

Utilizing Think-Aloud Protocols to Assess the Usability of a Test for Ethical Sensitivity in Construction

Mr. Kenneth Stafford Sands II, Virginia Tech

Kenneth S. Sands II is a doctoral candidate and graduate assistant in Environmental Design and Planning at Virginia Tech. His research focus is on professional ethics and its pedagogy.

Dr. Denise Rutledge Simmons, Virginia Tech

Denise R. Simmons, Ph.D., is an assistant professor in the Myers-Lawson School of Construction & Civil and Environmental Engineering at Virginia Polytechnic Institute and State University. She holds a B.S., M.S., and Ph.D. in civil engineering and a graduate certificate in engineering education – all from Clemson University. Until 2012, she was the director of the Savannah River Environmental Sciences Field Station. Dr. Simmons has nearly fourteen years of engineering and project management experience working with public utility companies, a project management consulting company, and a software company. She is a registered professional engineer, project management professional and LEED accredited professional. Her research interests are in investigating students' development of leadership skills and other professional competencies and in student involvement in co-curricular activities. Dr. Simmons is a NSF CAREER award winner for her research entitled, "Investigating Co-Curricular Participation of Students Underrepresented in Engineering."

Utilizing Think-Aloud Protocols to Assess the Usability of a Test for Ethical Sensitivity in Construction

Abstract

Unethical conduct in the construction industry has been so prevalent that research has revealed it is one of the most corrupt industries in international business. Additionally, accredited construction and construction engineering programs are expected to provide professional ethics instruction to its students to meet the standards established by the profession. However, no assessment tool measuring a student's ethical sensitivity to professional issues in construction exists to determine whether this instruction improves student recognition of ethical issues of the construction industry. The purpose of this paper is to outline the use of think-aloud protocols in assessing the usability of a test instrument, the Test for Ethical Sensitivity in Construction (TESC). This paper details the process taken to develop and administer think alouds, the results of the think-alouds, and how the results assisted refinement of the TESC. The think-aloud sessions helped researchers improve the usability of the TESC, making revisions that accounted for mechanical or structural and cognitive issues of the instrument. It was found that think-aloud sessions improved the TESC by helping to improve terminology, remove extraneous information, understand the length of time required to complete the TESC, and improve the scoring rubric.

Introduction and Motivation

Unethical conduct in the construction industry is so prevalent that Transparency International, a non-governmental organization that monitors and publicizes corporate and political corruption in international development, reveals that the construction industry is more corrupt than any other sector of international economy¹. The FMI Corp. performed a survey for the Construction Management Association of America (CMAA) and reports, "the construction industry, in general, is tainted by prevalent acts that are considered unethical," adding that it is "tainted by illegal acts"².

As a part of the effort to curb unethical behavior, the mandate of construction related accrediting bodies have instituted requirements for literacy of ethics in the curriculum. The American Council for Construction Education (ACCE) requires ethics integration in construction curriculum (at least 1 semester hour). The ACCE also states:

*In addition, oral presentation, business writing, and ethics must be integrated throughout the construction-specific curriculum. Example courses in this division include: Human relations, psychology, sociology, social science, literature, history, philosophy, art, language, political science, and other appropriate courses.*⁴

Additionally, the Accreditation Board for Engineering and Technology (ABET) has ethics education requirements for construction programs. In the 2012-2013 ABET criteria for accrediting engineering programs, two students' outcomes of the list labeled (a) through (k) relate directly to ethics education and states that students should possess:

(c) an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability and
(f) an understanding of professional and ethical responsibility.⁵

Construction programs must include some form of ethical component to satisfy both the construction industry and accreditation requirements; however, no benchmark has been identified in construction education literature that addresses whether ethics instruction in construction education actually produces more ethically sensitive professionals. Furthermore, there is no standard to teaching ethics in construction education and construction programs have a multitude of options to satisfy the construction industry and accreditation requirements. In order to address this, we must investigate ethical competencies of construction students. In order to assess the competencies of construction students, an instrument was developed to measure the minimum competency in ethical conduct, i.e. the ethical sensitivity of construction students.

Ethical sensitivity is derived from Rest's four component model of cognitive moral development which includes: *moral sensitivity*, moral judgment, moral intent, and moral behavior. Component one, moral sensitivity, involves "imaginatively constructing possible scenarios, and knowing cause-consequence chain of events in the real world; it involves empathy and role-taking skills"⁶. For the purpose of this paper, ethics are associated with the rules of conduct recognized in a particular profession or area of human life and morals are associated with personal values that drive ethical behavior⁷. Adopting Rest's moral sensitivity component and placing it in an ethical context specifically for the construction industry, the term ethical sensitivity shall refer to the ability of an individual to recognize ethical issues of the construction industry.

