A Dialectic Data Integration Approach for Mixed Methods Survey Validation

Mr. Nicholas D. Fila, Purdue University

Nicholas D. Fila is a Ph.D. candidate in the School of Engineering Education at Purdue University. He earned a B.S. in Electrical Engineering and a M.S. in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign. His current research interests include innovation, empathy, and engineering design.

Mr. Justin L Hess, Purdue University, West Lafayette

Justin Hess is a Ph.D. candidate at Purdue University's School of Engineering Education, Masters student in the School of Civil Engineering and a National Science Foundation Graduate Research Fellow. He received his Bachelor's of Science in Civil Engineering in 2011 with a minor in philosophy and anticipates receiving his MSCE in 2015, both from Purdue University. His research focuses on understanding engineers' core values, dispositions, and worldviews. His dissertation focuses on conceptualizations, the importance of, and methods to teach empathy to engineering students. He is currently the Education Director for Engineers for a Sustainable World and an assistant editor for Engineering Studies.

Dr. Senay Purzer, Purdue University, West Lafayette

enay Purzer is an Assistant Professor in the School of Engineering Education. She is the recipient of a 2012 NSF CAREER award, which examines how engineering students approach innovation. She serves on the editorial boards of Science Education and the Journal of Pre-College Engineering Education (JPEER). She received a B.S.E with distinction in Engineering in 2009 and a B.S. degree in Physics Education in 1999. Her M.A. and Ph.D. degrees are in Science Education from Arizona State University earned in 2002 and 2008, respectively.

A Dialectic Data Integration Approach for Mixed Methods Survey Validation

Abstract

In quantitative studies, surveys are often used as tools to gauge attitudes, knowledge, and behaviors. Often, when these surveys are used in new contexts or with new populations, they require validation procedures such as confirmatory factor analysis or comparison to similar measures. These methods, however, are bounded by the need for large sample sizes which are not always feasible. In this paper, we discuss the use of mixed-methods research for survey validation. We present an example study that incorporates both traditional quantitative data and qualitative data into the validation of a survey targeted at engineering students. First, we present the philosophical underpinnings of quantitative and qualitative validation and discuss connections that allow both traditions to be incorporated into the same survey validation. Second, we discuss how quantitative and qualitative data can be mixed to form a deeper understanding of the participants, their educational context, and how survey results might be interpreted in that context among those participants. This paper contributes to research in engineering education by providing a dialectic data integration approach to support survey validation through the use of mixed methods.

The Need for Mixed Methods Survey Validation

Surveys, tests, and other types of assessment instruments are often used as tools to collect data on students' attitudes, knowledge, and behaviors throughout engineering education research.[1] These instruments have significant implications for ensuring high quality research, and the validity of conclusions are often derived solely from quantitative data. Typically, these types of instruments are designed for a specific context and purpose which ties directly to where and when their usage may be considered valid.[1]

In order to ensure that students' scores on such instruments are interpreted accurately in the implementation context, and any action taken as a result of those scores are appropriate and meaningful, survey instruments need to go through rigorous validation procedures. These methods have traditionally been quantitative in nature with a heavy focus on psychometrics.[2] However, more recently mixed-methods approaches have been used for survey validation by incorporating qualitative data to various degrees through a variety of methodologies.[3-5] Two flexible frameworks for mixing qualitative and quantitative methods during development and validation have been proposed.[6,7]

Despite this growing trend in mixed methods research, several authors have suggested more work is needed in determining appropriate and useful ways methods can be mixed to improve survey development, especially with regard to understanding the concept of validity in mixed methods research.[7,8] These broader methodological questions are of direct interest to the growing field of engineering education research where the mixing of quantitative and qualitative is a growing trend but mixed modes of validation have not been extensively explored.[9]

Key methodological questions that we explore in this paper include, "What would a process of mixed methods survey validation look like within an engineering education research context?" and "What are the potential benefits of blending research methods from different research paradigms that have traditionally been considered incompatible?" This paper seeks to provide a response to these questions by presenting a methodological approach for incorporating quantitative and qualitative data into the validation of a survey targeted at engineering students. Alongside this theory, we provide an example of this approach as applied to validating an existing survey developed outside of engineering within the context of engineering.[10]

Previous Examples of Mixed Methods Survey Development/Validation

The instrument development process typically relies most heavily on quantitative methods. However, there are studies that incorporate a small qualitative piece to support or enhance the still predominant quantitative methods. For example, Daigneault and Jacob[4] added an open-ended item about the participant's perceived accuracy of the instrument to a pilot test of their instrument. They performed a thematic analysis[11] on these responses, which helped identify semantic and conceptual aspects of the survey items that affected the respondents' scores. The qualitative data were utilized to suggest minor revisions to the items. The authors then utilized quantitative techniques to provide validity evidence for the revised instrument. While the use of qualitative data was limited in their study, the qualitative data helped enhance their validity

arguments through better understanding of the participant's context.[12] Further, this study exemplified the benefits of maintaining researcher flexibility (i.e. incorporating all available data, even an unexpected qualitative component) in the survey development and validation process.

