122nd ASEE Annual Conference & Exposition
June 14 - 17, 2015
Seattle, WA
Making Value for Society
Paper ID #12226

Defining and Assessing Global Engineering Competency: Methodological Reflections

Prof. Brent K Jesiek, Purdue University, West Lafayette

Dr. Brent K. Jesiek is Associate Professor in the Schools of Engineering Education and Electrical and Computer Engineering at Purdue University. He is also an Associate Director of Purdue's Global Engineering Program, leads the Global Engineering Education Collaboratory (GEEC) research group, and is the recipient of an NSF CAREER award to study boundary-spanning roles and competencies among early career engineers. He holds a B.S. in Electrical Engineering from Michigan Tech and M.S. and Ph.D. degrees in Science and Technology Studies (STS) from Virginia Tech. Dr. Jesiek draws on expertise from engineering, computing, and the social sciences to advance understanding of geographic, disciplinary, and historical variations in engineering education and practice.

Prof. Sang Eun Woo, Purdue University

Sang Eun Woo is an assistant professor in the Department of Psychological Sciences at Purdue University. She received her PhD degree in industrial and organizational psychology from University of Illinois Urbana-Champaign. Her research interests include construct validation, psychological measurement (focusing on behavioral assessment), personality and individual differences, turnover and withdrawal, and alternative methodological approaches to organizational and psychological science.

Qin Zhu, Purdue University

Qin Zhu is a PhD student in the School of Engineering Education at Purdue University. His main research interests include global, comparative, and international engineering education, engineering education policy, and engineering ethics. He received his BS degree in materials science and engineering and first PhD degree in the philosophy of science and technology (engineering ethics) both from Dalian University of Technology, China.

Ms. Kavitha D Ramane, Purdue University
Neha Choudhary, Purdue University Programs

Neha choudhary is currently pursuing doctoral studies at Purdue university in Engineering Education department. Her research interest includes graduate studies, global engineering, design for community services.

Defining and Assessing Global Engineering Competency: Methodological Reflections

Introduction

Researchers face manifold challenges as engineering education continues to grow and evolve as a distinct field of scholarly activity. For instance, discussions about criteria for evaluating engineering education research have intensified, including through published commentaries focused on concepts like "research quality," "rigor," and "systematic research," as well as accompanying shifts in the various criteria used to evaluate funding proposals and peer reviewed papers. The field's topical foci are also something of a moving target given a long and episodic history of efforts to reinvent the form and content of engineering curricula. As the methods and desired outcomes of engineering instruction change, so does the engineering education research agenda. Further worth noting are rising pressures to relate research to practice, as reflected in mandates to identify the "broader impacts" associated with scholarly work in the field.

This paper speaks to these challenges through the lens of our team's recent experiences working on a research project supported by the National Science Foundation (NSF), one primary objective of which is to develop and validate a global competency assessment tool. This paper is organized in three major parts, each focused on a different phase of the project and a variety of associated questions and challenges. After giving a general overview of the larger research project, the first major section of the paper is focused on clarifying our main domain of interest, including our efforts to develop a working definition of "global engineering competency" (GEC) that is grounded in both the extant literature and our own qualitative data set. This part of the paper is particularly concerned with how we meaningfully delineate, define, and study specific domains or facets of engineering education and practice, including by working toward more nuanced interpretations of key attributes and capabilities important for current and future engineers. The second section of the paper describes our efforts to develop a large pool of assessment questions using a situational judgment test (SJT) format. Here we interrogate our rationale for picking this type of assessment tool, procedures for developing the scenario-based test questions, and efforts to bolster validity through use of an underlying theoretical framework and systematic item development procedures. Third and finally, we discuss our ongoing efforts to validate and deploy our large pool of SJT questions, with an emphasis on data collection bottlenecks, contrasting approaches to instrument validation, and the need for an online survey and assessment system.

It is worth emphasizing that this paper reports on research-in-progress, and is focused on a variety of methodological decisions – some which have already been made, and others still under consideration. One major goal of this paper is to make visible and critically reflect on the kinds of study design decisions which are often absent or simplified in more conventional research papers. We therefore see this work as contributing to broader discussions of methodology and method in engineering education,[1-3] while also following in the footsteps of colleagues who have published detailed discussions about the methodological decisions and challenges encountered in their own research projects.[4-5] We expect our commentary will be of interest to engineering education researchers who wish to explore contemporary methodological concerns that cut across multiple research traditions. This paper may also be of interest to scholars involved with efforts to define, develop, and/or assess other complex and multi-faceted learning outcomes such as innovation, teamwork, leadership, inter/trans/multi-disciplinarity, engineering design, etc.