Operationalizing Ethical Sensitivity to Test Student Ethical Sensitivity

To investigate the ethical sensitivity of construction engineering students, a test was developed called the Test for Ethical Sensitivity in Construction (TESC). This test is modeled after Clarkeburn's Test for Ethical Sensitivity in Science (TESS), which takes an open-ended questionnaire response approach to investigate the ethical sensitivity of university science students⁸. Such a test allows for qualitative spontaneous recognition of ethical issues that cannot be observed with check box methods of ethical sensitivity investigation.

The TESC is an eight vignette qualitative response test for engineering students in the construction industry with an original intended time of completion of 20 minutes. The TESC situates students in a mental avatar with the following original scenario:

You recently graduated and have been hired as a project engineer for Solid Construction, an established medium-sized commercial construction company in the US. As an entry-level employee, you must participate in a rotation program so that you are involved in various company operations.

The original eight-vignette test was developed from ethical concepts that were found to be major issues of the construction industry. The test embedded 14 issues of the construction industry into

the eight vignettes (i.e. 14 pilot test items). The 14 issues were: claims games, collusion, bid shopping, bid peddling, theft, abuse of client resources, unfair labor allocation through overtime, labor issues, frontloading, payment games, low competence of work performance, improper client relations, use of joint ventures to increase satisfaction of prequalification requirements, and bid rigging.

Response generation of the TESC requires students to read and reflect on the vignette and provide 3 statements or questions regarding each vignette. It is expected that within 3 statements or questions, a student would be able to identify the issue, or write statements or questions that showed a level of discomfort with the issues embedded in the vignette. Student response rating is based on a three level scoring rubric, indicating the level of student ability to recognize an issue:

- 1) Score of 0 = Inability to recognize issue
- 2) Score of 1 = Unable to name issue but demonstrates unfavorable response to issue
- 3) Score of 2 = Ability to recognize issue by name or thoroughly demonstrates unfavorable response to issue with explanation

Problem Statement/Aim

As a new instrument, the TESC requires refinement, review, and pilot testing. Certain unknowns of a new instrument need addressing. Is the TESC clear? How long will it take to administer the TESC? Essentially, instrument developers need to know the *usability* of the TESC prior to pilot testing. The aim of this paper is to answer these questions using the verbal report (i.e. think-aloud protocols) and to discuss benefits of testing the usability of an instrument.

Literature Review

Conceptualizing the Think-Aloud

The think-aloud is a type of verbal report method where useful research data is obtained via elicitation of test taker vocalization of self-generated ‘symbols’ while he/she performs a given task. The data elicitation method is referred to as ‘thinking aloud’ or ‘concurrent verbalization.’ This vocalization occurs at three different levels, vocalization of covert articulatory or oral encodings, description/explication of the thought content, and explanation of the thought processes or thoughts. The written transcripts of the verbalizations are called think-aloud protocols^{10,11}.

Adding to this, the think-aloud is an approach to cognitive assessment of a test taker, whereby test administrators are able to observe the thought processes of the test taker whilst performing a task (i.e. completing a test). It has been posed as a qualitative research tool designed to highlight test taker’s thought processes during the administration of a test. Additionally, think-alouds can yield valuable insights concerning perception, interpretation, and responses to test items¹² (p. 274).

Usability Testing and the Think-Aloud of the TESC

Assurance of instrument rigor should involve tests of usability. The most widely accepted definition of usability is from the International Organization for Standardization (ISO 9241-11) which defines it as:

The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use¹³ (p.1).

By another definition, usability is the ability of a user of a product to do what he or she wants to do with the instrument the way he or she expects to do it without hindrance, hesitation, or questions¹⁴ (p.4). For this paper, usability is the ability of test takers to complete the TESC as intended by the researchers with ease and efficiency. Satisfaction of the components of usability is important to rigorous data collection and results of actual TESC administration post piloting the TESC. Think-alouds are widely used for usability testing in various fields¹⁵ (p. 339) and are an essential element in achieving good results from small changes¹⁶. Barnum situates think-alouds for usability testing and identifies the key elements of a think-aloud protocol in usability tests as having the following:

- 1) *Defining the user profile*
- 2) *Creating task-based scenarios*
- 3) *Using think-aloud processes*
- 4) *Making changes and testing again¹⁶*

To be usable, an instrument must possess the ability of being useful, efficient, effective, learnable and satisfying^{14,16,17}. For the TESC, relevant components of usability require satisfaction prior to piloting. Table 1 summarizes each component of usability in general and how it relates to the TESC. As related to the TESC, each of these key elements will be discussed in another section.

