in the collection of evidence of validity for the Global Engineering Competence SJTs. Specifically, we seek to contribute evidence of validity for the GEC China SJT by administering the assessment to engineering students who did and did not have experience studying in China and comparing their SJT performance to the data collected from practicing engineers. Our approach provides evidence of validity by comparing SJT performance to scores on related tests, to other external variables (e.g., amount of travel), and across groups with different global experiences (American Educational Research Association, 2014; Educational Testing Service, 2015).

3.1 | GEC China SJT

As documented in more detail in previous publications, the development process for the GEC China SJT followed three steps: (i) identify potential scenarios, (ii) generate responses to each scenario, and (iii) determine how to score the SJT (Jesiek et al., 2015; Jesiek et al., 2020; Jesiek, Zhu, et al., 2014). Scenarios were conceptualized following a review of the literature for examples of global case studies and by collecting critical incidents from practicing engineers. This list of scenarios was compiled into a smaller group of scenarios that are able to elicit both potentially effective and ineffective responses from readers based on further inputs from SMEs. In the scenario development phase of the project, SMEs included practicing engineers familiar with the Chinese context. Both native and non-native SMEs were consulted regarding the relevance and utility of the scenarios, allowing the research team to refine the scenarios to their final versions (Jesiek et al., 2020). A broad range of SMEs were included at this stage of the process to improve both the variation of scenarios collected and relevance of final scenarios selected for the SJT.

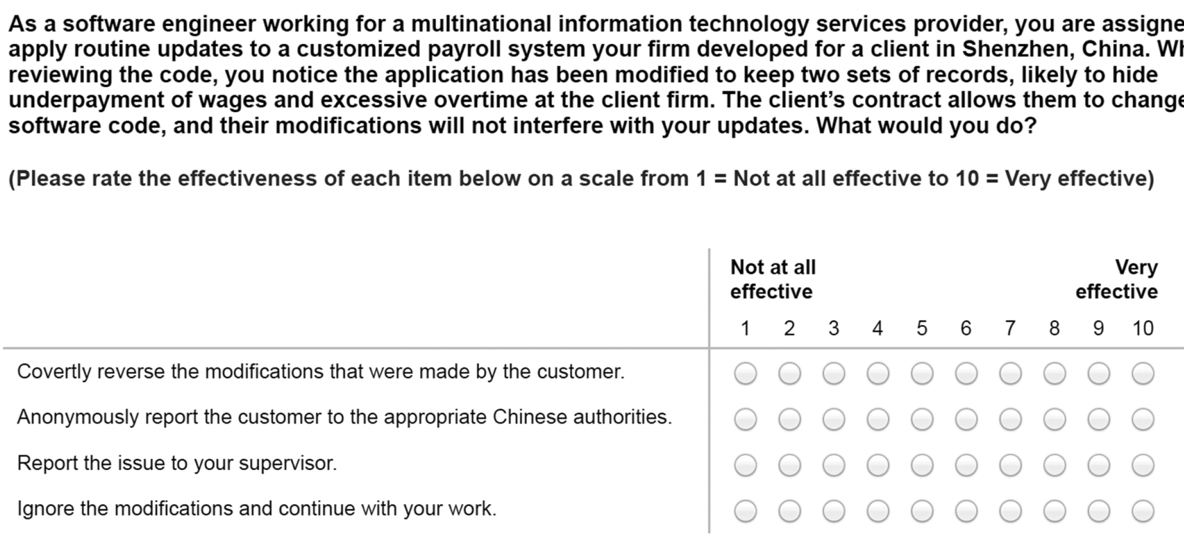
To generate responses, a different group of SMEs and a group of novice engineers were asked to describe how they would respond to each scenario. These potential responses were then tested with yet another group of SMEs who had multiple years of work experience in the Chinese context. These SMEs worked for multinational corporations, had "extensive interactions with Chinese contacts, frequent travel to China, work assignments in China, etc.," and 50% of them had lived in China for a period of at least 3 months (Jesiek et al., 2020). It is important to note that these SMEs were not Chinese themselves, but rather non-Chinese engineers working in the Chinese context. This decision was made because the SJT was designed to be used with this population rather than native Chinese engineers. During the SJT development process, it became clear that the best response in a particular situation may not always be the same for a native and non-native engineer in a given context. Thus, it was important that the SMEs who were rating the effectiveness of the responses aligned with the target population of the GEC SJT.

The resulting instrument consists of six scenarios (two aligned with each dimension of the GEC model described previously), each with four or five possible responses for participants to consider. Participants are asked to rate each possible response on a scale from 1 ("not effective at all") to 10 ("very effective"). For each item, a "best practice" rating was identified based on the responses from SMEs. Participant responses are scored by finding the difference between their rating and the average SME rating on each item, squaring these differences, finding the average of the squared values across all items, and multiplying these values by −1 so that the scores closest to zero are the closest to the SME scores. This composite score is the value used to compare participants' SJT performance with their responses on other related survey items and demographic questions (Jesiek et al., 2020). An example of one of the SJT scenarios from the GEC China SJT is shown in Figure 1.

3.2 | Data collection

The survey for our study is a shortened version of what was used in the Jesiek et al. (2020) study. That study included both the GEC China SJT and a wide variety of other self-report items that measure self-reported competencies, which potentially could relate to GEC for comparison purposes. As noted, the prior study's results suggested that several self-reported constructs did not relate to SJT performance as anticipated. For this reason, coupled with the fact that their survey took on average nearly 30 min to complete, we used a reduced form of the survey with the components shown in Table 2 in the present study.

The survey was designed to explore possible relationships between SJT performance and a variety of other potentially related factors such as personality, universal–diverse orientation, cultural intelligence, and individual background


As a software engineer working for a multinational information technology services provider, you are assigned to apply routine updates to a customized payroll system your firm developed for a client in Shenzhen, China. While reviewing the code, you notice the application has been modified to keep two sets of records, likely to hide underpayment of wages and excessive overtime at the client firm. The client's contract allows them to change the software code, and their modifications will not interfere with your updates. What would you do?

(Please rate the effectiveness of each item below on a scale from 1 = Not at all effective to 10 = Very effective)

	Not at all effective 1	2	3	4	5	6	7	8	9	Very effective 10
Covertly reverse the modifications that were made by the customer.	○	○	○	○	○	○	○	○	○	○
Anonymously report the customer to the appropriate Chinese authorities.	○	○	○	○	○	○	○	○	○	○
Report the issue to your supervisor.	○	○	○	○	○	○	○	○	○	○
Ignore the modifications and continue with your work.	○	○	○	○	○	○	○	○	○	○

FIGURE 1 Example of a situational judgment test (SJT) scenario from the GEC China SJT for the ethics, standards, and regulations dimension of GEC.

TABLE 2 Instruments included in the survey.

Instrument	Construct measured	Sample item	Items
GEC China SJT	GEC, in three dimensions (refer to literature review)	Refer to Figure 1	26
China Cultural Questions	Knowledge of Chinese Culture	How should you give a coworker work-related advice in China? (Multiple choice)	7
GEC Scale	GEC, in two scales: Cognitive and behavioral	I am familiar with cross-national/cultural differences in engineering practice. (Likert scale)	12
Cultural Intelligence Scale	CQ, in four dimensions: Cognitive, Metacognitive, Behavioral, Motivation	I am confident that I can socialize with locals in a culture that is unfamiliar to me. (Likert scale)	20
Miville–Guzman Universality–Diversity Scale (Short)	Universal–Diverse Orientation, three scales: Diversity of Contact, Relativistic Appreciation, Comfort with Differences	I attend events where I might get to know people from different racial backgrounds. (Likert scale)	15
Big Five Inventory	Big Five personality traits: Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness	I am someone who tends to be disorganized. (Likert scale)	44
Demographic Questions	Age, gender, race, ethnicity, nationality, work experience, personal travel, languages	Please select the types of engineering work experience you have had so far in your career: *None, Summer Internship, Part-Time Work, Co-op, Full-Time Work, Other*	8
Global Program Questions	Type of program, number of countries visited, length of trip, academic content	Please select the types of global programs you have participated in during college: *Short-Term Study Abroad (<3 months), Internship Abroad, [...]*	4
Academic Context Questions	Major, GPA, institution, location of institution, year in school	Please list your intended major or majors. (Open response)	5

characteristics. Universal–diverse orientation and cultural intelligence were selected because they are frequently used in research on international education as methods for more generally assessing global competence (Ogden, 2015). Further, research has found evidence of connections between these factors and individual performance in cross-cultural environments (Groves & Feyerham, 2011; Van Dyne et al., 2017). Similarly, prior research has suggested that personality (Caligiuri & Tarique, 2012; Huang et al., 2005), prior travel experiences (Bhaskar-Shrinivas et al., 2005; Caligiuri & Tarique, 2012), and cultural knowledge (Deardorff, 2006) may be predictors of intercultural competence and/or intercultural adjustment. In addition to these factors, we included academic and demographic questions to help determine the composition of the sample so that potential transferability to other settings can be determined.