Project Overview

This paper reflects on a larger, multi-year research project focused on enhancing the ability of engineers and other technical professionals to work across national and cultural boundaries. More specifically, the project is organized around three major objectives:

- O1: Generate an empirically based definition and developmental theory of global engineering competency.
- O2: Create and disseminate a valid and reliable SJT assessing 3-4 critical dimensions of global engineering competency in both paper-and-pencil and online formats.
- O3: Promote expansion of a scholarly community focused on the continued development and use of situational assessment techniques in engineering education.

The sections below are most directly linked to objectives one and two, organized sequentially to first discuss our efforts to clarify the domain of interest ("global engineering competency") using a qualitative approach. We then discuss how these findings informed the development and then validation of a situational judgment test, which respectively involved use of qualitative and then quantitative approaches. We also intend that our discussion will more generally support our third project objective, namely by contributing to ongoing conversations about assessment techniques in engineering education, particularly in terms of reliability and validity considerations.

The multiple methods used in this study provide considerable inspiration for methodological reflection. The first project stages were largely exploratory in nature, focused on a mix of what Elliott and Timulak refer to as "definitional" and "descriptive" kinds of research questions.[6] The processes we then used to develop assessment items drew on our initial empirical findings, while also benefitting from a considerable measure of creativity and use of an emerging theoretical framework. The latter stages of the research project are more quantitative in nature, with methods focused on developing scoring keys for the assessment questions and collecting validity evidence. Here, the associated questions and objectives are more about diagnosis and prediction, e.g., can our SJT reliably diagnose and predict levels of global engineering competency in various subject populations? As this overview suggests, each research phase described here has been designed to support our larger project objectives. Yet the varied approaches and procedures used along the way raise questions about how the parts of our study speak – or fail to speak – to one another. The following sections will more deeply explore these synergies and tensions.

Phase 1: Clarifying the Domain of Interest

The first objective of our research project centers on the important yet difficult task of clarifying the domain of interest, i.e., global competency in engineering. Prior commentary by Walther and Radcliffe discusses two major ways of developing such competency profiles.[7] The first approach, which is more common in educational settings, is to organize panels of experts that generate relatively wide-ranging sets of target graduate attributes. Strategies like the Delphi method are sometimes used to more systematically identify, clarify, cluster, and prioritize such lists.[8-9] While the 11 outcomes comprising Criterion 3 of ABET's Engineering Criteria 2000 (EC2000) accreditation framework is a particularly influential attribute list in engineering education,[10] the 24 outcomes presented in the American Society of Civil Engineer's *Civil Engineering Body of Knowledge for the 21st Century* also deserve mention as a more transparent, expansive, and nuanced product of a large expert panel.[11]

It is also instructive to note that many such reports and documents often say little about how the target attributes identified through such studies might specifically be implemented in extant courses and curricula, much less assessed. Indeed, one finds that many attempts to assess the ABET Criterion 3 outcomes have involved development and use of indirect measures, e.g., survey instruments and questions that ask students to self-assess their own capabilities for each target attribute.[12] Many assumptions undergird this entire process, including that: a) groups of practicing professionals and educators can accurately identify, through reflection and discussion, the specific knowledge, skills, attitudes, abilities, etc. that a practicing professional should possess, b) the process of bringing these attributes into courses and curricula is relatively straightforward, and c) students can reliably and accurately gauge and report on their own capabilities, such as by filling out Likert-scale surveys that list various target outcomes.

As Walter and Radcliffe discuss, a second and more robust approach to generating competency profiles can be found in industry settings. This typically involves creating very detailed characterizations for specific job roles or positions in an organization, often based on evidence collected through direct observations of work practices and critical incident interview techniques. Walther and Radcliffe further note that each of these approaches has its own advantages, reflecting different priorities for different contexts of use (e.g., broader educational goals in university degree programs versus higher levels of specification desired by employers).

Yet these authors also point to prior research suggesting only moderate agreement between the attributes identified by expert panels as compared to more direct studies of practice. On this point, one commentator with considerable experience in this domain has argued that "about 50 percent of the competencies identified by expert panels are validated by a full competency study."[13] As this author suggests, some of this disagreement is likely linked to experts lacking appropriate psychological vocabulary, and also proposing "folklore" competencies that may be inadequately defined, difficult to assess, and/or poorly linked to actual job performance.