Table 1.

Usability Components Relative to the TESC¹⁴

Component of Usability	In General	The TESC
Usefulness	Degree to which a product enables a user to achieve their goal/willingness to use the product/ease of use	The ease in which students complete the TESC
Efficiency	The quickness with which the user can accomplish their goal	The time in which the students complete the TESC
Effectiveness	The extent to which the product behaves in the way users expect it to	The construct validity of the TESC
Learnability	As part of effectiveness, refers to the user's ability to operate	Essentially, it is expected that training involves ethics

Component of Usability	In General	The TESC
	the system at some level of competence after training	education and thus the TESC should not require additional education to complete since the TESC is structured as a short answer test students should be accustomed to
Satisfaction	User's perception, feelings, and opinions of the product, usually captured through both written and oral questions	The follow-up survey post TESC think-aloud provides satisfaction data

Operationalizing the Think-Aloud in Research

There are different procedures of the think-aloud which uses various protocols to gain certain results; however, they tend to follow a general approach¹² (p. 274). This general approach has been used in a variety of disciplines for a variety of purposes. The think-aloud has been used in adjustment studies^{20,21}, clinical intervention^{23,24}, translational process research^{25,11}, cognitive psychology^{10,12}, and library studies¹⁵ as well as a host of other research studies.

Exemplar Think-Alouds in Engineering Education

In engineering education, various studies used think-aloud processes. A study by Atman and Bursic utilized the think-aloud protocol as a tool to document the processes engineering students use to approach design problems. The study was administered to two freshman engineering students and differences in approach to completing a design problem were documented. The study concluded that verbal protocol analysis (i.e. think-alouds) is a powerful tool for design process assessment.²⁶

Additionally, a study by Trenor and colleagues argues that the advantage of the think-aloud is to provide evidence that survey items are interpreted by participants as the test administrator intended, i.e. cognitive validity or construct validity. This study posed a cognitive validation model for use by engineering education researchers in survey design. Operationalizing the model, this study used ten participants across three iterations of think-alouds to see how they respond to taking a web-based survey identifying and quantifying engineering undergraduates' access to and activation of social capital in making decisions about entering and persisting in engineering. It was found that issues of the survey ranged from grammatical errors to "serious cognitive mismatches," which caused issues with interpretation and responses.⁹

The Think-Aloud Protocol

Think-alouds range in type based on its intended use. There is no archetype protocol for conducting think-aloud sessions. Literature provides general points of consideration for hosting a think-aloud session which can keep the session in line with the intended purpose. Some guiding principles include, telling the participant to voice confusion and assuring that the participant knows that the instrument is being evaluated and not the participant. Additionally, literature suggests that if a participant struggles with the test, the researcher should probe with particular questions to help fully understand the thought processes of the participant^{9,27,28}. However, the

protocol used for this study is minimally intrusive and does not probe participants with questions during the think-aloud sessions.

Adding to these key principles, Van Someren and colleagues outlines the practical procedures in obtaining think-aloud protocol data as summarized in Table 2²⁹.

Table 2.

*Summary of Practical Procedures in Obtaining Think-Aloud Protocols*²⁹

Procedure	Description
Setting	<ol style="list-style-type: none"> 1) Ensure subject is at ease, room is quiet, water is available 2) Focus should be on tasks with limited interference by experimenter for influence 3) Explain the purpose of research, protection/confidentiality of data, and that the interest is in the way problems are completed (the instrument) and not the ability of the test taker
Instructions	<ol style="list-style-type: none"> 1) Give instructions about the task, i.e. perform the task and say out loud what comes to your mind 2) Avoid elicitation of opinions, i.e. asking the participant to tell you what they think 3) Avoid making the instructions too long
Warming Up	<ol style="list-style-type: none"> 1) Allow participant time to practice thinking aloud <ul style="list-style-type: none"> • Assists in getting participant to verbalize thoughts and get comfortable with the task
Behavior of the Experimenter Prompting	<ol style="list-style-type: none"> 1) Experimenter should be restrained 2) Interfere ONLY when participant stops talking 3) If interference is required, experimenter should at most say: 'Keep on Talking'
Recording	<ol style="list-style-type: none"> 1) Think-alouds are typically recorded 2) Check and recheck your instrument prior to starting session 3) Inconspicuously check your instrument periodically to ensure that it works
Transcription of the Protocol	<ol style="list-style-type: none"> 1) Transcribe the session as verbatim as possible and code from transcription, not directly from audio tape 2) Everything in the session should be typed out (even interruptions as they influence results) 3) Note recognizable pauses and unusual silences in transcription by using dote, i.e. 'I guess...the answer' 4) Avoid unwarranted interpretation
Review	<ol style="list-style-type: none"> 1) When possible, reviewing protocol with participant can provide very useful information, especially when protocols are difficult to interpret <ul style="list-style-type: none"> • If review will be done, do it as soon as possible after the think-aloud session