Another common method when developing a survey instrument is to perform qualitative interview and observation studies to identify important topics and specific ideas for survey items.[13] For example, Crede and Borrego[3] utilized ethnographic observations and interviews to develop a survey about retention amongst graduate-level engineering students. Themes identified from this ethnographic study and previous research literature were used to develop survey constructs. The researchers then used participant excerpts to design specific survey items representative of the survey constructs, often using the direct language of the target population. The survey was revised through two rounds of pilot testing. The first of these rounds included review by members of the original ethnography sample and a panel of experts. The second of these rounds included pilot testing of the revised survey with a sample representative of the target population and review by an expert in the survey topic. The authors then performed checks of internal consistency on the final version of the survey across various demographics. Validity (or what the authors labeled as "rigor") in this example was attained through the utilization of effective qualitative and quantitative techniques during survey development iterations, as well as through traditional quantitative checks (internal consistency) on the survey's full implementation.[3] The particular strength of this method was that it utilized qualitative data to situate the survey in an authentic user context. Still, researchers who use this scale need to be cognizant of validity concerns if attempting to use the survey in a new, and potentially incompatible context. Further, the authors give little description or specific indication of how qualitative data factored into the revision and overall decision about the appropriateness of the final survey for the target population. Thus, this study is an example of sequential mixed methods where an initial qualitative study informed the development of quantitative survey validation.[9,13]

Convergent parallel mixed methods design[13] is another approach taken for survey validation. This approach attempts to incorporate qualitative and quantitative data on equal footing and truly integrate the findings rather than consider them in isolation.[6,7] Luyt (2012) applied a convergent parallel mixed methods design[13] to evaluate the use of three versions of a survey, each in a different language, with three different cultural groups.[6] Thematic analysis was used to identify conceptual, contextual, and semantic issues with the survey implementation with samples from the three distinct cultural groups. These findings were evaluated holistically with quantitative factor analysis and item analysis to evaluate and improve specific survey items.

Another approach was identified by Onwuegbuzie, Bustamante, and Nelson.[7] In their ten-step *Instrument Development and Construct Validation* framework, they discussed a number of approaches to writing survey items (e.g. literature review, Delphi study, personal reflection). These authors utilized both quantitative and qualitative data to validate and identify potential improvements to the survey. These phases include not only consideration of pure qualitative and quantitative data, but also crossover analysis of both types of data.[7] In other words, "quantized" qualitative data and "qualitized" quantitative data could be used to provide a holistic overview of the data.

Throughout these mixed methods approaches to survey validation, it is crucial to consider how incorporating and blending data from quantitative (postpositivist) paradigms and qualitative (constructivist, postmodernist, critical theory) paradigms might be accomplished most appropriately (e.g. resolving crossover validity issues, effective integration). Further, these frameworks present flexible, and thus under-prescribed methods, which could lead to inconsistent and potentially misappropriate application of the frameworks. Authors using these frameworks must critically consider the rationale for mixing data in their own context, and the appropriate techniques to utilize at each stage of their own survey development process. In this study we seek to generate a concerted discourse in this domain by presenting a framework for mixing qualitative and quantitative data for survey validation within engineering education.

Mixed Methods Survey Validation

According to Messick (1995, p. 741), "Validity is an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment."[14,15] The focus for establishing validity within the development of a survey is not necessarily on ensuring that the survey reliably measures a particular construct (e.g. through internal consistency reliability), but whether the meaning attributed to an assessment score within a particular context is appropriate. Thus validity is not a static property of an assessment instrument, but depends on the people being assessed and the assessment environment. Therefore, validation is a dynamic, iterative process that must carefully consider context in addition to theory.[1,6,8,12,14]

Messick (1995) acknowledged that all available evidence can and should be utilized when considering the validity of an assessment in a particular context, but also that certain evidence may be stronger than other types.[14] In particular evidence must be introduced to account for six particular aspects of the unified validity construct: content validity, substantive validity, structural validity, generalizability, external aspects of validity, and consequential validity. Figure 1 describes each of these validity types and demonstrates their interrelationships. Collecting data on multiple aspects of validity with a single methodological approach is challenging. A mixed-methods approach provides multiple means to collecting such evidence.

Messick (1995, p. 747) noted that "the validation process is scientific as well as rhetorical and requires both evidence and argument."[14] Thus, all collected validity evidence must be integrated to present a coherent argument for the survey's validity when applied to a target context. As demonstrated by Figure 1, Messick's aspects of validity are not simply a list of boxes that need to be checked on the way to validation, but a guide that helps one ensure sufficient validity evidence has been provided by considering the unique contribution of each of the six aspects when generating a holistic validity argument.

In the qualitative research space, Walther, Sochacka, and Kellam[12] take a similar approach to Messick. These authors describe six validation elements including theoretical validation, procedural validation, communicative validation, pragmatic validation, ethical validation, and process reliability. Walther and colleagues emphasize creating a conversation within any