One important consequence of this gap is what Walther and Radcliffe call a "competence dilemma", where degree programs establish target attributes for their graduates based in part on direct or indirect input from expert panels, but employers continue to bemoan a lack of preparation for workplace realities among the early career engineers they recruit. As the preceding overview suggests, relying on experts to identify desirable attributes by reflecting on the nature of practice may generate impoverished characterizations of engineering work. Indeed, a growing body of empirical studies presenting much "thicker" descriptions of engineering practice provide further support for this claim.[14] And yet, lists of attributes and outcomes generated by experts continue to inform development of curricula, resulting in poorly integrated coursework and scattershot rather than holistic treatment of learning outcomes.

Returning to the main focus of this paper raises the question of what strategies have been used to define global competency and related constructs in engineering. Our forays into the literature reveal that U.S. educators and policymakers have been discussing desirable capabilities and modes of training for engineers working internationally since at least the 1940s.[15] Yet as we have documented, more systematic efforts to define and assess global competency only started to emerge in the late 1990s and early 2000s,[16] and this movement has more recently been called an "emerging field of inquiry".[17] Intensified globalization trends – and industry demands in particular – have helped spur a variety of efforts to identify, promote, and assess various attributes deemed important for global professional practice. For instance, one review of the literature identified fourteen papers and reports published 1997-2012 that have proposed various definitions and lists of attributes for global engineers.[16]

As we have observed, some of these prior studies have adopted "expert panel" approaches to defining global competency. Many others have been based on weaker forms of evidence such as prior literature (which itself often lacks empirical grounding), the personal and/or professional experiences of the authors themselves, and/or pre-existing learning outcomes for global courses and programs. And even when experts are consulted to generate lists of attributes or definitions of competence, rarely are these informants described in detail, making it difficult to evaluate their qualifications. The typical study of this type offers a long and rather broad set of target attributes or outcomes, often without deep definitional or theoretical grounding. Direct studies of global engineering practice are comparatively sparse, albeit with a few exceptions.[18]

Nonetheless, our numerous forays into the extant literature allowed us to initially identify three major foci in existing discussions of global competency: (1) various technical and professional attributes that are not explicitly global in nature but are important for global engineers (e.g., communication and teamwork skills, domain-specific technical expertise), (2) attributes that are important for most any global professional (e.g., foreign language proficiency, intercultural competence, historical and cultural knowledge) but not are specific to the engineering profession, and (3) global engineering competency (GEC), or "those capabilities and job requirements that are uniquely or especially relevant for effective engineering practice in global context."[16]

Our decision to focus the present research project on the third of these areas was based on a number of factors, including a practical need to keep the scope of our efforts manageable, a relative lack of explicit attention to the third area in other studies, and comparatively large bodies of literature already focused on the first and second areas. When we initially proposed our project, we also postulated three more specific domains or dimensions of GEC based on local and regional differences in: "engineering cultures" (i.e., actual engineering work styles and practices), engineering ethics and professionalism, and standards and regulations (e.g., units of measure, legal requirements, documentation practices, technical codes and standards, etc.). We identified these areas based on our own prior research and knowledge of the literature, but also recognized that further empirical work was needed to more systematically identify and verify key facets or dimensions of GEC.

While we adopted a multifaceted approach to this challenge, philosophically our goal was to get as close to possible to actual situations and behaviors encountered by engineers working cross-nationally/culturally. We first conducted a wide-ranging search of the literature, with a focus on collecting case studies and other discussions of global work situations and practices from publications in engineering and other technical fields, business and management, and cross-cultural studies. We especially sought out sources that discussed one or more of the following: 1) globally competent professional behaviors, 2) situations and interactions involving multiple national, regional, and/or ethnic cultures, and 3) technical tasks and/or problem solving. We ultimately collected more than 50 relevant case studies from various sources including books, case studies collections, journal articles, and conference papers.

The research team collected further data via one-on-one interviews and small group (2-5 participant) discussions with about 25 total subjects. Most of these individuals held at least one engineering or other technical degree, and all had previous or current global work experience, typically in multinational firms. Rather than asking them what attributes or competencies were most important for global work, we placed much greater emphasis on specific work situations they had encountered. The data collection procedure first involved sensitizing interviewees to the domain of interest through discussion of the sample global scenario prompt presented in Figure 1. We then employed a combination of narrative and critical incident approaches to elicit stories of cross-cultural working experiences from the participants. In some instances, we also tried to energize and broaden conversations by suggesting additional topics and themes that had not yet been discussed by the participants (e.g., cross-cultural teamwork, ethics and professionalism, styles of communication, standards and regulations, local differences in technical practice, etc.).