Methodology

Situating and Using the Think-Aloud Protocol

The creation of a test is the “product of the thoughtful and sound application of established principles of test development”¹² and the test development process involves five fundamental stages as outlined in Figure 1.

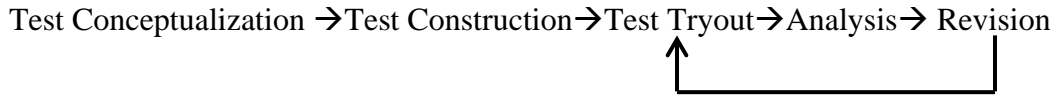


Figure 1: The Test Development Process¹²

Additionally, Trenor and colleagues summarize literature and presents a model for the instrument development process that highlights the concept of survey instrument cognitive validation. This model, in conjunction with the test development process, inspires the TESC instrument development process employed by this study (see Figure 2).

Participants

For the think-aloud sessions, three student participants were purposively selected. The small representative sample size was selected for the think-alouds are similar to other qualitative methods, seeking rich, in-depth data from a small sample³⁰ (p.432). Additionally, research has found that the first three participants detect the most severe problems of usability, additional participants are less likely to detect additional issues, and running fewer sessions saves time and money^{9,31}.

Recruitment

Student participants were initially recruited without promise of compensation; however, gift cards for lunch were provided at the end of the session for the participants’ time and contribution to the study (\$15 Panera Bread Gift Cards). The main consideration for recruitment was the student enrollment and class standing in a construction/construction engineering program. Special consideration was made to ensure two students were seniors,

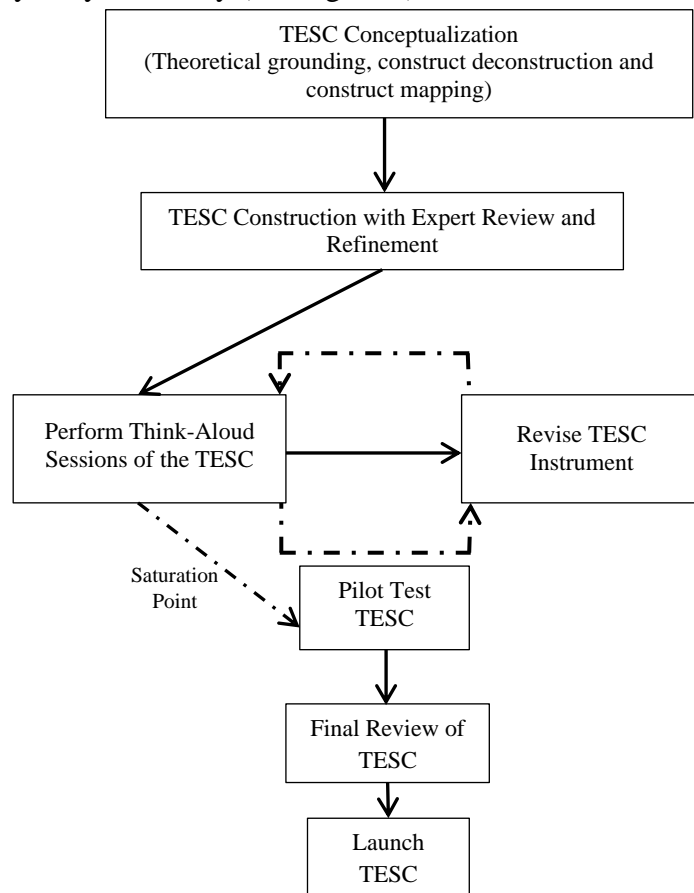


Figure 2: Think-Aloud Protocol in Context to the Development of the TESC (Derived from Trenor and colleagues⁹)

because the intended population of interest for the TESC is senior students who are expected to have thorough knowledge of construction ethics based on program requirements (see Table 3 for participant demographics).