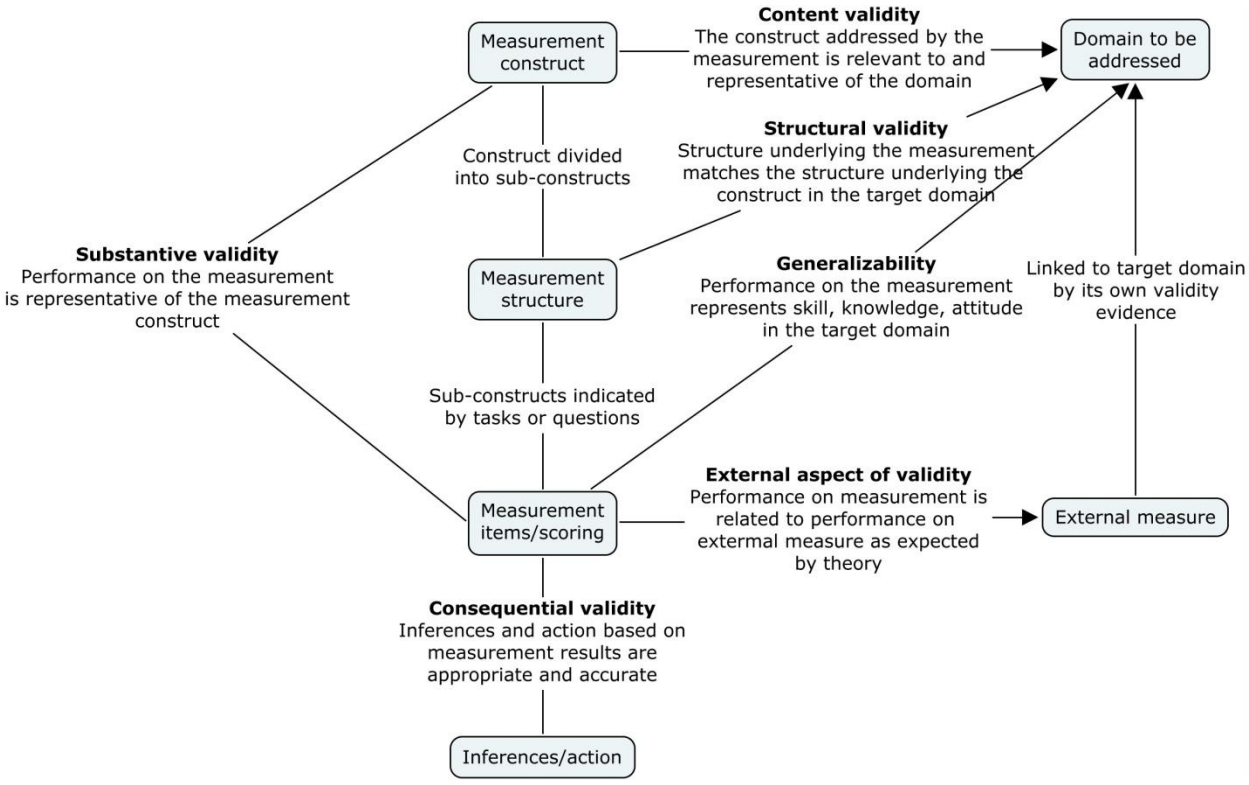

Figure 1. Mapping of Messick's unified theory of construct validity[14]

qualitative study to ensure each of these areas are addressed through "making" and "handling" data. They also stress researcher flexibility in finding the aspects that will support validation in their work. While the focus in Walther and colleagues' framework is not directly linked to survey validation, similar to Messick it emphasizes considering a complex array of aspects of validity and situating study findings in the authentic "social realities" of the participants, suggesting at least a common purpose for validation in both qualitative and quantitative domains.

Taking Messick's framework as a starting point, several scholars have utilized mixed methods research to collect validity evidence.[6-8,16] While these approaches have addressed construct validity through a variety of means, the focus is on integrating qualitative and quantitative data to lead to more comprehensive understanding of the survey and survey construct in the participant context.[2-7] In particular, qualitative and quantitative data can provide unique types of evidence—e.g., understandings of a what a particular construct looks like in the participants' context (qualitative) versus competency in the construct area as defined by survey authors who defined it based on literature and domain expertise (quantitative). The negotiation between these types of data serves as the basis for such improved understanding.[8]

Challenges to Integrating Data to Address Critical Aspects of Validity

When employing mixed methods research techniques, specifically when the data involves concurrent methods, paradigmatic considerations become a key concern. The central question

that must be addressed is, "Is it possible to embrace the same paradigm while implementing seemingly incompatible analysis techniques, and if so, how?" Traditionally, qualitative and quantitative paradigms have addressed different issues in relation to the validity of particular study.[8] Hence the challenge of mixing data to inform the validity of a survey in a particular context becomes understanding how both types of data can jointly inform the same conclusions.

Greene and Caracelli (1997) describe three viewpoints on mixing paradigms that would each suggest different methods for mixing data: purist, pragmatic, and dialectic.[17] A purist would suppose that data simply could not be mixed due to key differences between the data types. Thus, mixed methods survey validation would be inappropriate. On the other hand, mixed methods researchers embracing a pragmatic paradigm recognize that there are different philosophical traditions but are not concerned with potential incompatibilities between analysis procedures. Instead, the pragmatist focuses attention on the data itself and what analysis techniques will work best to answer the broader research questions. In the last case, the dialectical mixed methods researcher embraces each of the incompatible paradigms, and carefully navigates between them when working in the quantitative versus qualitative domains.

Previous approaches to mixed methods survey validation have focused on pragmatic methods. Here, researchers must provide the rationale for mixing the data and clarify how each piece of data informs the overall validity argument. Onwuegbuzie and colleagues (2010), for example, identified specific validity aspects met by piece of data (quantitative, qualitative, or mixed).[7] Luyt (2012) took a more holistic approach to integrating mixed data, where the focus of validation was on understanding what the survey content meant in the participants' context.[6]

A dialectic approach, as outlined in Figure 2, may help maximize utility of both qualitative and quantitative data.[17] Dellinger and Leech (2007) suggested the difficulty integrating data into a cohesive validity argument was primarily due to the quantitative necessity of a pre-defined construct and the qualitative practice of building understanding of a construct throughout a study.[8] In other words, the focus in qualitative work is understanding what the construct looks like in the participant context, thus ensuring a match between the meaning attributed to survey outcomes and the participant's social reality[12], while the focus in quantitative work is ensuring that the construct as defined in the target domain is accurately measured in the participant context.[6]

The dichotomy between these paradigms is precisely what makes both types of data useful in a validity argument. For example, a pre-defined construct represents the *measurement construct* (e.g. see Figure 1) which is the operationalized version of a construct from some target domain, and thus is defined based on previous literature, theory, and the researcher's judgment.[6,8] Through analysis of qualitative information, the researcher may determine how the measurement construct was interpreted in the participant's context, which may or may not match the measurement construct as initially operationalized. Through conversation between these perspectives, the nature of the construct in the participant context, and to some extent the more general content domain, becomes apparent. This dialogue becomes the basis of a validation of the survey, gauging whether the participant's interpretation aligns with the *a priori* theory for the survey construct, or if some core component of the participant's context that is tied to the research phenomenon is absent from the scale altogether.