The Chinese cultural knowledge questions and GEC Scale (GECS) were developed by Jesiek et al. (2020). The former was developed based on the SME interviews described previously and refined through an iterative process with collaborators of Chinese and non-Chinese descent (Jesiek et al., 2020). The latter is a self-report survey based on the GEC model with two dimensions (cognitive and behavioral) that were identified through exploratory and confirmatory factor analysis (Mazzurco et al., 2020). The remaining three sets of items are all widely used surveys with extensive evidence of validity. The Cultural Intelligence Scale (CQS) assesses four dimensions identified in the literature on Cultural Intelligence (CQ), and evidence of validity has been provided for use with college students across multiple cultural contexts (Ang et al., 2007). The MGUDS is a measure of universal–diverse orientation (UDO), which reflects an individual's awareness and acceptance of similarities and differences among people (Miville et al., 1999). This survey has been used extensively, significant evidence of validity has been collected, and a short version has been developed through exploratory and confirmatory factor analysis (Fuertes et al., 2000; Miville et al., 1999). The Big Five Inventory (BFI) is one of the major inventories based on Big Five personality theory, which was developed through the synthesis of decades of personality research (Goldberg, 1992; John et al., 2008). Despite its many items, the developers of the BFI report that it takes only 5 min to complete and suggest it as the best choice for researchers desiring a quick but thorough assessment of the Big Five personality traits (John et al., 2008). These scales were chosen because, based on literature and prior research, they were hypothesized to relate to participants' SJT performance.

We used several existing instruments where extensive evidence of validity has been established as comparison variables for the China SJT. Because it is important to check for measurement invariance when using an instrument across different populations (Gallagher & Brown, 2013), we conducted confirmatory factor analysis (CFA) on BFI, CQS, GECS, and the MGUDS – Short Form (MGUDS-S). In most cases, typical measures of fit (i.e., Comparative Fit Index [CFI], Tucker Lewis Index [TLI], root mean square error of approximation [RMSEA]) indicated that the scale model of the instruments aligned with the data at a good or excellent level. The GECS had a high RMSEA (e.g., >0.1), but otherwise we observed good fit indicators. We also calculated Chronbach's alpha for the scales on each instrument, and all were above the target value of 0.7 (Field et al., 2012; Gallagher & Brown, 2013).

3.3 | Participants and sampling strategy

For this study, we used existing data from the Jesiek et al. (2020) study of engineering professionals and collected new data from students. Jesiek et al. (2020) administered the survey to 400 practicing engineers at multiple companies. These participants were recruited by a professional recruiting company, Qualtrics, LLC, and were compensated for their participation. Participants were engineers working in a variety of industries (e.g., manufacturing, defense, information technology) and represented a wide variety of engineering disciplines and levels of experience. The participants had been in the workforce for an average of 22.25 years (SD = 13.58) and, on average, were 46.09 years of age (SD = 13.72). Participants came from a variety of company sizes (36% from companies smaller than 250 employees, 39% from companies over 3000 employees), and 42% reported post-graduate education. The sample was overrepresented in both female respondents (23%) and White respondents (83%) compared to representation in the engineering profession nationally (15% and 63%, respectively). Participants were recruited from companies based in the United States but had varying levels of experience working in the Chinese context (described subsequently).

For the new engineering student data, we collected 243 complete and usable responses. Because one goal of the study was to make comparisons across students who have a variety of global experiences, we used purposive sampling to recruit participants based on their global experiences in three groups: (i) students who have participated in global programs in China ($n = 46$); (ii) students who have participated in global programs in other contexts ($n = 104$); and (iii) students who have not participated in a global program ($n = 93$). To be counted in the *China Programs* or *Global Programs* groups, students needed to have traveled outside the country for at least 1 week as a part of a global education program (not personal travel) which may or may not have been engineering-focused. Students were recruited from US-based universities and primarily held US nationality. Only one participant was a Chinese citizen, and four reported being fluent in Mandarin; these students were grouped according to their program participation, the same as the other participants. Recruitment took place via an email that included a link to the survey, which was typically sent by the coordinator for the program in which the students participated to help boost response rates. Students for the *No Global Program* group were recruited through non-global campus programs (e.g., student organizations) and university Facebook pages. The final sample included students from 19 different universities across the United States, although students from two universities make up a majority of the sample. All students who completed the survey and provided their email address received a $10 Amazon gift card. Table 3 shows the composition of participants for this study based on demographic and global experience variables. The variables shown in Table 3 are those that were prominent later in the Results, and additional demographic information can be found in the Appendix (Table A1).

This study includes a higher percentage of women than most US engineering programs. This breakdown is unsurprising because women are overrepresented in study abroad programs but limits our ability to generalize our findings to all engineering students. There are also more women in the student sample than in the Jesiek et al. (2020) study, which is important to note in comparing the two datasets. Additionally, the student sample has a higher percentage of White students and lower percentage of Black and Hispanic/Latinx students compared to bachelor's degree recipients in engineering nationally (National Science Board, 2019), likely related to the recruitment populations. Recruitment for the *China Programs* and *Global Programs* groups targeted global programs of different types and lengths, with some success. Nevertheless, because short-term programs make up a significant portion of study abroad participation (and especially in engineering), it is unsurprising that a majority of students in the study participated in programs less than

TABLE 3 Characteristics of student participants ($n = 243$).

Variables	% of participants
Gender	
Women	46.1%
Men	53.5%
Not reported	0.4%
Prior engineering work experience	
Yes	70.0%
No	30.0%
Summer internship	52.2%
Part-time work	16.9%
Co-op	14.0%
Prior international travel (both personal and educational)	
I have never left my home country	6.2%
I have left my home country a few times (1–3) for short-term trips (each trip lasting 2 months or less)	41.6%
I have left my home country several times (4+) for short-term trips	26.8%
I have spent an extended period of time outside my home country (more than 2 months)	14.8%
I have lived outside my home country for at least a year	4.1%
I have lived outside my home country for several years	6.6%
Time abroad (on educational programs in college)	
None	38.3%
Less than a month	37.9%
1–3 months	11.9%
4–10 months	9.1%
A year	1.2%
More than a year	1.7%

3 months in length, which was consistent across program groups. Lastly, it is important to note that because recruitment was conducted based on convenience and ease of access to the researchers, most participants came from a limited number of institutions.

3.4 | Data analysis

To build on the work of Jesiek et al. (2020), we conducted analyses with data collected from engineering students similar in approach to their analyses of data from practicing engineers. We explored the relationships between the SJT performance score and potentially related variables by calculating bivariate correlations among all the continuous (using Pearson correlation) and ordinal (using Spearman correlation) variables collected for the study.

The second goal of this study was to explore differences in students' SJT performance based on their prior experiences and personal characteristics. First, we explored this question by considering differences between the three student groups (China Programs, Global Programs, No Global Program) using ANOVA to compare SJT performance across these groups. Second, we conducted three regression analyses to explore how demographic variables, global experience variables, and self-report measures relate to SJT performance. Decisions about which variables to include in the regression analyses were made based on results from prior research (including the Jesiek et al., 2020 study) and the theoretical conceptualization of GEC. All of the available variables could not be included in the regression analyses because of the sample size.

The third goal of this study was to explore differences in performance on the GEC China SJT between students and practicing engineers. We ran independent-samples *t*-tests to determine whether there were significant differences between students' and practicing engineers' overall SJT performance and scores on each of the six individual scenarios. Next, we broke the practicing engineers into three groups related to global experience and experience in China (similar to the students). Table 4 shows these groups, how they were defined, and the percentage of students for each group for comparison.

We used ANOVA to test for differences in SJT performance of students and practicing engineers across the three global experience groups shown in Table 4. Where differences were identified between groups, we conducted post hoc analyses using paired *t*-tests to determine which groups were significantly different in their SJT performance. Lastly, we ran identical regression analyses with the common variables between the two datasets—one analysis with the engineers dataset and one with the students dataset—to investigate whether these groups reveal different predictors of SJT performance. Before conducing the ANOVAs for this study, we used Levene's test to check for homogeneity of variance and identified that this was an issue for several of the variables (including overall score on the China SJT). Therefore, we used Welch's-F rather than the typical F-statistic (Field et al., 2012). For the regression analysis, we checked for multicollinearity by calculating the variance inflation factor (VIF). For two variables (*Number of Global Programs in College* and *Number of Countries Visited in College*), the VIF was near 5 and these variables were strongly correlated with each other and the *Time Abroad* variable ($r_s > 0.8$), so we dropped these variables from the analysis (Field et al., 2012).