As an employee in a large multinational corporation, you are temporarily assigned to your company's branch operations in Shanghai, China. You are a member of a team consisting of three Chinese engineers, all about the same rank as you. Your team reports to an engineering manager, who is also Chinese. You are in a team meeting where your manager proposes a solution to a difficult quality control problem. However, you are concerned that the proposed solution will fail. Consider these possible actions:

a) Have the entire team approach the manager together.
b) Bring up your concerns in the meeting.
c) Set aside your concerns and follow the manager's lead.
d) Discuss the issue with the manager later, in a private meeting.
e) Consult your Chinese team members about appropriate actions to take.
f) Discuss your concerns with a higher-ranking manager.

Which of these actions (a-f) would you MOST likely take? Which of these actions (a-f) would you LEAST likely take?

Figure 1. Sample/Pilot SJT Question for Global Engineering Competency

The literature and empirical data collected during this phase of research has served two main purposes. First, we discuss in more detail below how it has helped inspire the drafting of SJT scenarios. Second, it allowed us to reflect on and refine our list of sub-dimensions or facets of GEC. As reported previously, we specifically proposed the following GEC dimensions: *technical coordination* (in cross-cultural context), *understanding and negotiating engineering cultures*, and *navigating ethics, standards, and regulations*.[16] This revised set of dimensions involved two major changes from what we had originally proposed. First, we combined ethics with standards and regulations because we found that professional decisions about standards and regulations often have an ethical component. For instance, deciding whether to follow regulatory requirements set by one's home country or a different set of requirements prevailing in a host country location may test an engineer's moral and ethical commitments, such as in relation to safety standards, environmental protection, issues of social justice, etc. A second change to the GEC dimensions involved identifying technical coordination as a very broad category of situations that involve "getting things done" amidst cultural differences, but with technical issues or expertise as a major backdrop consideration. To further improve grounding of these dimensions in the extant literature, we related each dimension to previously published lists of desirable attributes for global engineers. The dimensions were in turn used to organize our draft SJT development efforts. Nonetheless, we also acknowledge that opportunities remain to further analyze and report on our literature collection and empirical data, including to verify the GEC dimensions already identified, as well as to seek out other potential dimensions or facets of GEC.

Still another challenge in our early project efforts centered on whether and how to focus our work on specific countries or regions. We initially proposed developing sets of SJT questions for a handful of general geographic regions, e.g., North America, East Asia, Western Europe, etc. Yet as we engaged with the literature and study participants, we were reminded that major historical and cultural differences *within* many regions made this approach infeasible. We therefore more narrowly targeted our efforts on select countries based on a variety of criteria, such as by considering what countries were: prevalent in our literature collection and empirical data set; among top U.S. trade partners; home to many multinational firms and subsidiaries; large R&D spenders; and among the world's largest economies. Given these criteria, we picked six target countries for our SJT, namely China, France, Germany, India, Japan, and Mexico.

Phase 2: Developing the Assessment Tool

While scenario-based and situational assessment techniques have been used in a handful of studies in engineering education, they have a long history of use and are regularly deployed in many other domains.[19] We find particular promise in the SJT format because it allows for measurement of behavioral tendencies, which are in turn linked to a test-takers foundational knowledge, skills, attitudes, and other personal characteristics. It is also difficult to guess or fake the "right" answers to SJT questions, and the format is highly scalable since data can be collected and evaluated automatically once scoring keys have been developed.

Our general approach to SJT development is consistent with the current literature.[20-23] More specifically, SJTs are typically developed in four stages.[22,24-26] The first involves item stem generation, where critical incidents related to real work situations are generated by subject matter experts (SMEs), then revised and grouped into categories to create item stems (i.e., realistic work scenarios or situations). Second, sets of response options are generated for each item stem, typically based on sample responses and/or direct input from SMEs and novices to cover a wide range of possible response options. Third, a scoring key is developed to allow systematic appraisal of test results, namely by establishing the effectiveness of each possible response option. Fourth and finally, additional large-scale validity tests may be carried out to establish the plausibility and predictive strength of the final items. In our own efforts we are implementing conventional procedures and best practices widely shown to be effective, e.g., by using survey and SME methods to generate scenarios and response options.[23,27-28]

Here we summarize initial development of the assessment tool in two phases. The first involved many iterations of creative writing and revision of draft SJT item stems. This effort was initially launched with a series of workshop-style meetings with members and affiliates of the lead author's research group. Examples, stories, and cases drawn from both our literature collection and empirical data set were compiled and shared with the writers, who were then asked to draft prompts inspired by the data and/or their own experiences. Writers were also given a sample SJT item, as in Figure 1, and were told that the scenarios should depict realistic situations that might be encountered by global engineers and handled in a variety of more-or-less effective ways. Draft item stems were shared and revised in pairs and/or larger group meetings until the research team was happy with the quality of each. The revisions often focused on clarity and consistency of language, and making sure the proposed scenario was realistic and appropriate in scope. Particularly in our early stages of item stem development, we brainstormed ways of handling some of the item stem situations to help establish whether it would later be feasible to write an associated set of item response options.

