

## Using Participant Voices to Inform Validity Evidence in the Survey Development Process

### Abstract

This study presents qualitative findings from a larger instrument validation study. Undergraduates and subject matter experts (SMEs) were pivotal in early-stage development of a survey focusing on the four domains of Validation Theory (academic-in-class, academic-out-of-class, interpersonal-in-class, interpersonal-out-of-class). An iterative approach allowed for a more rigorously constructed survey refined through multiple phases. The research team met regularly to determine how feedback from undergraduates and SMEs could improve items and if certain populations were potentially being excluded. To date, the research team has expanded on the original 47 items up to 51 to address feedback provided by SMEs and undergraduate participants. Numerous item wording revisions have been made. Support for content, response process, and consequential validity evidence is strong.

### Purpose

Instruments used in survey research should be created through a rigorous iterative process using multiple sources of validity evidence gathered from experts and potential participants during the development process (Sondergeld, 2020; Desimone & Le Floch, 2004; Willis, 2004). The purpose of this research is to describe initial steps implemented in and qualitative findings from a larger survey development project of the *Validation Theory Survey (VTS)* which is being designed to measure in- and out-of-class academic and interpersonal validation of undergraduate students. One overarching research question drives this study: *How do typical participant and expert panel perspectives inform various forms of VTS qualitative validity evidence (content, response process, consequential) and item development?* Explaining the steps taken to create an instrument on Validation Theory, which focuses on ways underrepresented students can access higher levels of achievement through academic and interpersonal validation (Rendón, 1994; 2002), ties to this year's *AERA* theme as the feedback provided by our participants allowed for a wider range of voices to be heard to determine if institutions, schools, or classes are validating all of their student's identities (Cho et al., 2013; Collins & Bilge, 2020).

### Brief Literature Review

#### Validation Theory

Rendón (1994) published seminal research on the importance of validating student success experiences, specifically for first generation, underserved, non-traditional, and low-income college students. According to Rendón (1994; 2002), Validation Theory is comprised of six elements: (1) supportive and affirming processes; (2) student empowerment; (3) student development; (4) in- and out-of-class engagement activities; (5) seeing validation as a piece of the developmental process rather than "end" outcome; and (6) the importance of early occurrence in student academic careers. Rendón (1994; 2002) has expanded upon the fourth element (in- and out-of-class engagement activities) by identifying two forms of validation: interpersonal and academic. Interpersonal validation encourages "personal and social adjustment...both inside and outside of class"

(Rendón, 1994, p. 42). While academic validation has been shown to promote “student attitudes and behaviors that led to academic development” (Rendón, 1994, p. 40). These two forms of validation, in tandem, can strengthen students' perceptions of their abilities in higher education spaces.

Through this theoretical model, scholars have further explored how institutions can provide validating experiences to develop supportive learning environments for college students (e.g., Allen, 2016; Bauer, 2014). Further, Validation Theory researchers have been more concerned with *how* validating experiences are embedded into institutional structures for the purpose of identifying validating processes and systems rather than understanding *what* validating events are offered. With this in mind, it is not surprising that until recently, most Validation Theory studies have used qualitative methods alone (Kelly et al., 2021). The primary quantitative methods for assessing Validation Theory are two sections of the *Diverse Learning Environments (DLE)* survey which have been shown to broadly measure Validation Theory. To capture more specific information aligned with Rendón's (1994; 2002) Validation Theory model and inform targeted programmatic or institutional practices, our research team has begun the process of developing and validating a new instrument entitled the *Validation Theory Survey (VTS)*.