3.5 | Limitations in data collection

There are several limitations to the data collection approaches used in this study. First, although the Jesiek et al. (2020) study provides preliminary evidence of validity and our study contributes to this process, the early stage of the process to collect validity evidence should be taken into consideration when interpreting results. Second, we used closed-ended questions to try to describe global experiences which may contribute to SJT performance. However, individuals may engage in the same global program together and have vastly different experiences and learn entirely different attitudes, knowledge, or skills (Davis, 2020; Davis & Knight, 2021). These nuances are lost in this quantitative analysis. Further, it is important to acknowledge that, even though using the SJT format is closer to assessing GEC, it still does not directly assess the participants' competence but only their awareness of how to behave in certain situations. Finally, this study is limited by its focus on the Chinese context alone because of the availability of the China SJT. Although SJTs in other contexts are under development, they were not ready at the start of this study. Therefore, the results may not be generalizable to other contexts with which students or engineers may be more familiar.

There are additional limitations to the sample of participants included in this study. Most notably, the majority of the participants were recruited from two universities with an overrepresentation of certain engineering majors (e.g., mechanical and engineering management are overrepresented). Although these two institutions represent different types (large public and small private) and our goal is not to make any claims about a program but rather to investigate the alignment between SJT and other measures, the institutional sample nevertheless limits the general experiences of the students in the sample. Similarly, the sample has an overrepresentation of White and female students

TABLE 4 Global experiences of practicing engineers and student survey respondents.

Engineer Group	Engineer Group Description	Percentage of practicing engineers	Comparable Student Group	Percentage of students
No/little professional experience abroad	Indicated that they had never lived abroad	36.8%	No Global Program	38.3%
Professional experience abroad	Indicated that they had lived abroad for any length of time	52.9%	Global Programs	42.8%
Professional experience in Chinese context/culture	Indicated that they either frequently attended professional meetings in China or spoke Chinese	10.2%	China Programs	18.9%

compared to engineering student populations generally. This characteristic of the sample is not surprising given that these groups are overrepresented in global education programs; however, it limits generalizability. Finally, this study is not an experimental design, and therefore we are not able to make causal claims about the influence of particular variables on SJT performance.

3.6 | Researcher positionality

The three authors of this paper are all US citizens and have been involved in coordinating global programs for engineering students. Our personal motivation in conducting this study was to better understand student learning in such programs while acknowledging time limitations of in-depth qualitative analysis for large-scale assessment. Our intent in designing the survey was to capture as much detail as was feasible about the prior experiences and identities of participants, but we acknowledge that the quantitative nature of these data reduces our understanding of participants' experiences and contexts. We reviewed the study design decisions with other researchers and prior research, but acknowledge that our own experiences and identities will always be present in the way we conduct a study.

4 | RESULTS

In this section, we address each research question in turn. The first two research questions were addressed using the dataset with only engineering students, and the analysis for the third research question also included data from the practicing engineers from Jesiek et al. (2020).

4.1 | RQ1: Relationship between SJT performance and similar instruments

Table 5 shows the Pearson correlation matrix comparing SJT performance to participants' scores on the BFI, MGUDS-S, CQS, GECS), and a series of seven questions focused on Chinese culture. Students' SJT performance is not significantly correlated with any of the dimensions from the BFI, MGUDS-S, CQS, GECS, or with their scores on the Chinese cultural questions. However, several of the self-report variables are significantly correlated with each other, in some cases quite strongly. For example, the CQS and MGUDS-S scores have several medium and strong correlations between their dimensions. The CQS *Motivational* dimension in particular is significantly correlated with most of the other variables in the study. Similarly, the *Agreeableness* dimension from the BFI is significantly correlated with both the other BFI dimensions and all the MGUDS-S dimensions. It is notable that the GECS scale is divided between the *Cognitive* dimension, which is only significantly correlated with the CQS scores, and the *Behavioral* dimension, which is also significantly correlated with the MGUDS-S scores. Overall, the self-report comparison instruments included in this study seem to be assessing related constructs to one another, but these instruments are not significantly correlated with either SJT performance or responses on the Chinese cultural knowledge questions. It is particularly notable that SJT performance is not significantly correlated with the Chinese cultural knowledge questions, in contrast to the findings of Jesiek et al. (2020), and also suggests that Chinese cultural knowledge may not assist participants in responding to the SJT scenarios.

4.2 | RQ2: Relationship between SJT performance and international experiences

Table 6 shows the Spearman correlations comparing SJT performance and scores on the global-competency-related instruments from the survey (MGUDS-S, CQS, GECS, Chinese culture questions) with five ordinal variables related to students' international experiences. *Total International Travel* refers to the self-reported time spent outside home country across a participants' entire life, which ranges from "I have never left my home country" to "I have lived outside my home country for several years." *Number of Global Programs in College* specifically counts reported participation in global education programs (either none, 1, 2, or >2). *Time Abroad in College* refers to time spent outside home country as a part of global programs in college, which ranges from "None" to "More than a Year." *Number of Countries Visited in College* counts countries on global education programs (either none, 1, 2, or >2). Lastly, *Time in China in College* is similar to the Time Abroad variable, but focuses only on time in China.

TABLE 5 Pearson correlations for continuous variables (RQ1).

Variable	n	M	SD	1	2A	2B	2C	2D	2E	3A	3B	3C	4A	4B	4C	4D	5A	5B	6
1—GEC SJT China Score	243	−6.19	2.90	–															
2A—Big Five Personality Extraversion	240	3.29	0.84	−0.15	–														
2B—Big Five Personality Agreeableness	242	3.96	0.57	−0.02	0.24*	–													
2C—Big Five Personality Conscientiousness	242	3.95	0.61	0.06	0.06	0.34***	–												
2D—Big Five Personality Neuroticism	242	2.69	0.77	0.03	−0.20	−0.36	−0.28**	–											
2E—Big Five Personality Openness	240	3.69	0.54	−0.08	0.21	0.26**	0.11	−0.16	–										
3A—MGUDS-S Diversity of Contact	241	4.38	0.89	0.05	0.25**	0.33***	0.14	−0.16	0.34***	–									
3B—MGUDS-S Relativistic Appreciation	242	4.90	0.70	−0.06	0.18	0.38***	0.17	−0.14	0.35***	0.65***	–								
3C—MGUDS-S Comfort with Differences	241	5.02	0.71	−0.04	0.12	0.25**	0.16	−0.16	0.12	0.45***	0.42***	–							
4A—CQS Cognitive	242	3.49	1.15	−0.07	0.14	0.00	0.03	−0.16	0.16	0.30***	0.13	0.08	–						
4B—CQS Meta-Cognitive	241	4.80	1.17	0.03	0.10	0.22	0.20	−0.24*	0.33***	0.49***	0.40***	0.31***	0.52***	–					
4C—CQS Behavioral	242	4.50	1.39	0.01	0.14	0.17	0.08	−0.14	0.21	0.37***	0.36***	0.19	0.34***	0.56***	–				
4D—CQS Motivational	243	5.20	1.16	−0.02	0.24*	0.35***	0.16	−0.25**	0.38***	0.66***	0.51***	0.46***	0.42***	0.59***	0.45***	–			
5A—GECS Cognitive	243	2.94	0.86	0.00	0.02	−0.01	0.00	−0.14	0.12	0.16	0.10	−0.02	0.48***	0.37***	0.30***	0.24*	–		
5B—GECS Behavioral	242	4.01	0.58	−0.05	0.04	0.18	0.19	−0.16	0.19	0.32***	0.31***	0.26**	0.32***	0.40***	0.34***	0.48***	0.32***	–	
6—China Cultural Knowledge Questions	243	3.16	1.43	0.15	−0.13	−0.03	−0.13	0.03	0.00	0.10	0.08	0.09	0.07	0.08	0.00	0.04	0.12	0.08	–

*$p < .05$; **$p < .01$; ***$p >< .001$.

TABLE 6 Spearman correlations for ordinal variables (RQ2).