Table 3.

Participant Demographics

Descriptor	Participant 1 (P1)	Participant 2 (P2)	Participant 3 (P2)
Program of Study	Construction Engineering	Building Construction	Building Construction
Student Standing	Senior	Junior (4 th Year)	Senior
Age	22	22	22
Sex	Male	Female	Male

Administration of the TESC Think-Aloud Sessions

Following the practical procedures of the model as provided by Van Someren and colleagues, with slight modifications, the following sections describe and discuss the TESC think-aloud procedures:

Setting(s)

The physical settings for hosting the think-aloud sessions were quiet spaces for recording with conditions conducive for test taking. To ensure participants' comfort, statements of purpose, their rights and administrator behavior were explained. Participants were advised that this was a test of the instrument and not of their ability. They were also advised that if they felt uncomfortable, they could discontinue the session at any time. The participants were encouraged to ask questions during the think-aloud, but warned that the administrator will not answer questions.

Additionally, participants were warned that the administrator will be checking the time frequently and this should not be misconstrued as being rude or otherwise, this is simply to collect data. Lastly, participants were warned about prompts when there is a lull in the session. Instead of verbally prompting participants to keep talking, participants were advised that taps on the desk (with example tapping) would be an indicator for them to keep talking.

Instruction(s)

Participants were verbally instructed to say aloud whatever comes to their mind during the process of performing the TESC. Due to the nature of the TESC, the participants were asked to express any questions or concerns about the instrument during the session but were told that the administrator would not be allowed to answer them. In addition, participants were told that the administrator will not tell them when the session is complete or when to stop providing responses to the TESC, and that they must determine this on their own.

Warming Up

In order to warm up, a simple exercise for thinking aloud was used. Participants were asked to spend a minute describing the room where the think-aloud session was held. The method employed is different from the task and was intentional, as the administrator did not want to hint at the type of task the participant would be performing.

Behavior of Observer

During the think-aloud, the administrator of the TESC was minimally intrusive, only prompting participants to continue talking via desk taps. The observer also avoided looking at the participant and attempted to take limited handwritten notes to avoid distracting the participant.

Recording

The think-aloud session was recorded with an audio recording device.

Transcription of the protocol

The protocol was immediately transcribed from the audio recording with as many audio cues transcribed as possible.

Revision and Feedback

After each session, the TESC was revised to account for issues discovered during the session. Issues were resolved and changes were made after the first think-aloud and prior to administering the next think-aloud session (see Table 3). After completion of three iterations of think-aloud sessions and prior to pilot testing the instrument, participant 1 was asked to review the final version of the TESC and provide feedback.

Results

The outcomes of the think-aloud sessions revealed various usability issues with the TESC (see Table 4 for examples issues revealed). Issues participants revealed during the think-aloud sessions were mechanical/structural issues (S) (typos, grammar) and/or cognitive (C) (relative information not recalled) issues, with the cognitive issues being more severe^{9,27}. Cognitive issues may impact a student's ability to correctly respond to the vignettes of the TESC. For the TESC, cognitive issues were more prevalent than structural issues. Seven cognitive issues were found with the TESC (see Table 4).

Think-aloud sessions also revealed attributes of the TESC that were simply indeterminable during initial survey construction. The attributes includes time to complete, determination if participants answer as intended, and elicitation of affective reasoning in processing responses.

Table 4.

Example Usability Issues Revealed by Think-Alouds

Participant	Revealed Issues	Correction Procedure
P1	(C) Project specifications too specific, distracts intent of	Changed specific project from airport renovation to simply 'renovation' for

Participant	Revealed Issues	Correction Procedure
	vignette[s]	vignettes 1-5
	(C) Specific documents (i.e. AIA documents) confused participant and extended time to complete	Removed unnecessary details irrelevant to responses sought for vignette
	(C) Ambiguous and confusing details regarding project duration for vignette 6	Removed ambiguity for project details
	(S) Typographical Error in Vignette 1	Adjustment made to punctuation and period placement
	(C) Descriptor items 4 and 5a were confusing, indirect and did not contain all responses desired of participant	Word “introduced” changed to “taught.” Added co-op work experience to 5a and changed ‘level’ to ‘amount of construction experience.’
P2	(C) Instructions are unclear	Simplified instructions
	(C) Vignette 8 contains unnecessary information and extends time to understand vignette, too large of a distractor	Removed unnecessary information not required to answer question
P3	(C) Collusion issue could not be identified by participant with ‘Joint Venture’ distractor in the vignette	Joint Venture distractor removed

Time to Complete

The average time of completion for the TESC for the think-aloud sessions was twenty-seven minutes. Participant P1 (as shown in Table 4) took approximately forty minutes to complete the TESC while P2 and P3 had completion times of twenty-one minutes each. Additionally, P1 on average took four and a half minutes to respond to each vignette, and 3 minutes to complete the descriptors section. P2 averaged a completion time of two and quarter minutes for each vignette and two minutes to complete the descriptors section. P3 averaged a completion time of two and half minutes, with a completion time of roughly 2 minutes for the descriptors.