In parallel with initial scenario development we also revised and finalized both the dimensions and countries of interest for the scenarios, as noted above. We used this new information to categorize all of the scenarios already developed, in some cases discarding or modifying scenarios that did not fit into our final list of target dimensions or countries. This also pushed us to more systematically start filling in gaps in our item matrix (i.e., particular combinations of country and dimension for which we had few items), with the goal of having at least three draft scenarios for every country-dimensions combination, or at least 54 scenarios total.

Subsequent scenario development occurred country-by-country, with pairs of researchers generating and revising draft scenarios based on any and all sources of inspiration they could find, including our pre-existing literature and empirical data, and further insights gleaned from other books, articles, and web sites (e.g., *Harvard Business Review*). In some cases we created multiple versions of a given scenario simply by changing the country context. Writing the last few scenarios needed to complete the matrix proved to be a very difficult and creative process, as this tended to involve specific combinations of country and dimension for which we lacked evidence or examples to inspire our writing efforts. This often led to extensive forays onto the web, searching many different keyword combinations in hopes of turning up useful ideas. For example, searching keywords such as "engineering", "ethics", and "French" led us to a case study from the UK about how to handle questionable workplace behaviors when working abroad in France.[29] We used the case to draft an ethics scenario for France. In summary, we generated 74 total scenarios, 66 of them unique. Item distribution by country is as follows: China (18 items), Japan (10), India (11), Germany (14), France (11), and Mexico (10). Item distribution by domain is as follows: Technical Coordination (26), Engineering Cultures (21), and Ethics, Standards, and Regulations (27).

The second major phase of our SJT item development process involved generating response options for each scenario. To begin, we first asked small numbers of novices and experts to individually respond to each of our draft scenarios in an open-ended online survey format, both to see how the situations were being interpreted and as a source of ideas for developing fixed response options. To keep this process manageable, we only asked respondents to complete one set of scenarios at a time, organized by country. Novice respondents were usually undergraduate or graduate engineering students with little or no special knowledge of the country or culture in question. Experts were typically practitioners (i.e., global engineers) with some prior exposure to the country or culture in question. For a large majority of scenarios we collected at least 3 novice and 5 expert responses. While our main goal was to get a good variety of responses, we could have strengthened this step by collecting more data about our respondents (demographics, etc.).

The research team next worked to develop response options for all draft scenarios, country-by-country. This typically involved sending a batch of 3-5 scenarios, along with the associated open-ended novice and expert responses, to a pair of researchers, including undergraduate students, graduate students, and faculty members. Those involved with this project have typically worked individually and then in pairs to analyze scenarios and write response options. Early in this process we also realized that the scenarios and response options could usually be viewed in terms of three major considerations: cultural knowledge, domain/technical knowledge, and cultural sensitivity. This insight was aligned with the developmental model we had originally proposed for the project, as well as theoretical models proposed by other researchers. Based on this insight, we started encouraging those who were writing response options to reflect deeply on what kinds of knowledge and sensitivity considerations might be relevant for a given scenario, and then brainstorm possible ways of responding to the scenario based on both their analysis and the sample open-ended responses.

Table 1. Situational Judgment Test (SJT) Item Matrix

	Dimensions of Global Engineering Competency (GEC)			
National Context	***Technical Coordination***	***Engineering Cultures***	***Ethics, Standards, and Regulations***	***No. Fully Developed***
China	9 items	3 items	6 items	16 (of 18)
Japan	3 items	3 items	4 items	9 (of 10)
India	3 items	3 items	5 items	12 (of 12)
Germany	4 items	6 items	4 items	11 (of 14)
France	4 items	3 items	4 items	TBD
Mexico	3 items	3 items	4 items	TBD

The researchers were also told to classify or code each draft response for relative levels of knowledge and sensitivity. In many cases this provided inspiration for creating novel response options to fill in gaps and promote wider range of both more and less effective response options. For example, we found that many scenarios permitted a low knowledge/high sensitivity response option that involved seeking additional information or help.

Throughout this process, sets of draft scenarios and response options were also brought into larger group meetings for continued discussion and edits, typically followed by one final round of review and edits by the lead researcher and at least one graduate student. When possible, individuals with native cultural knowledge and/or prior experience working with individuals from a given country/culture were asked to help review and revise scenarios and response options based on their own first-hand insights.