Variable	*n*	*M*	SD	1	2A	2B	2C	2D	2E	3A	3B	3C	4A	4B	4C	4D	5A	5B	6	7A	7B	7C	7D	7E
1—GEC SJT China Score	243	−6.19	2.90	–																				
7A—Total International Travel	243	1.89	1.25	0.07	0.04	0.02	−0.01	−0.08	−0.05	0.29***	0.08	0.07	0.35***	0.31***	0.17	0.29***	0.20	0.21	0.12	–				
7B—Number of Global Programs in College	243	0.80	0.78	0.20	0.04	0.17	0.07	−0.19	0.10	0.30***	0.18	0.14	0.29***	0.41***	0.30***	0.38***	0.25**	0.32***	0.16	0.42***	–			
7C—Time Abroad in College	243	1.02	1.12	0.21	0.00	0.12	0.06	−0.13	0.07	0.29***	0.15	0.12	0.27**	0.38***	0.25*	0.34***	0.22	0.27**	0.17	0.52***	0.90***	–		
7D—Number of Countries Visited in College	243	1.09	1.10	0.16	0.04	0.14	0.06	−0.11	0.09	0.26***	0.16	0.08	0.24*	0.37***	0.28**	0.30***	0.25**	0.29***	0.12	0.33***	0.89***	0.82***	–	
7E—Time in China in College	243	0.28	0.67	0.17	−0.03	−0.07	−0.12	−0.03	−0.05	0.04	−0.02	0.00	0.27**	0.17	0.13	0.19	0.22	0.13	0.31***	0.26**	0.38***	0.38***	0.21	–

Note: 2A - Big Five Personality Extraversion, 2B - Big Five Personality Agreeableness, 2C - Big Five Personality Conscientiousness, 2D - Big Five Personality Neuroticism, 2E - Big Five Personality Openness, 3A - MGUDS-S Diversity of Contact, 3B - MGUDS-S Relativistic Appreciation, 3C - MGUDS-S Comfort with Differences, 4A - CQS Cognitive, 4B - CQS Meta-Cognitive, 4C - CQS Behavioral, 4D - CQS Motivational, 5A - GECS Cognitive, 5B - GECS Behavioral, 6 - China Cultural Knowledge Questions.
*$p < .05$; **$p < .01$; ***$p < .001$.

Once again, participants' SJT performance is not significantly correlated with the other variables in this analysis. However, the international experience variables (7A–7E in Table 6) are significantly correlated with many of the self-report survey variables; in particular, *Time in China* is significantly correlated with the participants' scores on the Chinese cultural knowledge questions (and this is the only significant correlation with those questions). These correlations are not as strong as those among the self-report surveys. Some exceptions are the CQS *Meta-Cognitive* and *Motivational* dimensions, which have medium-strength correlations with several international experience variables, and *Number of Global Programs in College*, which is correlated with most of the survey variables. These findings align with prior studies, where these self-report surveys have been used to assess student development during global programs. It is notable, however, that the MGUDS-S dimensions *Relativistic Appreciation* and *Comfort with Differences* are not significantly correlated with any of the international experience variables.

Students were recruited for the study who fell into three different groups based on their international experiences while in college: those who had participated in a global program in China (*China Programs*), those who had participated in a global program outside China (*Global Programs*), and those who had participated in *No Global Programs*. We therefore also compared participants' SJT performance across these three groups and found significant differences across groups (F [2139] = 9.31, $p < .001$). Post hoc analyses revealed that the differences are significant both between the *China Programs* and *No Global Programs* groups ($p < .001$) and between the *Global Programs* and *No Global Programs* groups ($p = .002$). Figure 2 summarizes the average scores for each group and how these scores compare with the score of SMEs consulted in the development of the GEC China SJT. The average SJT score across all student participants was −6.19 (SD = 2.90), with a range from −19.5 to −1.62. To achieve this average score, a student would choose options on the SJT items that were 2–3 points away from the SME choice (on a scale from 1 to 10). In comparison, students performing better than the average were 1–2 points from SME choices, and students below the average were 4–6 points from SME choices.

Three regression analyses explore the relationship between SJT performance and various demographic, international experience, and survey variables (refer to Table 7). Variables were selected for inclusion in the analysis based on prior research results and the GEC theoretical framework. Because of limited sample size and to run a parsimonious model, only certain dimensions of the surveys were included (e.g., only the *Cognitive* and *Behavioral* dimensions of the CQS); these dimensions are most closely related to the types of knowledge we believe to be measured by the GEC China SJT. The results of the earlier Jesiek et al. (2020) study were also consulted in making these determinations. Demographic variables were added in the first model, the international experience variables in the second model, and the survey variables in the third model. We did not include race/ethnicity in the model because of the overall lack of diversity of the sample with respect to race/ethnicity.

In the first model, all independent variables are significantly related to SJT performance except for gender, and the model has an adjusted R^2 value of 0.05. *Age* and *GPA* are positively related to SJT performance, but a binary variable describing whether students have had prior *Engineering Experiences* (e.g., internships) is negatively related to SJT performance. The adjusted R^2 value of the second model increased to 0.10. Both *GPA* and *Engineering Experiences* remain significantly related to SJT performance, and a binary variable describing students' participation in any *Global Programs*

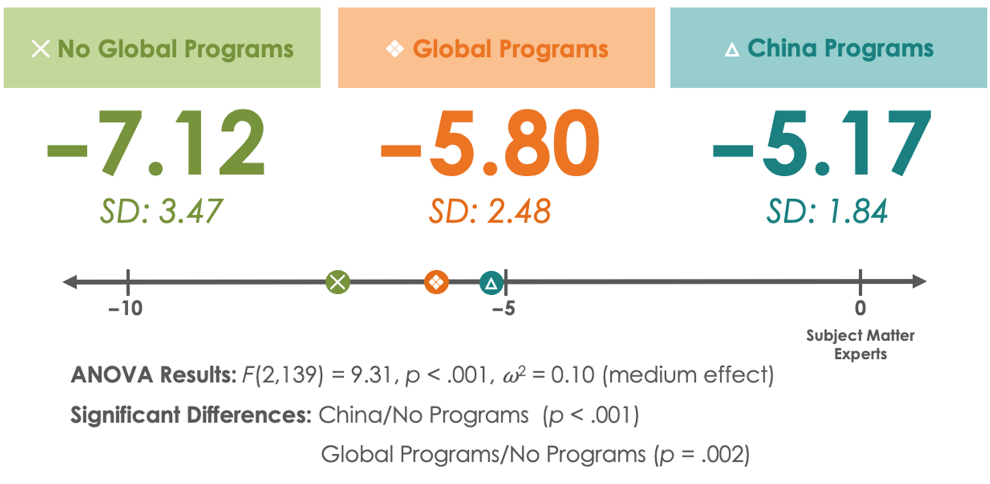


FIGURE 2 Average SJT performance across student groups. Scores closer to zero represent closer alignment with the subject-matter experts scoring key and therefore indicate better performance on the GEC-SJT instrument.

TABLE 7 Regression models of students' SJT performance.

Variable	Regression 1 Beta	SE	p-Value	Regression 2 Beta	SE	p-Value	Regression 3 Beta	SE	p-Value
(Intercept)		3.45	<.001***		3.59	<.001***		4.32	.014*
Gender (Women = 1)	−0.12	0.36	.076	−0.12	0.36	.053	−0.13	0.41	.080
Age	0.17	0.15	.015*	0.12	0.15	.094	0.11	0.16	.148
GPA	0.17	0.41	.008**	0.16	0.42	.014*	0.18	0.45	.009**
Engineering experience (Yes = 1)	−0.15	0.42	.030*	−0.14	0.41	.044*	−0.13	0.44	.068
Prior travel				−0.07	0.17	.332	−0.05	0.18	.529
Global programs (Yes = 1)				0.21	0.54	.023*	0.22	0.59	.030*
Time abroad				−0.01	0.26	.926	0.00	0.26	.970
Programs in China (Yes = 1)				0.11	0.50	.117	0.08	0.55	.319
China cultural questions							0.06	0.14	.388
Extraversion							−0.11	0.24	.107
Agreeableness							0.01	0.35	.912
Openness							−0.01	0.38	.876
CQS—cognitive							−0.04	0.21	.635
CQS—behavioral							0.03	0.15	.647
GECS—cognitive							0.06	0.26	.471
GECS—behavioral							−0.17	0.37	.026*
MGUDS-S Diversity of contact							0.02	0.27	.812
Adjusted R^2	0.05**			0.10***			0.10**		

*$p < .05$; **$p < .01$; ***$p < .001$.

in college is also significant. The adjusted R^2 value remained 0.10 in the full model. In this model, both *GPA* and *Global Programs* variables remained significant predictors from the prior models, but *Engineering Experience* was no longer significant. One of the survey variables (GECS *Behavioral* dimension) was also a significant predictor, albeit in the negative direction. This finding aligns with the results of the Jesiek et al. (2020) study, which found negative relationships between SJT performance and several global competency-related survey variables.