Post Test Questionnaire Responses

A post-think-aloud questionnaire was immediately administered to participants once they determined that they had completed the TESC. The post-think-aloud questionnaire involved eight, five-point Likert-type scale items (1 = strongly agree to 5 = strongly disagree), two binomial items (yes/no), and an open-ended feedback item. Results of the post-think-aloud questionnaire are in Table 5.

Table 5.

Results of Post-Think-Aloud Questionnaire

Questionnaire Item	P1	P2	P3
The purpose of this test was clear	1	2	1
I found the test unnecessarily complex	4	4	4

Questionnaire Item	P1	P2	P3
I thought the test was easy to use [complete]	3	3	1
I found the various functions in the test were well integrated	2	3	2
I thought there was too much inconsistency in this test	4	5	4
I would imagine that most people would complete this test quickly	4	4	3
I found the test cumbersome to complete	3	4	5
I need to learn a lot of things before I could get going with this test	1	3	4
Was the test too long?	No	No	No
Did any of the questions make you feel uncomfortable?	No	No	No

Additional feedback received regarding the TESC included:

P1 – Neutralize the questions [Vignettes], i.e. could remove specific project documents, or project specifications.

P2 – I am a fast test taker, so it didn't feel long to me, but I know some of my peers would take a while completing it. Possibly have students choose one vignette to respond from sets of two vignettes.

P3 – It is hard thinking aloud, not a commonly done practice. That is all.

P1 Pre -Pilot Feedback

After all think-alouds were completed and prior to sending the instrument out for pilot testing, the first participant was asked to provide a review of the revised TESC and provide feedback concerning each vignette (see Table 6). Based on the feedback of the original participant, the modifications made to the TESC assisted mainly in its clarity and ability to complete the test.

Table 6.

P1 Pre-Pilot Feedback

Vignette	Feedback
1	“This problem is set up much clearer than the last; very few people know much about airports or their building so this is easier to think on.”
2	“Still clear and understandable”
3	“Again well written, easy to comprehend.”
4	“Doesn't seem to have changed much for this scenario [vignette].”
5	“Good change, makes it easier to understand what's going on.”
6	“I think if you changed ‘anger’ to ‘angry’, it would make the sentence flow a little better, making it easier to understand.”
7	“By changing the scenario [vignette] to be about a project the company has no experience in, it makes it a little more easy to understand. Last time was a little too specific.”
8	“This scenario is set up to be MUCH clearer than previously. Good change.”

Discussion

To guide the discussion of this section, Rubin and Chisnell's usability components are used. As discussed, the usability components are: usefulness, efficiency, effectiveness, learnability, and satisfaction. Unintended affective responses emerged from think-aloud sessions and will be discussed.

Usefulness

The feedback of P1 reveals that the vignettes have improved and are clearer than the original TESC. Additionally, the follow-up questionnaire highlights an improvement in ease of use of the instrument. The responses of P1 and P2 indicate that with revisions to the TESC, the test became easier to use.

Efficiency

Based on the three think-aloud sessions, the TESC took an average of twenty-seven minutes to complete. The time to complete the TESC was vastly different from P1 (forty minutes) to P2 and P3 (twenty-one minutes each). These results may suggest that many issues of clarity in understanding and responding to the vignettes were resolved after the first session. Additionally, the follow-up questionnaire reveals that students did not find the test administration to be too long, and with each revision, the test is less cumbersome to complete.

Effectiveness

Even though the instrument had various cognitive issues, it still did not detract from the effectiveness of the TESC. Observations revealed that students were able to respond to vignettes in ways intended by the researchers with some exceptions.