For the four countries completed so far, we have retained 47 of 53 scenarios. Each complete question written to date has 5-9 response options, with most having 6 or 7. Typical reasons for dropping scenarios include a lack of relevant knowledge and/or sensitivity considerations in the situation itself, pilot data suggesting a given scenario is ambiguous or unrealistic, and/or not being able to generate a large enough pool of responses that are both plausible and distinct.

As this overview suggests we strove to develop our SJT items by going beyond more traditional approaches, including by incorporating some methodological innovations previously introduced in the SJT literature to further improving the quality of our efforts.[20,30] For example, when constructing hypothetical GEC scenarios and generating corresponding behavioral response options, we keep a set of specific underlying psychological constructs in mind that have a firm theoretical foundation. We also adopt more data-driven (empirically based) approaches such as utilizing in-depth, qualitative interviews and case studies available from the existing literature in deriving such higher-level constructs. This combination of bottom-up (inductive) and top-down (deductive) approaches to developing scenarios and response options is achieved through multiple iterations of SME focus group discussions across diverse areas of domain expertise.

Phase 3: Validating the Assessment Tool

The ultimate test of our SJT development efforts centers on a critical third phase of the research project, where we turn to validating the draft test questions. This process has two main goals, namely to establish whether experts think the scenarios and response options are realistic and plausible, and to establish some standard or benchmark that can be used when scoring responses to the SJT questions. One of our original motivations for developing a very large pool of SJT questions – in our case at least 2-3 SJT questions for every combination of country and GEC dimension – is to provide a buffer in case some questions prove impossible to validate. Yet there is no guarantee that this process will produce a complete set of validated test questions.

While we are just launching our validation effort, it is worth commenting on some critical methodological issues related to the two main approaches we are now pursuing. The first approach is among the most widely used for scoring SJT items. It involves utilizing a small group of SMEs (i.e., job incumbents with extensive global experience) who identify best and worst options, or rate each response option on a continuum using a Likert-type scale (e.g., from 1=least desirable behavior/action to 5=most desirable behavior/action). A test-taker's answers will then be compared to the SME ratings; the more similarities between SME ratings and the test-taker's answers, the higher scores s/he would receive. This presumes that responses collected from SMEs responding to a set of SJT questions reflect some level of consensus on what are the most and least effective ways of handling each situation. Some low consensus response options may be used as distractors or thrown out, but the whole question might be discarded if expert consensus is limited or nonexistent. The question of what counts as consensus is also debatable.

Additionally, we plan to explore use of various other scoring strategies that have been proposed in the SJT literature, comparing them for their relative strengths and weaknesses in terms of diagnostic accuracy, as well as strongest predictive validity.[20-21,30-31] Doing so will likely require that we collect considerably more data from respondents, such as measures of personality, cultural knowledge and sensitivity, and self-assessment of global engineering competency, as well as detailed background and demographic information. One challenge of this approach, however, is that it requires relatively long surveys (up to 30 minutes in length) completed by relatively large numbers of SMEs (in the hundreds) to improve the likelihood that the results can be used to generate robust numerical models that reliably predict test-taker performance.

Conclusion

This paper offers an in-depth look at an in-progress research project that aims to develop and validate an SJT for global engineering competency, or GEC. We have discussed a number of important processes and decisions that might be omitted or glossed over in a more traditional journal paper format. In terms of methodological considerations, our efforts involve multiple stages, each with partially distinct assumptions about the nature of knowing. Our effort to clarify the domain of global engineering competency follows an interpretive approach, with primary emphasis on "the participants' views of the situation being studied."[32] Compiling a body of case study literature and interacting with more than two dozen SMEs provided us with a pool of data that could be analyzed inductively, including by identifying patterns in the stories and case studies that helped us identify and clarify three sub-dimensions of GEC. This phase also allowed us to generate theory that builds conceptual links between various foundational attributes (e.g., domain knowledge, cultural knowledge, cultural sensitivity) and real-world behaviors and actions. Potential limitations at this stage include whether we: recruited a suitably representative pool of SMEs; accurately identified important sub-dimensions of GEC; overlooked other notable sub-dimensions of GEC; and built an inclusive and well-grounded theoretical model.