4.3 | RQ3: Comparing SJT performance between students and engineers

The students have better average SJT performance ($M = -6.19$, SD = 2.90) than the professional engineers ($M = -6.76$, SD = 4.03) in the combined dataset according to a *t*-test analysis ($t = -2.08$, $df = 618$, $p = .038$), but the effect size is small (Cohen's $d = 0.17$).

We also compared the engineers' SJT performance across the three groups based on their international work experience (Figure 3). The results of this analysis run counter to what we found in the student data: engineers who have *Traveled to China* for work have the most different SJT performance from the SMEs, those who have *Lived Abroad* elsewhere are next, and those who have *Not Lived Abroad* or traveled to China have the most similar SJT performance (F [2104] = 6.08, $p = .003$); post hoc analyses reveal significant differences between the *Travel to China* and *Not Lived Abroad* groups ($p = .002$). The average SJT score across all engineers was −6.76 (SD = 4.03), with a range from −28.66 to −1.53. To achieve the average score, an engineer would choose options on the SJT items that were 2–3 points away from the SME choice (on a scale from 1 to 10). In comparison, engineers performing better than average were 1–2 points from SME choices and those below average were 4–6 points from SME choices.

Finally, we ran two regression analyses with common variables—one using the student data and the other using the practicing engineer data. Similar to the earlier analysis with student data, we ran three regression analyses for each dataset to explore the relationship between SJT performance and demographics (first model), global experiences (second model), and survey variables (third model). Variables were selected for inclusion in the analyses based on which variables were available in both the student and practicing engineer datasets. We were able to include all of the survey dimensions for the CQS and GECS for the students in this case because we did not include the *GPA*, *Engineering Experience*, *Prior Travel*, *Time Abroad*, and *Programs in China* variables from Table 7 since they did not exist in the practicing engineer dataset. In the first model, *Gender* and *Age* are significant for both groups, but *Age* is a stronger predictor for the engineers, and the model for the students is not significant ($p = .052$). The second regression model indicates that *Global Programs* are significant predictors of students' SJT performance, whereas *Age* remains the main predictor for the engineers; global experiences have a negative relationship with SJT performance for the engineers. In the second model, the students' adjusted R^2 value jumps from 0.02 to 0.08, and the model is now significant ($p < .001$). In contrast, the final model for engineers sees a significant jump in the adjusted R^2 from 0.07 to 0.27. Although *Gender* and *Age* remained significant, two of the survey variables (*China Cultural Questions* in a positive direction and CQS *Cognitive* in a negative direction) were the strongest predictors of SJT performance for the engineers. For the students' model, adding the survey variables contributed little to model performance; global program participation remained a significant predictor of SJT performance. These results are summarized in Tables 8 and 9.

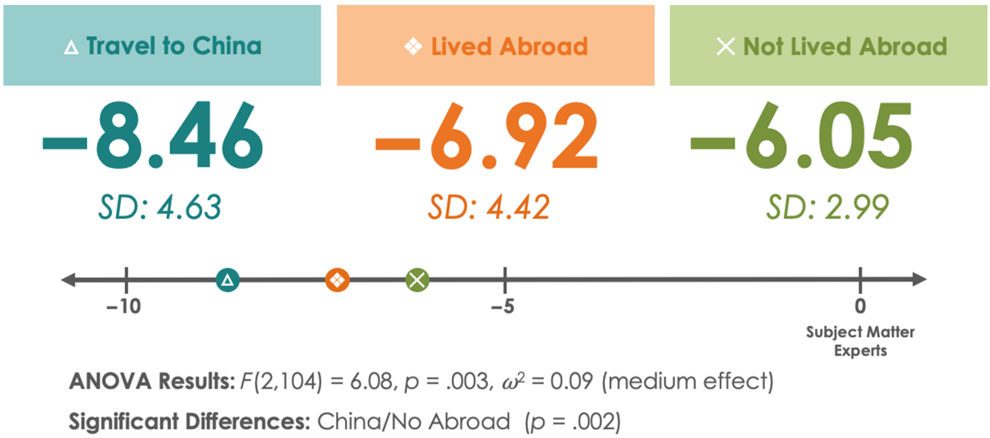


FIGURE 3 Average SJT performance across professional engineer groups. Scores closer to zero represent closer alignment with the subject-matter experts scoring key and therefore indicated better performance on the GEC-SJT instrument.

TABLE 8 Regression models of engineer SJT performance for comparison.

	Regression 1			Regression 2			Regression 3		
Variable	**Beta**	**SE**	***p*-Value**	**Beta**	**SE**	***p*-Value**	**Beta**	**SE**	***p*-Value**
(Intercept)		0.78	<.001***		0.84	<.001***		1.59	.010
Gender (Women = 1)	0.11	0.51	.041*	0.09	0.50	.087	0.09	0.45	.042*
Age	0.24	0.02	<.001***	0.22	0.02	<.001***	0.12	0.01	.016*
Global experience (Yes = 1)				−0.11	0.42	.029*	−0.05	0.40	.314
Experience in China (Yes = 1)				−0.11	0.68	.037*	−0.08	0.61	.069
Second language (Yes = 1)				−0.03	0.52	.613	0.02	0.48	.627
China cultural questions							0.33	0.12	<.001***
CQS—cognitive							−0.24	0.29	<.001***
CQS—behavioral							−0.06	0.34	.243
CQS—meta-cognitive							−0.09	0.40	.124
CQS—motivational							−0.02	0.34	.730
GECS—cognitive							0.00	0.32	.984
GECS—behavioral							−0.05	0.37	.427
Adjusted R^2	0.05***			0.07***			0.27***		

*$p < .05$; ***$p < .001$.

TABLE 9 Regression models of student SJT performance for comparison.

Variable	Regression 1 Beta	Regression 1 SE	Regression 1 *p*-Value	Regression 2 Beta	Regression 2 SE	Regression 2 *p*-Value	Regression 3 Beta	Regression 3 SE	Regression 3 *p*-Value
(Intercept)		2.69	<.001***		2.60	.003**		3.01	.086
Gender (Women = 1)	0.13	0.37	.051	0.12	0.36	.062	0.11	0.37	.096
Age	0.08	0.13	.223	0.02	0.13	.789	0.00	0.14	.945
Global programs (Yes = 1)				0.19	0.41	.008**	0.25	0.45	.002**
Programs in China (Yes = 1)				0.14	0.50	.052	0.10	0.54	.179
Second language (Yes = 1)				−0.15	0.43	.023*	−0.12	0.47	.080
China cultural questions							0.13	0.13	.060
CQS—cognitive							−0.05	0.22	.594
CQS—behavioral							0.02	0.16	.778
CQS—meta-cognitive							−0.01	0.23	.914
CQS—motivational							−0.06	0.21	.526
GECS—cognitive							−0.04	0.26	.636
GECS—behavioral							−0.07	0.37	.382
Adjusted R^2	0.02			0.08***			0.08**		

*$p < .05$; **$p < .01$; ***$p < .001$.

5 | DISCUSSION

GEC and related constructs have been defined, modeled, and assessed in a variety of different ways. In many cases, assessment of global competence has relied on self-report instruments where participants rate themselves on a list of items using a Likert scale (e.g., IDI, GPI, CQS, MGUDS). These instruments are frequently the focus of studies in international education (Ogden, 2015), but such methods can only capture students' perceptions of their competence or learning, at best. Our study explored the use of a situational judgment test as an alternative approach for assessing outcomes related to GEC. Although the SJT is still not directly assessing student behavior, it assesses students' knowledge of behaviors that have been deemed appropriate in a given situation by SMEs. We compared SJT performance to scores on self-reported global competency instruments and participants' international experiences to continue the process of collecting evidence for the validity of this instrument.

A significant finding in this study is that students' performance on the GEC China SJT did not relate to their responses on self-report instruments (e.g., CQS, MGUDS). This finding contrasts with results from Jesiek et al.'s (2020) study, where the professional engineers' SJT performance negatively correlated with the self-report instruments. The results of both studies suggest that the GEC China SJT may be measuring a different underlying construct than the self-report instruments traditionally used to assess and research global programs. One possible explanation for this finding can be found in the weaknesses of self-report instruments already documented in the literature on student learning (e.g., Bowman, 2011; Porter, 2013). Although many of these critiques focus on self-reported learning gains, many of the same concerns could potentially apply to self-reported knowledge or competence. Our own work on engineering students' systems thinking competencies has shown that self-reports align with one another but do not align with students' performance on multiple scenario-based assessments of systems thinking (Davis et al., 2023). Along these lines, Anderson et al. (2017) suggest that SJTs may help mitigate instances of stereotype threat and response-style biases as compared to self-report surveys as a result of their less test-like structure and more interactive nature. However, further research will be necessary to determine the construct that the GEC China SJT is assessing and why it does not relate to self-report style assessments of global competence and related constructs. Ideally, studies could be conducted that compare SJT performance to more direct assessments of GEC, such as ratings from supervisors or performance while role-playing in global engineering scenarios. At the very least, our findings point out the importance of describing the specific construct that researchers and practitioners are seeking to measure (e.g., knowledge, attitude, skill, belief, and behavior) and identifying linkages between those constructs and the inherent limitations associated with different forms of measurement (e.g., self-reports, lab/no stakes setting like an SJT, behavioral measure in a real-stakes setting).