There appeared to be an issue with three items embedded in various vignettes. The items included use of joint ventures to increase satisfaction of prequalification requirements, unfair labor allocation/use, and abuse of public/client resources. What was revealed is that the participants did not correctly understand these three items. For the joint venture item, a reason for this could be that it is an issue relevant to a particular region and is not prevalent in the United States. Additionally, possible reasons unfair labor allocation/use, and abuse of public/client resources was not understood is because these items may have had more dominant distractors in the vignette, or that the issues were not explicit enough. The lack of understanding of these items was accounted for in the scoring and analysis of the TESC.

Additionally, an improvement was made to the scoring rubric based on the think-alouds. Due to the use of words such as 'plain clothes', a response generated by P1 revealed student recognition of an issue of safety. In responding to vignette five, P1 stated, "if the kid's there just to look and watch, um, make sure he is dressed in the proper PPE [personal protective equipment]", P2 stated, "also ask that he wear PPE", and, P3 also identified that this vignette may involve a "safety hazard." Consideration is being made to include this as an issue for scoring the TESC. Changes will not be made to this vignette prior to pilot testing the TESC.

Learnability

There were no issues of learnability associated with the TESC. Participants were able to begin and complete the TESC without issue, under their own ability, without additional training. What is a concern for learnability is P1's response to the post TESC questionnaire item, "I need to learn a lot of things before I could get going with this test," where he responded 'strongly agree.' In future work, it may be interesting to engage such participants and inquire what things needed to be learned before a participant could get going with the test.

Satisfaction

Based on the responses of the post TESC questionnaire, participants seem satisfied with the completion of the instrument. On average, the participants rated 'somewhat disagree' that the test had too much inconsistency. They also 'somewhat agree' that the various functions [components] of the TESC were well integrated. Overall, there were no opinions of dissatisfaction with the TESC.

Unintended Affective Response

Even though the original intention of the think-aloud sessions for the TESC is to test the usability of the instrument, another attribute of the TESC emerged. The TESC is a survey instrument intended to measure the cognitive ability of students to recognize ethical issues in the construction industry; however, the think-aloud sessions reveal that the TESC also elicits affective reasoning in processing responses to the TESC not explicitly written as a response on the TESC. Example statements include:

P1: "I mean, I am really not a fan of this owner, and am kind of curious as to why am working here"

P2: "I would be questioning, my boss' morals because it seems like, everything seems to be easily falling into place and now that we've received this bid, he's trying to, bring in other companies for different prices, so upon questioning his morals"

"the company does not have an ethical focus, as an individual, I would be wondering, whether I would want to remain with the company with such low ethical standards"

Conclusion & Future Work

The think-alouds were useful in determining usability characteristics of the TESC that were indeterminable during original instrument development (i.e. usefulness, efficiency, and effectiveness). For the TESC, it was found that major cognitive issues hampered usability characteristics. Each think-aloud session offered unique contributions to the development of the TESC.

Overall, the think-aloud process improves usability by helping researchers discover mechanical/structural errors in the survey instrument. Even more importantly, it assists

researchers account for, and make revisions to the instrument for cognitive issues that are more severe when trying to make observations that reflect what the researcher truly intended.

The benefits of think-aloud sessions warrants its use in survey development for instrument refinement and to obtain measures of instrument usability. It should be a part of a researcher's repertoire of instrument development techniques.

As research suggests, there are other steps to complete to gain sufficient evidence of cognitive validity of the TESC beyond think-alouds, but these protocols are a start⁹. The next step for the TESC is to pilot test the refined instrument.