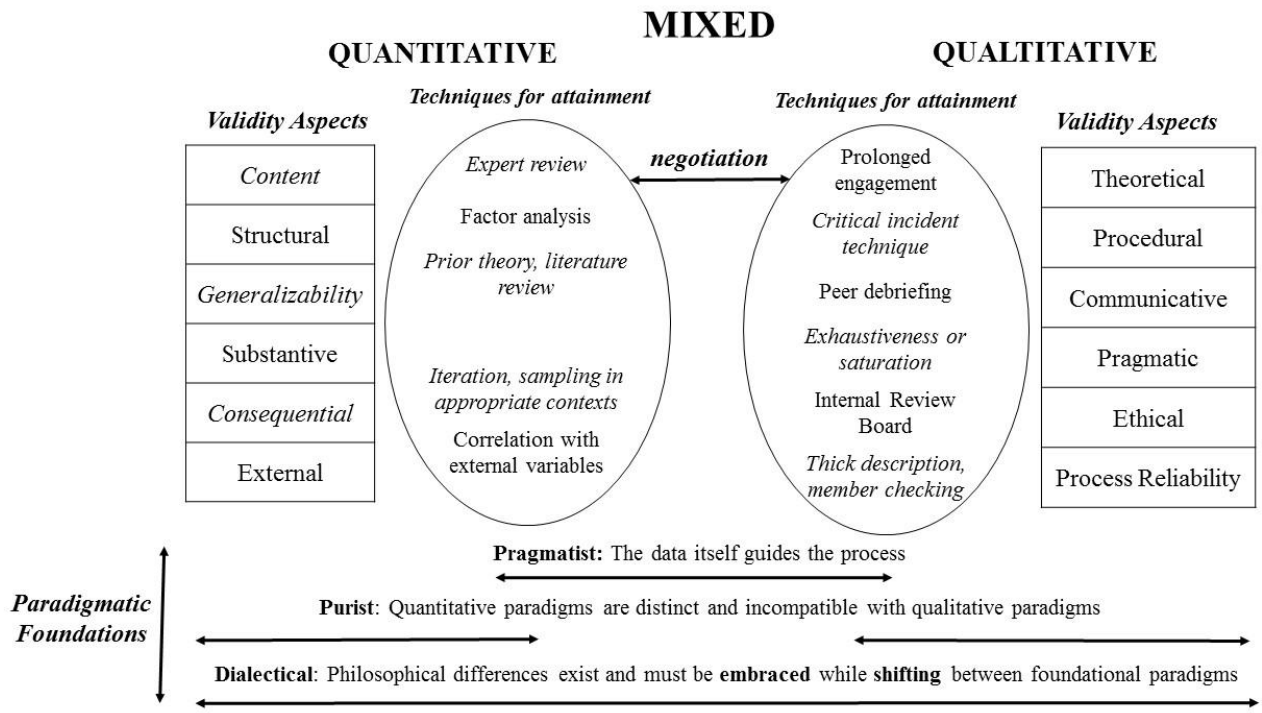

Figure 2. Negotiation between Qualitative and Quantitative Data

Dialectic Approach to Integrating Quantitative and Qualitative Findings

While previous pragmatic approaches provide flexible guidelines for integrating qualitative and quantitative data into survey validation, our dialectic approach placed the researcher as negotiator between the implementation context of the survey and previous literature and theory. In this respect, the researchers also acted as negotiator between quantitative evidence and qualitative evidence in light of their unique perspectives and the state of knowledge of the construct. This negotiation includes considering quantitative and qualitative data at two levels: item and concept, which are outlined in Table 1.

Table 1: Analysis Overview

Level	Guiding Questions	Quantitative	Qualitative
Concept	Is the construct relevant in the participant context? Does it share the same structure?	Factor analysis	Thematic analysis
Item	Are the survey items representative of the survey concepts and are they accurately interpreted in the context?	Item analysis	Content analysis

Item Level

Analysis begins at the level of survey items. Survey items are constructed to detail actions, conceptions, or beliefs that are intended to indicate presence of the construct in the participant domain. As Luyt noted, survey items can contain both semantically and contextually inappropriate wording in a specific participant context.[6] A particular word or phrase in the survey item may carry an unintended meaning in the participant context. It is also possible that the content of a survey item, while applicable to a different domain, may not represent the survey construct in the participant domain. For example, taking time each week to work on side projects may be linked to innovation among professional engineers,[18] but this action may carry different meaning to engineering students who may not link side projects to engineering innovation.

Item analysis is a quantitative technique for understanding the functioning of individual items in the overall survey.[2] Quantitative data take the form of item difficulty (represented by mean score among survey respondents). High item difficulty, especially in comparison to similar items, can indicate wording that is challenging or inappropriate in the participant context.[2] Corresponding qualitative data may take the form of content analyses of participant interpretations of the meaning of individual or groups of survey items.[19] This analysis can highlight aspects of the item wording that confuse or misdirect participants, or whether the item content is relevant to the construct in the participant context.