Developing the SJT item stems was a more creative process that assumed we could translate the case study and SME data into relatively brief vignettes, each of which is realistic and plausible, primarily focuses on one dimension of GEC, is relatively unambiguous, and can be linked to some manageable number of plausible response options. Developing response options involved similar considerations, but also involved analyzing how response options could be mapped onto our theoretical framework (e.g., each response representing relatively higher or lower levels of knowledge and sensitivity). Perhaps most significantly, developing response options was rather reductionist in nature, where more-or-less nuanced responses from novices and experts – usually ranging in length from a sentence to a paragraph – were carefully boiled down into much more succinct and accessible response options, never more than a sentence in length. There is a danger here that complex, real-world situations are being oversimplified in ways that take us too far away from the actual practice of global engineering, in all of its messy, complex, nuanced glory. On the other hand, we would contend that even if the item stems are simplified, they are much closer to the reality of practice than long and rather abstract lists of potentially relevant attributes.

Ultimately, our success in developing realistic item stems and response options will be tested through the validation process. One critical first step in this regard involves establishing whether SMEs find the SJT scenarios plausible, and then whether there is consistency in how they evaluate the relative effectiveness of the various response options. If such consensus can be established, there remains the question of whether the SJT questions can be used to reliably differentiate novices from experts, predict current or future job performance, and perhaps even accurately measure multiple underlying facets of competency (e.g., knowledge and sensitivity) based on a given individual's unique response patterns.

If some minimally adequate level of validation can be established for a substantial subset of the SJT questions, the issue of disseminating the assessment tool becomes another concern. In order to proactively address this issue, we are currently building an online assessment platform that can be used to configure, administer, and automatically score the types of SJT assessment questions being developed through this project. Ultimately, we see tremendous value in making such a tool available, including to enable real-time, behavior-based assessments and diagnostics for students and professionals participating in global learning experiences and pursuing global careers. Yet we are acutely aware that the ultimate potential of these efforts is linked to a long chain of methodological assumptions and critical design decisions, none of which is ironclad.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1160455. The authors also thank other members and affiliates of the Global Engineering Education Collaboratory (GEEC) research group for their input on previous drafts of this paper.

References

[1] Borrego, M., Douglas, E. P., and Amelink, C. T. (2009). Quantitative, Qualitative, and Mixed Research Methods in Engineering Education. Journal of Engineering Education, 98(1): 53-66.

[2] Baillie, C. and Douglas, E. P. (2014). Confusions and Conventions: Qualitative Research in Engineering Education. Journal of Engineering Education, 103(1): 1–7.

[3] Johri, A. and Olds, B. (2014). Cambridge Handbook of Engineering Education Research. New York, NY: Cambridge University Press.

[4] Crede, E. and Borrego, M. (2013). From Ethnography to Items: A Mixed Methods Approach to Developing a Survey to Examine Graduate Engineering Student Retention. Journal of Mixed Methods Research 7(1): 62–80.

[5] Pawley, A. (2013). "Learning from Small Numbers" of Underrepresented Students' Stories: Discussing a Method to Learn about Institutional Structure Through Narrative. Proceedings of the 2013 ASEE Annual Conference and Exposition, Atlanta, GA, June 23-26, 2013.

[6] Elliott, R. and Timulak, L. (2005). Descriptive and Interpretive Approaches to Qualitative Research. In Miles, J. and Gilbert, P. (Eds.), A Handbook of Research Methods for Clinical and Health Psychology (pp. 147-159). Oxford, UK and New York, NY: Oxford University Press.

[7] Walther, J. and Radcliffe, D. F. (2007). The Competence Dilemma in Engineering Education: Moving Beyond Simple Graduate Attribute Mapping. Australasian Journal of Engineering Education, 13(1): 41-51.

[8] 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.

[9] Eskandari, H., Sala-Diakanda, S., Furterer, S., Rabelo, L., Crumpton-Young, L., and Williams, K. (2007). Enhancing the Undergraduate Industrial Engineering Curriculum: Defining Desired Characteristics and Emerging Topics. Education + Training, 49(1): 45-55.

[10] ABET, Inc. (2014). Criteria for Accrediting Engineering Programs. Baltimore, MD: ABET, Inc. Available at http://www.abet.org/eac-criteria-2015-2016/