### **Developing and Validating Educational Instruments**

To address an inherent concern about instrumentation rigor in the social sciences, *The Standards for Educational and Psychological Testing* (2014) were collaboratively developed and released by American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME). *The Standards* (AERA et al., 2014) discuss the need to collect, evaluate, and document multiple forms of validity evidence for results and interpretations of quantitative instruments to be judged suitable for specified intents. Furthermore, stronger inferences may be drawn regarding instrumentation soundness when more validity evidence is gathered (AERA et al., 2014; Kane, 2016). While volumes of literature discuss numerous forms of validity evidence that may be obtained, *The Standards* encourage developers of educational and psychological assessments to minimally evaluate five specific types: content, response processes, internal structure, relationship to other variables, and consequential (AERA et al., 2014). See Table 1 for each validity evidence type operationally defined with typical supporting evidence aligned. For the purpose of this study, we will only focus on qualitative validity evidence – content, response process, and consequential – as initial quantitative field-testing data have not been analyzed yet.

## **Methods**

### **Study Context**

The Department of Homeland Security (DHS) and the National Science Foundation (NSF) administer the CyberCorps Scholarship for Service (SFS) program to recruit and train the next generation of government cybersecurity professionals. As of 2021, there are 84 colleges and universities that offer CyberCorps programs. The University serving as the site for this study runs an NSF-funded CyberCorps program called the CyberCorps Mentoring Scholarship Program (CMSP). CMSP specifically aims to strengthen the government's workforce by recruiting top talent with an emphasis on

recruiting women, underrepresented minorities, and veterans. To become a CMSP student, applicants must be a U.S. citizen or lawful permanent resident, enrolled full-time as an on-campus student, enrolled in a cybersecurity major, have a GPA of a least 3.0, be a pre-junior, and be able to obtain an appropriate security clearance. CMSP students participate in weekly cohort meetings which consist of various activities led by program faculty members. These activities include hands-on cybersecurity training, guest speakers from government agencies, resume writing, job-interview practice, goal setting, and many other activities explicitly designed to foster student validation. In terms of funding, CMSP students receive a stipend of \$6,250 during academic terms, full-tuition support, and \$6,000 in professional development (PD) funds. This PD allowance can be used for cybersecurity certification training, cybersecurity books, and travel to cybersecurity conferences or SFS hiring fairs.

### **Data Sources, Sample, & Analysis**

Multiple data sources were used to collect information from various samples to inform the three validity evidence components under investigation in this study.

#### ***Content***

An expert panel ( $n=9$ ) of SMEs in relevant content areas including higher education, student development, educational diversity and inclusion, and psychometrics completed a questionnaire. This questionnaire consisted of all *VTS* items, and the expert panel was tasked with indicating 1) item-theory alignment and 2) describing any specific issues/concerns they found with specific *VTS* items. Conceptual content analysis (Christie, 2007) was conducted to determine which items needed modification and how. *VTS* items were revised and then used in response process and consequential validity evidence data collection.

#### ***Response Process***

A diverse group of CMSP students ( $n=6$ ) completed cognitive interviews with the research team over Zoom following Willis's (2004) guidelines. After a brief training, participants were asked to "think-aloud" (i.e., speak out their thinking process) as they read and responded to *VTS* items. Additional probing items were asked about specific words and potential areas of confusion related to some items. The research team met regularly during this phase to review student transcripts and discuss the need for item revision. Student responses were coded for alignment with item intention (aligned or did not align). All feedback was considered, and additional *VTS* item revision was conducted after each cognitive interview.

#### ***Consequential***

At the end of the CMSP student ( $n=6$ ) cognitive interviews, participants were asked if there were any items they felt uncomfortable answering truthfully. Conceptual content analysis was again conducted to determine if any items needed to be modified based on participant negative perception or identified bias.