Notably, students' international experiences correlated with self-report instruments, which aligns with the many studies using these instruments to study the impact of global programs (Ogden, 2015; Vande Berg & Paige, 2012). In light of the weaknesses with using self-reports for assessment of competence, these results may be interpreted to suggest that students feel more positive, aware, and/or confident in dealing with cultural differences as a result of participation in global programs—however, that finding is a different interpretation than claiming that students are, in fact, more globally competent. All the self-report instruments significantly related to each other with fairly strong correlations. This finding is unsurprising given that similar types of instruments were likely used to provide evidence of validity for these instruments, but this correlation alone does not provide evidence that the instruments measure global competence. Such instruments are capturing perceptions of learning outcomes that occur during global programs, but these perceptions may not align with development of the actual learning outcomes or with actual behaviors when individuals are in cross-cultural situations.

Both the ANOVA and regression analyses revealed that SJT scores were closer to those of SMEs for students who had participated in global programs in college compared to those who had not. Because this result aligns with expectations based on theory and prior literature, it provides evidence of validity for the SJT. However, contrary to expectations, students who had traveled to China did not perform differently from students who had traveled elsewhere. Several models of intercultural competence have included knowledge of a specific cultural context as a potentially important dimension (Deardorff, 2006; Spitzberg & Changnon, 2009). This inconsistency may be a result of the fact that the GEC China SJT does not directly assess students' behaviors but only their knowledge of what behavior may work in a given situation. It is possible, for example, that when an individual actually interacts in a cross-cultural situation, specific contextual knowledge may be essential in their responding effectively. However, in selecting responses to the relatively short scenario situations in the GEC China SJT, a more general knowledge of cultural differences may be sufficient. Further exploration across individuals with greater variation in their experiences both abroad and in China may be helpful in interpreting this finding.

The most surprising result of this study was that the students performed more similarly to the SMEs than did the practicing engineers on the GEC China SJT. Further, engineers who had traveled to China for work performed significantly worse on the SJT than engineers who had never lived abroad. Upon further reflection, however, these findings may not be so surprising when viewed in light of study abroad research on cross-cultural learning. Engineers who have traveled and lived abroad could be said to have experienced an "immersion" in another culture, but possibly without the kinds of educational interventions that are viewed as central in modern study abroad programs (Ogden, 2015; Vande Berg & Paige, 2012). Studies in global education have found ample evidence that providing students with training before going abroad, cultural mentoring while abroad, and reflection opportunities upon their return home can all improve student outcomes (Vande Berg et al., 2012). The assumption that immersion is sufficient for learning has long since been debunked by global education researchers, so it should be no surprise that immersion on its own may not lead to significant global learning for practicing engineers. On the flip side, the significantly better SJT performance of students compared to engineers may be a result of their engagement with these types of global education interventions, and so they may be more attuned to thinking about intercultural differences. Only the students who had not participated in any global programs performed at a level comparable to the engineers.

5.1 | Limitations of this study

There were several limitations in this study which should be considered in interpreting the results. In particular, the student sample included a much larger percentage of women than the sample of practicing engineers. Prior research suggests that women tend to have higher levels of global competence (Jesiek et al., 2012), so it is possible that the different composition of these groups influenced the cross-group comparison results. Further, the student sample had a significant number of students from two universities, whereas the sample of practicing engineers was more diverse in terms of their backgrounds and affiliations. Future studies with a more diverse population of students will help build on the results here to determine whether they apply across engineering students more broadly. Lastly, this study focuses specifically on the Chinese context, which may be more familiar to a younger generation because of the increasing globalization of US universities. As GEC SJTs in other contexts become available, it will be important to conduct similar studies to see how those results compare with those found in this study.

5.2 | Implications for practice

The results of this study suggest several implications for individuals seeking to assess global engineering programs. First, it is important to use multiple forms of assessment, as it is not clear that self-report surveys that currently dominate evaluation approaches in global education provide the information that is typically claimed or desired (Deardorff, 2006; Fantini, 2009). Although this study suggests that scores on the self-report instruments correlate with students' time abroad, our findings do not indicate that they are measuring competence or knowledge, at least as measured by a scenario-based assessment. We do not have evidence yet to suggest that the GEC China SJT measures these outcomes either, but using multiple methods can provide program evaluators with a more holistic understanding of what students are experiencing and learning through their programs (Davis & Knight, 2023).

Second, the results comparing SJT performance of students and engineers provides preliminary support for the interventionist methods that have been advocated by global education researchers (Lou et al., 2012; Ogden, 2015; Vande Berg et al., 2009). Students' participation in organized global programs during college was related to SJT performance, whereas engineers' experiences both working and living abroad did not relate to SJT performance. This finding, if replicated in future work, may also suggest implications for engineering companies in terms of providing similar educational and training support for their employees who work abroad or collaborate regularly with team members or clients in other countries. Future research could explore this possibility by asking practicing engineers whether they received any training in global competence before engaging in global work environments.

5.3 | Implications for research

This study is just one step in the overall process of building evidence of validity for the GEC China SJT. Although comparing students' and engineers' SJT performance provides some insight into what this instrument may be measuring, other comparisons will also be important. For example, thus far the SJT has only been administered to individuals who are engineers or engineering students. To claim that this instrument assesses GEC specifically and not intercultural competence more generally, it will be important to conduct a study where the comparison groups also include non-engineering students and professionals. Non-engineers should theoretically not score as high if the SJT measures GEC, but similar scores would be expected if it measures a more general intercultural competence construct. This design would allow researchers to determine whether there is any engineering-specific knowledge required to respond to these scenarios. The current study was also limited to the Chinese context, given the availability of the GEC China SJT and its scoring rubric at the time of the study. However, GEC SJTs for other cultural contexts are now available, so future studies could include scenarios from across contexts. Future research that compares the same participants across multiple GEC SJTs would be helpful in interpreting the results of the current study. For example, it would be useful to know whether these SJTs assess knowledge that is context specific (i.e., participants perform differently across contexts) or general (i.e., participants perform similarly across contexts). This future research question is particularly motivated by the results of this study, where there were few differences between the students who had been to China and students who had been to other countries.

Pike (2011) suggests that when collecting evidence of validity for instruments in cases where direct measures of a construct are hard to obtain (as is true for GEC), it is important to rely on theory to interpret results. In this study and Jesiek et al.'s (2020) study, comparison measures were selected based on theory and prior work, but the results did not generally align with anticipated findings predicted by the theories. However, there are weaknesses with this approach in this case because the comparison measures are all self-report instruments, which collectively have known limitations that we are seeking to address through the creation of the SJT. Although it will be more complicated, it would be helpful if future studies include other forms of assessment beyond self-report surveys in their research designs as comparisons for SJT performance. In the simplest version, other SJTs or scenario-based assessments for related constructs could be included in lieu of self-report instruments. To understand the types of knowledge and experiences that participants draw on while completing the GEC SJTs, a study could have participants complete the scenarios and then participate in an interview in which they talk through their logic and reasoning for each scenario. It would be particularly helpful to compare students and engineers using this approach, as this study design may reveal insights into the performance differences found in this study. Finally, a moambitious project could attempt to obtain more direct measures of participants' actual GEC through supervisor performance reviews or role-playing scenarios. In summary, future evidence of validity collected for the GEC SJTs and similar measures should move beyond comparison with self-report surveys and instead incorporate a variety of data collection methods.

In addition to collecting evidence of validity for the scenarios, such evidence is also needed for the scoring keys used in the calculation of SJT performance. The current keys are based on both theoretical understandings of culture and the responses of SMEs; items were selected for inclusion based on alignment between the SME responses, which typically confirmed the theoretical expectations. These SMEs were engineers who were not Chinese but had a significant amount of professional experience working in the Chinese context (Jesiek et al., 2020). In our study, engineers who had more experiences living abroad and experience in the Chinese context performed worse on this instrument relative to engineers who had not lived abroad. It is possible that this finding occurred because the participants had less experience than the SMEs. Nevertheless, the results of this study call into question whether this method of developing the scoring key is ideal. Future research should pursue evidence of validity for the scoring keys by seeking input from other sources about the accuracy of the SME responses. For example, different types of SMEs could be included, such as researchers who have studied Chinese business practices. Further, the scenarios could be written from the opposite perspective and given to Chinese engineers to ask "how would you prefer an international colleague to respond in this situation?" A comparison of responses across these various types of SMEs would provide additional evidence of validity for the GEC SJT scoring keys.