Concept Level

The concept level determines whether the constructs as framed in the survey are applicable in the participant context. This level goes beyond considerations of individual item wording and focuses on the constructs that are represented by linked groups of items or the entire survey.

Thus this level asks whether the core ideas addressed by the survey in the participant context accurately reflect critical constructs in the target domain. For example, does a survey meant to address a specific innovative mindset represent an innovative mindset to participants and do the collection of survey items reflect this mindset? The match is important in both the overall definition of the construct and the underlying structure of the construct.

At this level, quantitative data derive from factor analysis. Exploratory factor analysis, especially, is used when the researcher suspects the fit between the expected factor structure does not match the factor structure among a new group of participants.[2] Key data from factor analysis include the overall factor structure of the survey, i.e., whether all items load onto their intended factors, and factor loadings of individual items, i.e., items with high factor loadings represent key aspects of the factor in the participant context.

Qualitative data include the outcomes of thematic analysis.[11] Thematic analysis identifies the latent meaning demonstrated in a set of data. When thematic analysis is applied to interview responses representing participant interpretations of the intended survey constructs and the application of content described by the survey, themes can represent the underlying structure and interpretation of the survey constructs in the participant domain. In other words, what do the themes have to say about how participants interpret the constructs and what do connections between the themes have to say about the underlying structure? Thematic analysis would approach this question by giving voice to the participant, whereas factor analysis would answer this question through exploring the loading of items onto that factor. Both approaches can be open-ended (e.g. exploratory factor analysis or inductive coding) or seeking to justify if an *a priori* theory fits (e.g. confirmatory factor analysis or deductive coding). It should be noted that findings at the item level can also inform findings at the concept level by providing understanding of specific applications/examples of the broader construct.

Integrating Data at Each Level

Integration of qualitative and quantitative data requires considering what each piece of data means in an overall validity argument. This includes considering each piece of data at face value, demonstrating its unique contribution, as well as a negotiation of the meanings of the data in light of other corresponding data. Questions to ask at each stage are listed below. The answers to these questions structure understanding of the survey construct in the participant context, and thus help build an integrated understanding of the meaning of the survey. For ease of viewing, data can also be compiled in mixing tables.[13]

1. Considered in isolation, what do the qualitative data indicate about the appropriateness and applicability of the survey construct in the participant context?
2. Considered in isolation, what do the quantitative data indicate about construct among the participants and their context?
3. Considered in light of the quantitative data, what do the qualitative data indicate of the appropriateness and applicability of the survey construct in the participant context?
4. Considered in light of the qualitative data, what do the quantitative data indicate about construct among the participants and their context?

5. What do the quantitative data and qualitative data indicate about the construct-participant dyad?

A final stage of data integration is jointly considering conclusions at both the item and concept level to form a holistic understanding of the survey in the participant context and make recommendations for improving the validity of the survey in that context.

Applying the Dialectic Approach

To demonstrate and test the approach we outlined above, we applied this process to data collected during an implementation of an existing survey, the Innovation Behavior Scales survey,[20] to a sample of engineering students. This survey was initially developed based on extensive qualitative study of "innovative entrepreneurs" and quantitatively validated (EFA and external validity) on a representative sample.[20] Thus the current implementation represents a jump in context, both from professionals to students and entrepreneurs to engineers, requiring collection of new validity evidence. This survey addressed innovation competency via four sub-constructs: questioning, experimenting, networking, and observing. Each of these sub-constructs was addressed by four to six items outlining tendency of the survey respondent to demonstrate a behavior linked to the sub-construct.

Data Collection

The survey was completed electronically by 162 engineering students at a large research university. Demographically, all 14 engineering majors available at the university were represented in the survey. Further, students from all academic years, including master's and doctoral students, completed the survey. The survey sample was 43% female. Individual item scores were collected on a Likert-type scale from 1 to 5, 1 indicating that the item does not describe the participant well and 5 indicating that the item does describe the participant well.

Nine of the 162 students also completed in-person interviews during the same semester they completed the survey. Interviews lasted from forty-two to ninety minutes in length, with the average time a little over an hour. The interviews consisted of seven phases, four of which are relevant to the survey in this study. During the first phase, students discussed their perspectives and experiences related to innovation. The next phase focused on students perspectives and experiences related to four discovery behaviors and the survey items that described them. Students received a handout with the 19 survey items grouped into nameless dimensions representing the four discovery behaviors. Students were asked to describe nameless constructs. Next, students received the same handout with the dimensions being named and given a brief summary paraphrased from an account of the construct as intended by the survey developers.[20] The interviewees reacted to this information and discussed how they would or would not utilize the behaviors and their applicability to engineering work. During the third phase, students were asked to predict and react to their survey scores, as well as mean scores of their academic major and all of engineering. The next three survey phases modeled the first three, but focused on a different survey intended to measure empathy. The final phase asked students to describe any potential connections between the four discovery behaviors and empathy in terms of innovative

engineering design. Follow-up questions were asked to provide additional detail or clarity. Figure 3 describes the interview process.