## Condensed VTS Qualitative Validity Evidence Findings

### Content

All expert panel participants reported each VTS item aligned with its intended Validation Theory construct component (i.e., academic-in-class, academic-out-of-class, interpersonal-in-class, interpersonal-out-of-class). There were, however, 10 item revisions made based on expert panel item-level feedback. Most ( $n=5$ ) item revision feedback was related to *wording choice* or *clarity*. For instance, instead of asking if instructors demonstrated “concern” for teaching students the panel suggested a shift towards a different word such as “interest”. All word changes added clarity to items. In addition to wording or clarity, expert panel participants suggested *adding three items*. Two items suggested were related to student disability status – 1) acceptance by instructor and 2) accessibility needs being met by instructors. These two items addressed a demographic of students that was not previously being assessed in the survey. Finally, it was suggested that *two items be broken apart* to better capture participant experiences. For example, a broad question related to student educational pursuit after college was broken into items focusing more specifically on student 1) professional aspirations and 2) career paths. A shift in this question allowed students to express the difference between short- and long-term goals. Item changes from this stage were incorporated before the first undergraduate cognitive interview.

### Response Process

Nearly all undergraduate responses (91.2%) aligned with researcher intended item meaning suggesting participants largely understood the items in the same way researchers intended. During cognitive interviews, undergraduates provided feedback for item revision in three main ways: clarity, examples needed, and response options.

In an attempt to improve item *clarity*, the phrases “when I am in-class” and “when I am out-of-class” was added to the start of every question. Before, this framing was provided at the beginning of the item set, rather than at the start of each item, and some undergraduates were responding with examples that did not coincide with the intended validation domain. For example, one undergraduate shared a misaligned response of how their previous academic experiences helped during job interviews while responding to an item for academic-in-class Validation Theory.

Providing additional *examples* in some items was requested by undergraduates to help them better respond to the intended item meaning. One instance of this was when undergraduates were asked if curriculum reflects their personal background, some were left unsure of what this meant. Two undergraduates shared common language of their personal background being “life and academic experience,” so an example was added to this item and drawn directly from cognitive interviewing feedback.

Many undergraduates were found to select the *response option* of “not applicable” in times that did not align with their explanation. For example, on the item “while I am out-of-class, instructors/staff actively reach out to students to help get them involved in college activities” an undergraduate shared this “never happened” and they would respond with “not applicable” or “disagree”. A different undergraduate shared that they would respond “agree” or “not applicable” to whether their “sexual orientation or gender identity was accepted by their instructor in-class”. Undergraduates appeared to

be using “not applicable” in lieu of responding to questions that they had an answer for when they were given the chance to talk aloud.

### **Consequential**

All undergraduates willingly complete the survey during their cognitive interview and did not skip any of the questions being asked of them even though they knew participation was voluntary and they could stop at any time for any reason, including if they felt uncomfortable. Further, no undergraduates mentioned perceiving any items to be biased or personally problematic for any reason other than those listed in the response process section.

### **Scholarly Significance of the Work**

In our study, incorporating the voices of SMEs through an expert panel, and giving undergraduates a chance to think aloud through items in an iterative survey development process allowed for greater clarity of items, a survey that will produce results with greater validity evidence (AERA et al., 2014), and more inclusive items that work to address intersectional identities (Cho et al., 2013; Collins & Bilge, 2020) of participants. Feedback provided in the expert panel (Sireci & Faulkner-Bond, 2014) and cognitive interviews of participants (Padilla & Benitez, 2014) gave important stakeholder insights that allowed for revision of items to match researcher intention. Furthermore, certain items were expanded or elaborated on to provide greater clarity to participants. During cognitive interviews, the opportunity to hear the thought process of undergraduate participants provided an understanding of how items were interpreted. Cognitive interviewing allows for research teams to hear from diverse perspectives in a population and should be incorporated in any survey development to improve outcome validity and reliability (Desmione & Le Floch, 2004; Willis, 2004). Overall, this qualitative component of the larger validation study has demonstrated strong content, response process, and consequential validity evidence and using a similar process should be considered by others interested in developing high-quality surveys.