Bibliography

1. Transparency International (2005). "Global corruption report 2005." *Report Prepared for Transparency International*, Berlin, Germany, 31–155.
2. FMI (2004). "Survey of construction industry ethical practices, ethical practices today." <http://www.accehq.org/documents/ethics_survey.pdf> (June 1, 2013).
3. Construction Management Association of America (CMAA) (2013). "About CMAA." <<http://cmaanet.org/about-cmaa-0>> (Nov. 25, 2013).
4. American Council for Construction Education (ACCE) (2010). "Document 101 Accreditation Manual." <<http://acce-hq.org/accreditationprocedures.htm>> (Oct. 1, 2012).
5. Accreditation Board for Engineering and Technology (ABET) (2012). "Accreditation Policy and Procedure Manual (APPM), 2012-2013" <<http://www.abet.org/DisplayTemplates/DocsHandbook.aspx?id=3146>> (Oct. 2, 2013).
6. Rest, J. R. (1984). "The major components of morality." *Morality, moral behavior, and moral development*, John Wiley & Sons, Inc., Hoboken, N.J, 24-38.
7. Fellows, R. (2003) "Professionalism in construction: Culture and ethics." *Proc., CIB TG23 International Conference, CIB*, Hong Kong, China.
8. Clarkeburn, H. (2002). "A test for ethical sensitivity in science." *Journal of Moral Education*, 31(4), 439-453.
9. Trenor, J., Miller, M. K., and Gipson, K. G. (2011) "Utilization of a Think-Aloud Protocol to Cognitively Validate a Survey Instrument Identifying Social Capital Resources of Engineering Undergraduates." *Proc., American Society for Engineering Education*, American Society for Engineering Education.
10. Ericsson, K. A., and Simon, H. A. (1980). "Verbal reports as data." *Psychological review*, 87(3), 215.
11. Jaaskelainen, R. (2010). "Think-aloud protocol." *Handbook of Translation Studies*. Amsterdam: John Benjamins Publishing, 371-373.
12. Cohen, R. J., Swerdlik, M. E., and Phillips, S. M. (2013). *Psychological testing and assessment: An introduction to tests and measurement*, McGraw Hill Companies, New York, NY.
13. Jokela, T., Iivari, N. and Tornberg, V. (2004). Using the ISO 9241-11 Definition of Usability in Requirements Determination: Case Study. In HCI 2004, Vol. 2 (Eds, Dearden, A. and Watts, L.). Research Press: Leeds, pp. 155-156.
14. Rubin, J., and Chisnell, D. (2008). *Handbook of usability testing: how to plan, design, and conduct effective tests*, Wiley.
15. Van Den Haak, M., De Jong, M., and Jan Schellens, P. (2003). "Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue." *Behaviour & Information Technology*, 22(5), 339-351.
16. Barnum, C. M. (2010). *Usability testing essentials: ready, set... test*, Elsevier.
17. Bevan, N., and Macleod, M. (1994). "Usability measurement in context." *Behaviour & Information Technology*, 13(1-2), 132-145.
18. 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.
19. Anderson, N. J., Bachman, L., Perkins, K., and Cohen, A. (1991). "An exploratory study into the construct validity of a reading comprehension test: Triangulation of data sources." *Language Testing*, 8(1), 41-66.

20. Kendall, P. C., Williams, L., Pechacek, T. F., Graham, L. E., Shisslak, C., and Herzoff, N. (1979). "Cognitive-behavioral and patient education interventions in cardiac catheterization procedures: the Palo Alto Medical Psychology Project." *Journal of Consulting and Clinical Psychology*, 47(1), 49.
21. Sutton-Simon, K., and Goldfried, M. R. (1979). "Faulty thinking patterns in two types of anxiety." *Cognitive Therapy and Research*, 3(2), 193-203.
22. Davison, G. C., Vogel, R. S., and Coffman, S. G. (1997). "Think-aloud approaches to cognitive assessment and the articulated thoughts in simulated situations paradigm." *Journal of Consulting and Clinical Psychology*, 65(6), 950.
23. Haaga, D. A., Davison, G. C., McDermut, W., Hillis, S. L., and Twomey, H. B. (1993). "'States-of-mind' analysis of the articulated thoughts of exsmokers." *Cognitive Therapy and Research*, 17(5), 427-439.
24. Schellings, G., Aarnoutse, C., and van Leeuwe, J. (2006). "Third-grader's think-aloud protocols: Types of reading activities in reading an expository text." *Learning and Instruction*, 16(6), 549-568.
25. Gerloff, P. (1986). "Second language learners' reports on the interpretive process: Talk-aloud protocols of translation." *Interlingual and intercultural communication. Tübingen: Gunter Narr*, 243-262.
26. Atman, C. J., and Bursic, K. M. (1998). "Verbal protocol analysis as a method to document engineering student design processes." *Journal of Engineering Education*, 87(2), 121-132.
27. Willis, G. B., Royston, P., and Bercini, D. (1991). "The use of verbal report methods in the development and testing of survey questionnaires." *Applied Cognitive Psychology*, 5(3), 251-267.
28. Stone, D., Jarrett, C., Woodroffe, M., and Minocha, S. (2005). *User interface design and evaluation*, Morgan Kaufmann.
29. Van Someren, M. W., Barnard, Y. F., and Sandberg, J. A. (1994). *The think aloud method: A practical guide to modelling cognitive processes*, Academic Press London.
30. Fonteyn, M. E., Kuipers, B., and Grobe, S. J. (1993). "A description of think aloud method and protocol analysis." *Qualitative Health Research*, 3(4), 430-441.
31. Virzi, R. A. (1992). "Refining the test phase of usability evaluation: how many subjects is enough?" *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34(4), 457-468.