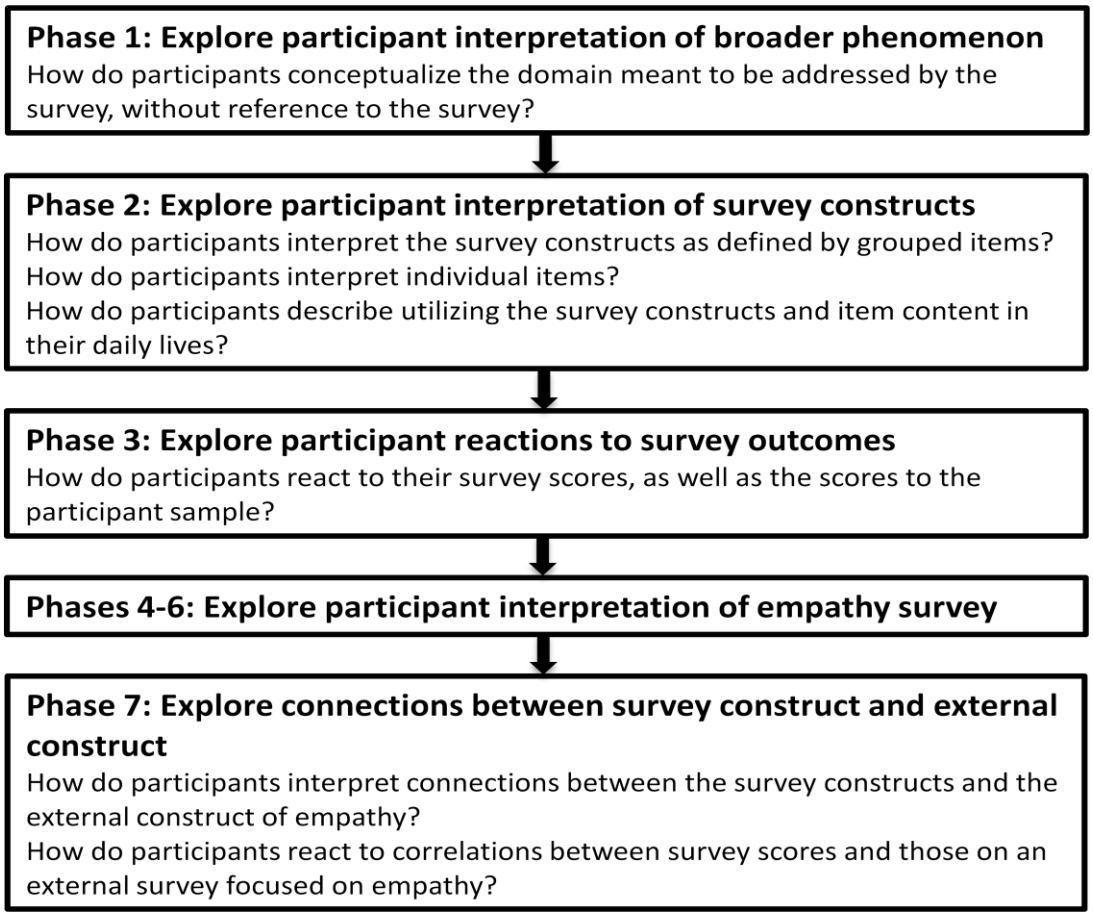

Figure 3. Outline of survey respondent interview procedure

Data Analysis - Quantitative

We performed an exploratory factor analysis and item analysis on all student survey responses. Exploratory factor analysis revealed a factor structure nearly identical to the initial survey implementation. This was supported through both the little jiffy criterion (all factors with eigenvalues above 1) and the scree plot. A four factor model, as intended by the initial survey, explained sixty-three percent of the variance in participant responses, which is an acceptable amount but would ideally be higher. This suggests a potential improvement to the survey/factor model in the participant context. All items aligned with their intended factors above the selected threshold of 0.4.

Item analysis included identifying the mean score among all participants for each item, and the percent of all participants who indicated that the item did not describe them well (indicated by a score of 1 or 2). Items that had a mean below 3 and/or a percent disagreed above 40% were considered as potentially troubling items.

Data Analysis – Qualitative

Thematic analysis was performed on the nine interview transcripts in order to identify how engineering students in the sample conceptualized the four survey sub-constructs.[11] Thematic analysis is useful in uncovering the latent meaning among a group of participants. By focusing on latent meaning, we were able to extract conceptualizations of the four sub-constructs beyond the wording of the survey. Thematic analysis was performed independently for each sub-construct and then integrated at a later stage. Qualitative content analysis[19] approached the interview data from a semantic, rather than latent, perspective to understand how students described the survey items and indicators.

Data Mixing

Following the guidelines described in the *Dialectic Approach to Integrating Qualitative and Quantitative Data* section, qualitative and quantitative data were integrated to form a comprehensive understanding of the survey constructs in the participant context at each level. These data were first viewed in a mixing table, which is meant to jointly present key findings from quantitative and qualitative analysis. Summaries of the findings from each individual analysis were encapsulated in the mixing tables to provide a visual overview and facilitate comparison. Table 2 is an example of a mixing table for all data related to the networking sub-construct. This was chosen as it represented the worst fit between the domain construct and the participant context.

Item Level Analysis

Quantitative Data

Quantitative item analysis revealed that students struggled with three of the four items. It should be noted that these represented the three lowest item means among all 19 questions of the study. The only item with a sample mean above 3 was: "I have a network of individuals whom I trust to bring a new perspective and refine new ideas." This item is similarly worded to the item that read: "I have a large network of contacts with whom I frequently interact to get ideas for new products and services," which had higher item difficulty. The primary difference is that the latter item linked networking to products and services. This result might suggest that the language related to products and services is not a critical aspect of networking in the student context.