6 | CONCLUSION

This study sought to provide evidence of validity for a scenario-based assessment of global engineering competency. The GEC China SJT was administered to 243 engineering students alongside several existing self-report instruments for global competence and questions about the students' prior international experiences. Analyses revealed that SJT performance was not related to the existing self-report instruments; students who had participated in global programs in college responded more similarly to SMEs compared to students who had not. Students were also found to respond more similarly to SMEs relative to practicing engineers based on data collected in an earlier study (Jesiek et al., 2020). In combination, these findings suggest that multiple forms of measurement of global competency, beyond self-report instruments, may be needed. Although collecting evidence of validity for the GEC China SJT is an ongoing process, this data collection technique may provide new insights on global engineering competency compared to traditionally used assessments. Future studies seeking to contribute additional evidence of validity for this assessment approach will need to employ alternative forms of assessment for comparison and should also incorporate SJTs from contexts other than China.

FUNDING INFORMATION

The data collected for the engineer participants was funded through the U.S. National Science Foundation under Grant Numbers 1160455 and 1254323.

DATA AVAILABILITY STATEMENT

The data are not available to share publicly.

ORCID

Kirsten A. Davis https://orcid.org/0000-0002-9929-5587
Brent K. Jesiek https://orcid.org/0000-0003-3056-5144
David B. Knight https://orcid.org/0000-0003-4576-2490

REFERENCES

Allert, B. I., Atkinson, D. L., Groll, E. A., & Hirleman, E. D. (2007). Making the case for global engineering: Building foreign language collaborations for designing, implementing, and assessing programs. *Online Journal for Global Engineering Education*, *2*(2), 1–14. https://digitalcommons.uri.edu/ojgee/vol2/iss2/1

American Educational Research Association. (2014). *Standards for educational and psychological testing*. https://www.aera.net/Publications/Books/Standards-for-Educational-Psychological-Testing-2014-Edition

Anderson, R., Thier, M., & Pitts, C. (2017). Interpersonal and intrapersonal skill assessment alternatives: Self-reports, situational-judgment tests, and discrete-choice experiments. *Learning and Individual Differences*, *53*, 47–60. https://doi.org/10.1016/j.lindif.2016.10.017

Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., & Chandrasekar, N. A. (2007). Cultural intelligence: Its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance. *Management and Organization Review*, *3*(3), 335–371. https://doi.org/10.1111/j.1740-8784.2007.00082.x

Bhaskar-Shrinivas, P., Harrison, D. A., Shaffer, M. A., & Luk, D. M. (2005). Input-based and time-based models of international adustment: Meta-analytic evidence and theoretical extensions. *Academy of Management Journal*, *48*(2), 257–281. https://doi.org/10.2307/20159655

Bowman, N. (2011). Examining systematic errors in predictors of college student self-reported gains. *New Directions for Institutional Research*, *2011*(150), 7–19. https://doi.org/10.1002/ir.386

Braskamp, L. A., Braskamp, D. C., & Merrill, K. C. (2009). Assessing progress in global learning and development of students with education abroad experiences. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *18*(1), 101–118. https://doi.org/10.36366/frontiers.v18i1.256

Caligiuri, P., & Tarique, I. (2012). Dynamic cross-cultural competenices and global leadership effectiveness. *Journal of World Business*, *47*(4), 612–622. https://doi.org/10.1016/j.jwb.2012.01.014

Davis, K. A. (2020). *Pursuing intentional design of global engineering programs: Understanding student experiences and learning outcomes* [Dissertation, Virginia Tech]. https://vtechworks.lib.vt.edu/handle/10919/97979

Davis, K. A., Grote, D. M., Mahmoudi, H., Perry, L., Ghaffarzadegan, N., Grohs, J. R., Hosseinichimeh, N., Knight, D. B., & Triantis, K. (2023). Comparing self-report assessments and scenario-based assessments of systems thinking competence. *Journal of Science Education and Technology*. https://doi.org/10.1007/s10956-023-10027-2

Davis, K. A., & Knight, D. B. (2021). Comparing students' study abroad experiences and outcomes across global contexts. *International Journal of Intercultural Relations*, *83*, 114–127. https://doi.org/10.1016/j.ijintrel.2021.05.003

Davis, K. A., & Knight, D. B. (2023). Assessing learning processes rather than outcomes: Using critical incidents to explore student learning abroad. *Higher Education*, *85*(2), 341–357. https://doi.org/10.1007/s10734-022-00836-6

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, *10*(3), 241–266. https://doi.org/10.1177/1028315306287002

DeBoer, J., Stump, G. S., Carter-Johnson, F., Allen, G., & Breslow, L. (2013). *Developing direct measures of global competence*. Paper presented at the ASEE Annual Conference and Exposition, Atlanta, GA. http://doi.org/10.18260/1-2–19412

Downey, G. L., Lucena, J. C., Moskal, B. M., Parkhurst, R., Bigley, T., Hays, C., Jesiek, B. K., Kelly, L., Miller, J., Ruff, S., Lehr, J. L., & Nichols-Belo, A. (2006). The globally competent engineer: Working effectively with people who define problems differently. *Journal of Engineering Education*, *95*(2), 107–122. https://doi.org/10.1002/j.2168-9830.2006.tb00883.x

Educational Testing Service. (2015). *ETS standards for quality and fairness*. https://www.ets.org/pdfs/about/standards-quality-fairness.pdf

Fantini, A. E. (2009). Assessing intercultural competence: Issues and tools. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 456–476). Sage Publications.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage Publications.

Friedman, T. L. (2005, April 3). It's a flat world, after all. *The New York Times*. https://www.nytimes.com/2005/04/03/magazine/its-a-flat-world-after-all.html

Friedman, T. L. (2017). *Thank you for being late: The optimist's guide to thriving in the age of accelerations*. Farrar, Straus, and Giroux.

Fuertes, J. N., Miville, M. L., Mohr, J. J., Sedlacek, W. E., & Gretchen, D. (2000). Factor structure and short form of the Miville-Guzman Universality-Diversity Scale. *Measurement and Evaluation in Counseling and Development*, *33*(3), 157–169. http://doi.org/10.1080/07481756.2000.12069007

Gallagher, M. W., & Brown, T. A. (2013). Introduction to confirmatory factor analysis and structural equation modeling. In T. Teo (Ed.), *Handbook of quantitative methods for educational research* (pp. 289–314). Sense Publishers.

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, *4*(1), 26–42. https://doi.org/10.1037/1040-3590.4.1.26

Groves, K. S., & Feyerham, A. E. (2011). Leader cultural intelligence in context: Testing the moderating effects of team cultural diversity on leader and team performance. *Group & Organization Management*, *36*(5), 535–566. https://doi.org/10.1177/1059601111415664

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The intercultural development inventory. *International Journal of Intercultural Relations*, *27*(4), 421–443. https://doi.org/10.1016/S0147-1767(03)00032-4

Huang, T.-J., Chi, S.-C., & Lawler, J. J. (2005). The relationship between expatriates' personality traits and their adjustment to international assignments. *International Journal of Human Resource Management*, *16*(9), 1656–1670. https://doi.org/10.1080/09585190500239325

Jesiek, B. K., Haller, Y., & Thompson, J. (2014). Developing globally competent engineering researchers: Outcomes-based instructional and assessment strategies from the IREE 2010 China research abroad program. *Advances in Engineering Education*, *4*(1), 1–31. https://advances.asee.org/?publication=developing-globally-competent-engineering-researchers-outcomes-based-instructional-and-assessment-strategies-from-the-iree-2010-china-research-abroad-program

Jesiek, B. K., Shen, Y., & Haller, Y. (2012). Cross-cultural competence: A comparative assessment of engineering students. *International Journal of Engineering Education*, *28*(1), 144–155. https://www.ijee.ie/contents/c280112.html

Jesiek, B. K., Trellinger, N. M., & Mazzurco, A. (2016). *Becoming boundary spanning engineers: Research methods and preliminary findings*. Paper presented at the ASEE Annual Conference and Exposition, New Orleans, LA. http://doi.org/10.18260/p.26370

Jesiek, B. K., Woo, S. E., Parrigon, S., & Porter, C. (2020). Development of a situational judgement test (SJT) for global engineering competency (GEC). *Journal of Engineering Education*, *109*(3), 1–21. https://doi.org/10.1002/jee.20325

Jesiek, B. K., Woo, S. E., Zhu, Q., Ramane, K. D., & Choudhary, N. (2015). *Defining and assessing global engineering competency: Methodological reflections*. Paper presented at the ASEE Annual Conference and Exposition, Seattle, WA. http://doi.org/10.18260/p.23781

Jesiek, B. K., Zhu, Q., Woo, S. E., Thompson, J., & Mazzurco, A. (2014). Global engineering competency in context: Situations and behaviors. *Online Journal for Global Engineering Education*, *8*(1), 1–14. http://digitalcommons.uri.edu/ojgee/vol8/iss1/1

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). Guilford Press.