## References

- Allen, T. O. (2016). (In) validation in the minority: The experiences of Latino students enrolled in an HBCU. *The Journal of Higher Education*, 87(4), 461-487.
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Artino, A. Jr., La Rochelle, J., Dezee, K., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE guide no 87. *Medical Teacher*, 36, 463–474. <https://doi.org/10.3109/0142159X.2014.889814>
- Bauer, K. (2014). Black male community college students and faculty-student engagement: Differences in faculty validation and time status. *Journal of Progressive Policy & Practice*, 2(2), 157-164.
- Beckman, T. J., Cook, D. A., & Mandrekar, J. N. (2005). What is the validity evidence for assessments of clinical teaching? *Journal of General Internal Medicine*, 20(12), 1159-1164.
- Bostic, J. D., & Sondergeld, T. A. (2015). Measuring sixth-grade students' problem-solving: Validating an instrument addressing the mathematics common core. *School Science and Mathematics Journal*, 115(6), 281-291.
- Christie, C. (2007). Content analysis. In Baumeister, R., Vohs, K. (Eds.), *Encyclopedia of social psychology* (p. 176). Sage.
- Cho, S., Crenshaw, K., & McCall, L. (2013). Toward a field of intersectionality studies: Theory, applications, and praxis. *Signs*, 38(4), 785-810. <https://www.doi.org/10.1086/669608>
- Collins, P. H., & Bilge, S. (2020). *Intersectionality*. Polity Press.
- Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, 26(1), 1-22.
- Kane, M. T. (2016). Explicating validity. *Assessment in Education: Principles, Policy & Practice*, 23(3), 198-211.
- Kelly, B. T., Raines, A., Brown, R., French, A., & Stone, J. (2021). Critical validation: Black women's retention at predominantly White institutions. *Journal of College Student Retention: Research, Theory & Practice*, 23(2), 434–456. <https://doi.org/10.1177/1521025119841030>
- Padilla, J. L., & Benitez, I. (2014). Validity evidence based on response processes. *Psicothema*, 26(1), 136-144.
- Rendón, L. I. (1994). Validating culturally diverse students: Toward a new model of learning and student development. *Innovative Higher Education*, 19, 35-51. <https://www.doi.org/10.1007/BF01191156>
- Rendón, L. I. (2002). Community college Puente: A validating model of education. *Educational Policy*, 16(4), 642-667.
- Sireci, S., & Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema*, 26(1), 100-107. <https://doi.org/10.7334/psicothema2013.256>
- Sondergeld, T. A. (2020). Shifting sights on STEM education instrumentation

- development: The importance of moving validity evidence to the forefront rather than a footnote. *School Science and Mathematics Journal*, 120(5), 259-261.
- Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema*, 26(1), 108-116.
- Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications.

## Appendix

**Table 1**

*Validity Evidence Types Operationally Defined with Aligned Data Sources*

<b>Validity Evidence</b>	<b>Operational Definition</b>	<b>Typical Supporting Evidence</b>
<b>Content</b>	Instrument item alignment (content) with the construct to be measured (theoretical trait).	Subject matter experts (SMEs) evaluating item-to-construct alignment and can be logical (quantitative) or empirical (qualitative) (Sireci & Faulkner-Bond, 2014).
<b>Response Process</b>	Participant responses or performance alignment with the assessment construct.	Cognitive interviews, think alouds, or focus group interviews, using a sample of typical respondents to verify that they interpret items and respond in ways developers imagined they would (qualitative) (Padilla & Benitez, 2014).
<b>Internal Structure</b>	Extent to which items and components of instrument reflect the construct.	Psychometric related to: (a) instrument dimensionality, (b) measurement invariance, and (c) instrument reliability (quantitative) (Rios & Wells, 2014).
<b>Relationship to Other Variables</b>	Instrument outcome associations with other variables hypothesized to be related (either positively or negatively).	Statistical testing between instrument outcomes and potentially associated variables (quantitative) (Beckman et al., 2005).
<b>Consequential</b>	Negative impact from completing assessment or item/ instrument bias.	Participant perceptions of instrument impact on them (qualitative) (Bostic & Sondergeld, 2015).