Qualitative Data

Qualitative content analysis revealed that students connected to language relating to a "large network of contacts," but did not often discuss networking outside their home discipline and debated the applicability of conferences as representative of their networking. Instead, there was frequent mention of networking in local contexts among members of their own discipline and especially peers and classmates.

Table 2: Networking Mixing Table

	Quantitative	Qualitative
Item-Level	• Low item difficulty scores (Mean,% Disagree) ○ "I have a large network of contacts with whom I frequently interact to get ideas for new products and service." (2.83/45.4) ○ "I initiate meetings with people outside of my industry to spark ideas for a new product, service, or customer base." (2.64/49.1) ○ "I attend many diverse professional and/or academic conferences outside of my academic discipline/profession." (2.67/52.1). • One item without low item difficulty scores: ○ "I have a network of individuals whom I trust to bring a new perspective and refine new ideas." (3.39/21.5)	• Frequent and positive response to "Large network of contacts" language • Students indicated conferences were not useful for idea networking. • Infrequent mention of networking outside discipline; frequent mention of networking within discipline, especially classmates
Concept-Level	• Item factor loadings (all above .4 threshold) ○ .860 – "I have a network of individuals whom I trust to bring a new perspective and refine new ideas." ○ .749 – "I initiate meetings with people outside of my industry to spark ideas for a new product, service, or customer base." ○ .687 – "I have a large network of contacts with whom I frequently interact to get ideas for new products and services." ○ .405 – "I attend many diverse professional and/or academic conferences outside of my academic discipline/profession." (Also loads above .4 onto observing) • Items from other factors with correlations above .3 to networking ○ .464 – "I regularly observe peoples' use of products and services to get new ideas." (from observing) ○ .396 – "I am constantly asking questions to understand why products and projects underperform." (from questioning)	• Variety of modes, contexts, and purposes for networking often based on age/experience • Themes: ○ Interpersonal interaction: people-orientation to build and maintain networks, communication, interpersonal understanding ○ Active pursuit to building and maintain networks ○ Being able to find the right people based on trustworthiness, experience, variety of perspectives and expertise, and openness

Integrating Quantitative and Qualitative Data

Both the quantitative and qualitative data revealed evidence regarding the link between the networking construct and the language/actions/purposes used to represent it in the survey items. Four item-level findings included:

- Large network of contacts as an indicator of networking (sufficient item score, frequent mention in interview)
- Academic conferences as a locale for networking (low item score, negative mention in interview)
- Products and services as a purpose for networking (low item scores, no explicit mention in interview)
- Networking as local and within discipline (frequent mention in interviews)

Each of these findings was the result of different combinations of data. The "large network of contacts" language was identified through content analysis as being a strong indicator of networking in the participant context. The strongest item also referenced this phrase, indicating an appropriate link between the construct of networking in the target domain and its application in the participant context. Thus, this aspect of survey item wording represents a strong existing bond in the participant-construct dyad.

Academic conferences were specifically noted by the interview participants as events they did not find particularly useful for networking, especially with relation to sharing and developing ideas. This represented a potential mismatch between networking in the participant's context and networking as indicated in the survey. As mean score for the survey item related to academic conferences was low, there was also evidence that attendance of academic conferences is not a strong indicator of the general construct of networking among this group of respondents. Thus, this aspect of survey item wording represents a weak existing bond in the participant-construct dyad.

The issue with "products and services" was indentified initially through quantitative data. Specifically, comparison of mean scores on similar items revealed that the item containing the phrase may have played a role in the low score. Hence quantitative data suggests that networking toward products and services is not a strong indicator of networking among survey respondents. Content analysis revealed only one mention of products and services that was not in direct relation to networking. Thus, this phrasing may be linked to networking in the participant context and participants, on average, did not participate in that aspect. More evidence is needed to better understand the "products and services" element of the participant-construct dyad related to networking.

Conversely, content analysis suggested a prevalence of networking activities within one's academic discipline and local peer group. Thus, local interaction may be a significant aspect of the networking construct as projected onto the participant domain. No survey items, however, directly address local contexts. Thus, more evidence is needed to better understand the local networking element of the participant-construct dyad related to networking, specifically with relation to whether local networking as indicated by participants is a sufficient indicator of the more general construct of networking.

Concept Level Analysis

Quantitative Analysis

Exploratory factor analysis revealed that all four items intended as describing networking correlated to the networking factor above the .4 threshold. The same was true for other factors, though it should be noted that items from observing and questioning also demonstrated strong correlations to networking and one item from networking demonstrated a strong correlation to observing. Both external items that correlated to networking, and two strongly correlated networking items, referred to "products and services", indicating a potential connection to networking. Further, the most strongly correlated item was: "I have a network of individuals whom I trust to bring a new perspective and refine new ideas," which may be most closely aligned with how the overall networking construct projects into the participant domain.

Qualitative Analysis

Broadly, thematic analysis revealed that specific conceptualizations of networking (e.g., purpose and contexts used) among the participant group differed with respect to the age and experience levels of participants. However, three consistent themes representing the latent meaning attributed to networking were revealed. These included: the requirement of strong interpersonal interaction and communication, the importance of actively pursuing and maintaining networks, and the need to identify the right mix of people for a network in relation to the effective functioning of the network. These themes seem to align with the intended conceptualization of networking upon which the survey is based.[20]

Integrating Quantitative and Qualitative Analysis

Both the quantitative and qualitative data revealed evidence regarding the link between the networking construct as operationalized in the survey and the networking construct as it manifests in the participant context. Three concept-level findings included:

- Building a network of trusted others is a critical aspect (highest factor loading, qualitative theme)
- People-orientation characterizes networking (high internal factor loadings, external factor loading, qualitative theme)
- More "products and services" confusion (high internal factor loadings, external factor loading, qualitative theme, variety in qualitative themes)

The item with the strongest factor loading, as well as another high loading item, both had to do with building a network of trusted contacts, indicating an element of the survey construct that connected in the participant context. Further, thematic analysis revealed that identifying and actively pursuing the right mix of people for a network was a core element of networking. These represent a strong match between the way networking is operationalized in the survey and manifests in the participant domain.