Johri, A., & Jesiek, B. K. (2014). Global and international issues in engineering education. In A. Johri & B. M. Olds (Eds.), *Cambridge handbook of engineering education research* (pp. 655–672). Cambridge University Press.

King, P. M., & Baxter Magolda, M. B. (2005). A developmental model of intercultural maturity. *Journal of College Student Development*, *46*(6), 571–592. https://doi.org/10.1353/csd.2005.0060

Knight, D. B., Davis, K. A., Kinoshita, T., Soledad, M. M., & Grohs, J. R. (2017). *Assessing students' global and contextual competencies: Three categories of methods used to assess a program with coursework and international modules*. Paper presented at the ASEE Annual Conference and Exposition, Columbus, OH. http://doi.org/10.18260/1-2–27622

Levonisova, S. V., Huang, S., Streiner, S. C., Cunningham, S., Ragusa, G. R., Besterfield-Sacre, M. E., Shuman, L. J., & Matherly, C. (2014). *Moving towards a research informed conceptual model of engineering global preparedness*. Paper presented at the ASEE Annual Conference and Exposition, Indianapolis, IN. http://doi.org/10.18260/1-2–22850

Levonisova, S. V., Savage, R. E., Streiner, S. C., McCave, E. J., Ragusa, G. R., Matherly, C., Benson, L., Besterfield-Sacre, M., & Shuman, L. J. (2015). *Identifying factors that enhance undergraduate engineering students' global preparedness*. Paper presented at the ASEE Annual Conference and Exposition, Seattle, WA. http://doi.org/10.18260/p.24211

Lohmann, J. R., Rollins, H. A., Jr., & Hoey, J. J., IV. (2006). Defining, developing and assessing global competence in engineers. *European Journal of Engineering Education*, *31*(1), 119–131. https://doi.org/10.1080/03043790500429906

Lou, K. H., Vande Berg, M., & Paige, R. M. (2012). Intervening in student learning abroad: Closing insights. In M. V. Berg, R. M. Paige, & K. H. Lou (Eds.), *Student learning abroad: What our students are learning, what they're not, and what we can do about it* (pp. 411–419). Stylus.

Mazzurco, A., Jesiek, B. K., & Godwin, A. (2020). Development of global engineering competency scale: Exploratory and confirmatory factor analysis. *Journal of Civil Engineering Education*, *146*(2), 1–10. https://doi.org/10.1061/(ASCE)EI.2643-9115.0000006

Miville, M. L., Gelso, C. J., Pannu, R., Liu, W., Touradji, P., Holloway, P., & Fuertes, J. N. (1999). Appreciating similarities and valuing differences: The Miville-Guzman Universality-Diversity Scale. *Journal of Counseling Psychology*, *46*(3), 291–307. https://doi.org/10.1037/0022-0167.46.3.291

National Science Board. (2019). *Higher education in science and engineering* (NSB-2019-7; Science and Engineering Indicators 2020). https://ncses.nsf.gov/pubs/nsb20197/data

Ogden, A. (2015). *Toward a research agenda for U.S. education abroad* (AIEA Research Agendas for the Internationalization of Higher Education). Association of International Education Administrators. http://www.aieaworld.org/assets/docs/research_agenda/ogden_2015.pdf

Ogilvie, A. M., Knight, D. B., Long, T. C., Adams, S. G., Amelink, C. T., & Sanderlin, N. (2015). *Rising sophomore abroad program: A model for professional formation of globally competent engineers*. Paper presented at the ASEE Annual Conference and Exposition, Seattle, WA. http://doi.org/10.18260/p.24685

Parkinson, A. (2009). The rationale for developing global competence. *Online Journal for Global Engineering Education*, *4*(2), 1–15. http://digitalcommons.uri.edu/ojgee/vol4/iss2/2

Pike, G. R. (2011). Using college students' self-reported learning outcomes in scholarly research. *New Directions for Institutional Research*, *2011*(150), 41–58. https://doi.org/10.1002/ir.388

Porter, S. R. (2013). Self-reported learning gains: A theory and test of college student survey response. *Research in Higher Education*, *54*, 201–226. https://doi.org/10.1007/s11162-012-9277-0

Ragusa, G. R. (2014). Engineering global preparedness: Parallel pedagogies, experientially focused instructional practices. *International Journal of Engineering Education*, *30*(2), 400–411. https://www.ijee.ie/contents/c300214.html

Spitzberg, B. H., & Changnon, G. (2009). Conceptualizing intercultural competence. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 2–52). Sage Publications.

Stevens, R., Johri, A., & O'Connor, K. (2014). Professional engineering work. In A. Johri & B. M. Olds (Eds.), *Cambridge Handbook of Engineering Education research* (pp. 119–137). Cambridge University Press.

Trevelyan, J. (2010). Reconstructing engineering from practice. *Engineering Studies*, *2*(3), 175–195. https://doi.org/10.1080/19378629.2010.520135

Van Dyne, L., Ang, S., & Tan, M. L. (2017). *Cultural intelligence* (Oxford bibliographies). Oxford University Press.

Vande Berg, M., Connor-Linton, J., & Paige, R. M. (2009). The Georgetown Consortium Project: Interventions for student learning abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *18*(1), 1–75. http://doi.org/10.36366/frontiers.v18i1.251

Vande Berg, M., & Paige, R. M. (2012). Why students are and are not learning abroad. In V. Berg, R. M. P. Michael, & K. H. Lou (Eds.), *Student learning abroad: What our students are learning, what they're not, and what we can do about it* (pp. 29–58). Stylus.

Vande Berg, M., Paige, R. M., & Lou, K. H. (2012). Student learning abroad: Paradigms and assumptions. In M. V. Berg, R. M. Paige, & K. H. Lou (Eds.), *Student learning abroad: What our students are learning, what they're not, and what we can do about it* (pp. 3–28). Stylus.

Walther, J., & Radcliffe, D. (2007). The competence dilemma in engineering education: Moving beyond simple graduate attribute mapping. *Australasian Journal of Engineering Education*, *13*(1), 41–52. http://doi.org/10.1080/22054952.2007.11464000

AUTHOR BIOGRAPHIES

Kirsten A. Davis is an Assistant Professor of Engineering Education at Purdue University, 701 West Stadium Ave., West Lafayette, IN 47907, USA; kad@purdue.edu

Brent K. Jesiek is a Professor of Engineering Education and Electrical and Computer Engineering at Purdue University, 701 West Stadium Ave., West Lafayette, IN 47907, USA; bjesiek@purdue.edu

David B. Knight is a Professor in the Department of Engineering Education at Virginia Tech, 635 Prices Fork Rd, Blacksburg, VA 24060, USA; dbknight@vt.edu

How to cite this article: Davis, K. A., Jesiek, B. K., & Knight, D. B. (2023). Exploring scenario-based assessment of students' global engineering competency: Building evidence of validity of a China-based situational judgment test. *Journal of Engineering Education*, 1–24. https://doi.org/10.1002/jee.20552

APPENDIX

Additional demographic and experience information for student participants.

TABLE A1 Additional student characteristic variables.

Variables	Prcentage of participants
Race/ethnicity	
American Indian/Alaska Native	0.4%
Asian	9.5%
Black or African American	1.7%
Hispanic or Latinx	4.9%
Native Hawaiian or Pacific Islander	0.4%
Two or more races	5.4%
White	77.0%
Not reported	0.8%
Year in school	
First year	15.2%
Second year	33.3%
Third year	24.3%
Fourth year	22.2%
5th+ year undergraduate	2.5%
Master's students	0.8%
PhD students	0.8%
Not reported	0.8%
Engineering major	
Aerospace/aeronautical engineering	5.8%
Biological systems engineering	2.1%
Biomedical engineering	2.1%
Chemical engineering	5.8%
Civil engineering	10.7%
Computer engineering	3.7%
Computer science	4.1%
Electrical engineering	5.4%
Engineering management	18.9%
Environmental engineering	4.1%
Industrial and systems engineering	9.9%
Mechanical engineering	22.2%
Other engineering	5.4%