[11] Body of Knowledge Committee of the Committee on Academic Prerequisites for Professional Practice. (2008). Civil Engineering Body of Knowledge for the 21st Century: Preparing the Civil Engineer for the Future (Second Edition). Reston, VA: American Society of Civil Engineers (ASCE).
[12] Strauss, L. C. and Terenzini, P. T. (2005). Assessing Student Performance on EC2000 Criterion 3.a-k. Proceedings of the 2005 ASEE Annual Conference and Exposition, Portland, OR, June 12-15, 2005.
[13] Spencer, L. M. (1997). Competency Assessment Methods. In Bassi, L. J. and Russ-Eft, D. (Eds.), Assessment Development and Measurement (pp. 1-36). Alexandria, VA: American Society for Training and Development.
[14] Stevens, R., Johri, A., and O'Conner, K. (2014). Professional Engineering Work. In Johri, A. and Olds, B. (Eds.), Cambridge Handbook of Engineering Education Research (pp. 119-140). New York, NY: Cambridge University Press.
[15] Jesiek, B. K. and Beddoes, K. (2010). From Diplomacy and Development to Competitiveness and Globalization: Historical Perspectives on the Internationalization of Engineering Education. In Downey, G. L. and Beddoes, K. (Eds.), What is Global Engineering Education For?: The Making of International Educators (pp. 45-76). San Rafael, CA: Morgan and Claypool.
[16] Jesiek, B. K., Zhu, Q., Woo, S. E., Thompson, J., and Mazzurco, A. (2014). "Global Engineering Competency in Context: Situations and Behaviors." Online Journal of Global Engineering Education, 8(1).
[17] Lohmann, J., Rollins, H., and Hoey, J. (2006). Defining, Developing, and Assessing Global Competence in Engineers. European Journal of Engineering Education, 31(1): 119-131.
[18] Johri, A. (2008). Boundary Spanning Knowledge Broker: An Emerging Role in Global Engineering Firms. Proceedings of the 38th Annual Frontiers in Education Conference, Saratoga Springs, NY, October 25-28, 2008.
[19] Jesiek, B. K. and Woo, S. E. (2011). Realistic Assessment for Realistic Instruction: Situational Assessment Strategies for Engineering Education and Practice. Proceedings of the SEFI Annual Conference 2011, Lisbon, Portugal, September 27-30, 2011.
[20] Bledow, R. and Frese, M. (2009). A Situational Judgment Test of Personal Initiative and Its Relationship to Performance. Personnel Psychology, 62: 229-258.
[21] Christian, M. S., Edwards, B. D., and Bradley, J. C. (2010). Situational Judgment Tests: Constructs Assessed and a Meta-Analysis of their Criterion-Related Validities. Personnel Psychology, 63(1): 83-117.
[22] Lievens F., Peeters, H., and Schollaert, E. (2008). Situational Judgment Tests: A Review of Recent Research. Personnel Review, 37(4): 426-441.
[23] Ployhart, R. E., and MacKenzie, W. I. (2010). Situational Judgment Tests. In Zedeck, S. (Ed.) The Handbook of Industrial and Organizational Psychology. Washington, DC: American Psychological Association.
[24] McDaniel, M. A. and Nguyen N.T. (2001). Situational Judgment Tests: A Review of Practice and Constructs Assessed. International Journal of Selection and Assessment, 9(1/2): 103-113.
[25] Peus, C., Braun, S., and Frey, D. (2013). Situation-Based Measurement of the Full Range of Leadership Model: Development and Validation of a Situational Judgment Test. Leadership Quarterly, 24(5): 777-795.
[26] McDaniel, M. A., Hartman, N. S., Whetzel, D. L., and Grubb, W. L. (2007). Situational Judgment Tests, Response Instructions, and Validity: A Meta-Analysis. Personnel Psychology, 60(1): 63-91.
[27] Lievens, F., and Sackett, P. R. (2007). Situational Judgment Tests in High-stakes Settings: Issues and Strategies with Generating Alternate Forms. Journal of Applied Psychology, 92(4): 1043-1055.
[28] Oswald F. L., Schmitt, N., Kim, H. B., Ramsay, L. J., and Gillespie, M. A. (2004). Developing a Biodata Measure and Situational Judgment Inventory as Predictors of College Student Performance. Journal of Applied Psychology, 89: 187–207.
[29] Fotheringham, H. (2007). French Engineers. Engineering Ethics Case Studies Database. Leeds, UK: University of Leeds. Available at http://www.engsc.ac.uk/downloads/scholarart/ethics/frenchengineers.pdf
[30] Motowidlo, S. J., and Beier, M. E. (2010). Differentiating Specific Job Knowledge from Implicit Trait Policies in Procedural Knowledge Measured by a Situational Judgment Test. Journal of Applied Psychology, 95(2): 321-333.
[31] Ascalon, M., Schleicher, D., and Born, M. (2008). Cross-Cultural Social Intelligence: An Assessment for Employees working in Cross-National Contexts. Cross Cultural Management, 15(2): 109-130.
[32] Creswell, J. W. (2003). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2nd Ed.). Thousand Oaks, CA: Sage.