The three highest loading survey items all explicitly characterize purposeful and frequent interactions with other people, indicating another element of the survey construct that connected to the participant context. Even items that were intended and contained content towards other survey constructs demonstrated a strong correlation with the intended networking factor. Once again, thematic analysis also revealed that strong interpersonal interaction is a necessity of networking. Thus, the people-orientation element demonstrates a strong match between the way networking is operationalized in the survey and manifests in the participant domain.

Findings regarding the notion of networking toward a product or service outcome were again mixed at the concept level. Factor loadings of networking and external items indicated that this aspect aligned with the networking construct as translated to the participant context. However, thematic analysis revealed a variety of purposes for networking, and working towards the development of products and services was not seen as an immediate effect. Thus, at least according to participants, the development of products and services was not an important element of networking. A preliminary interpretation of these findings would be that other elements tied to networking in the participant-construct dyad are represented by the items.

Conclusions

In order to ensure that students' scores, and any action taken as a result of those scores, are accurate and meaningful in the implementation context, surveys often go through rigorous validation procedures, such as rigorous psychometric analysis.[1,2] While these methods have a longstanding tradition, incorporation of qualitative data in a holistic mixed methods manner can improve understanding of the survey in the desired educational context. This paper presented a methodological approach for incorporating both traditional quantitative data and qualitative data into the validation of a survey targeted at engineering students.

Analysis, in addition to occurring at quantitative and qualitative levels, occurs at the item and concept levels. The results of integration of all findings helps develop a better understanding of matches and mismatches between the target construct as operationalized in the survey and the construct as it manifests in the participant domain. Consideration of all data in unison provides a holistic understanding of the survey constructs in the participant contexts and implications of using the survey as an assessment tool. These results can then be used to suggest potential revisions to the survey to better address the intended construct in the participant context.

This approach can work to complement previously developed approaches and provide specific guidelines and examples that may be useful for engineering education scholars who are developing survey instruments. Further work will focus on connections to larger topics of validity in mixed methods research, appropriate and utile data integration, and how a mixed methods survey analysis approach may fit into a larger framework for survey validation in educational research.

Acknowledgements

This material is based upon work supported by the National Science Foundation under EEC 1150874. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Douglas, K. A., & Purzer, Ş. (2015). Validity: Meaning and relevancy in assessment for engineering education research. *Journal of Engineering Education*, *104*(2), 1-11.
2. Hong, T., Purzer, Ş., & Cardella, M. E. (2011). A psychometric re-evaluation of the Design, Engineering and Technology (DET) survey. *Journal of Engineering Education, 100*(4), 800-818.
3. Crede, E., & 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.
4. Daigneault, P.-M., & Jacob, S. (2014). Unexpected but most welcome mixed methods for the validation and revision of the participatory evaluation measurement instrument. *Journal of Mixed Methods Research, 8*(1), 6-24.

5. Ungar, M., & Liebenberg, L. (2011). Assessing resilience across cultures using mixed methods: Construction of the child and youth resilience measure. *Journal of Mixed Methods Research*, 1558689811400607.
6. Luyt, R. (2012). A framework for mixing methods in quantitative measurement development, validation, and revision a case study. *Journal of Mixed Methods Research, 6*(4), 294-316.
7. Onwuegbuzie, A. J., Bustamante, R. M., & Nelson, J. A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research, 4*(1), 56-78.
8. Dellinger, A. B., & Leech, N. L. (2007). Toward a unified validation framework in mixed methods research. *Journal of Mixed Methods Research, 1*(4), 309-332.
9. Borrego, M., Douglas, E. P., & Amelink, C. T. (2009). Quantitative, qualitative, and mixed research methods in engineering education. *Journal of Engineering Education, 98*(1), 53-66.
10. Dyer, J., Gregersen, H., & Christensen, C. M. (2011). *The Innovator's DNA: Mastering the Five Skills of Distruptive Innovators*. Boston, MA: Harvard Business Press.
11. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.
12. Walther, J., Sochacka, N. W., & Kellam, N. N. (2013). Quality in interpretive engineering education research: Reflections on an example study. *Journal of Engineering Education, 102*(4), 626-659.
13. Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Los Angeles, CCA: Sage.
14. Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist, 50*(9), 741.
15. Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (pp. 13-103). New York: Macmillan.
16. Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality & Quantity, 43*(2), 197-209.
17. Greene, J. C., & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. *New directions for evaluation, 1997*(74), 5-17.
18. Berends, H., Vanhaverbeke, W., & Kirschbaum, R. (2007). Knowledge management challenges in new business development: Case study observations. *Journal of Engineering and Technology Management, 24*(4), 314-328.
19. Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115.
20. Dyer, J. H., Gregersen, H. B., & Christensen, C. (2008). Entrepreneur behaviors, opportunity recognition, and the origins of innovative ventures. *Strategic Entrepreneurship Journal, 2*(4), 317-338.