ASEE's 123rd Annual • Conference & Exposition • New Orleans, LA • June 26-29, 2016

New Orleans JAZZED about Engineering Education

Paper ID #14697

Exploring Interviews as Validity Evidence for the Engineering Professional Responsibility Assessment

Dr. Nathan E Canney, Seattle University

Dr. Canney teaches civil engineering at Seattle University. His research focuses on engineering education, specifically the development of social responsibility in engineering students. Other areas of interest include ethics, service learning, and the role of the public in engineering decisions. Dr. Canney received bachelors degrees in Civil Engineering and Mathematics from Seattle University, a masters in Civil Engineering from Stanford University with an emphasis on structural engineering, and a PhD in Civil Engineering from the University of Colorado Boulder.

Dr. Angela R. Bielefeldt, University of Colorado - Boulder

Angela Bielefeldt is a professor at the University of Colorado Boulder in the Department of Civil, Environmental, and Architectural Engineering (CEAE). She serves as the ABET assessment coordinator for the department. Professor Bielefeldt is the faculty director of the Sustainable By Design Residential Academic Program, a living-learning community where interdisciplinary students learn about and practice sustainability. Bielefeldt is also a licensed P.E. Professor Bielefeldt's research interests in engineering education include service-learning, sustainable engineering, social responsibility, ethics, and diversity.

Greg Rulifson P.E., University of Colorado - Boulder

Greg Rulifson is a Civil Engineering doctoral candidate focused on qualitative engineering education research while also completing the Engineering in Developing Communities certificate. Greg earned his bachelor's degree in Civil Engineering with a minor in Global Poverty and Practice from UC Berkeley where he acquired a passion for using engineering to facilitate developing communities' capacity for success. He earned his master's degree in Structural Engineering and Risk Analysis from Stanford University. His upcoming dissertation will focus on how student's connections of social responsibility and engineering change throughout college and what are the social responsibility-related reasons some students choose to leave engineering.

Exploring interviews as validity evidence for the Engineering Professional Responsibility Assessment

Abstract

This research paper explores the use of interviews as validity evidence for a survey instrument, the Engineering Professional Responsibility Assessment (EPRA). The EPRA tool uses 50 Likert items to assess engineering students' attitudes toward personal and professional social responsibility. Validity evidence for EPRA based on internal structure has been previously examined using structural equation modeling and multidimensional item response theory; both showed strong evidence. This paper expands the body of validity evidence, specifically evidence based on relations to other variables: interview responses.

Data came from interviews with 24 engineering students after they had completed the EPRA survey. To compare interview data to Likert items, a coding rubric correlating to Likert scores was developed with feedback from engineering education experts. Once language for the rubric was solidified, two researchers coded each interview, resulting in a score for each dimension for each participant. Interview and survey scores were compared using Spearman's rank order correlation coefficient and the Wilcoxon signed-rank test. Results showed that four of the 24 respondents had significant correlation ($p<0.05$) and two had suggestive correlation ($p<0.10$) between their scored interviews and EPRA scores across all dimensions. Eighteen respondents rejected the hypothesis of difference ($p>0.05$). Across the eight dimensions, three had strong correlation ($p<0.05$) and three rejected the hypothesis of differences ($p>0.05$). Only one dimension showed both correlation and a rejection of difference. The process of using interview data as evidence of validity for a survey instrument is appealing. Surveys tend to compress complex issues into bin-able categories, perhaps oversimplifying the nuances of attitudes and beliefs. This exploration steps through one way in which validity evidence may be explored, by coding interviews using a rubric and comparing scores with survey results. Suggestions for producing better results in future studies, such as more targeted interviews, are given.

Introduction

As engineering education continues to broaden the skills, knowledge, and attitudinal dispositions that it hopes to foster in students, the need for a broad range of assessment approaches and tools also grows. The ABET criteria focused on professional skills have challenged engineering departments to find new ways to assess difficult concepts like an understanding of ethical and professional responsibility and an understanding of the social context of engineering projects[1,2]. Moreover, many of these skills and dispositions may be reliant upon students' past experiences before coming to college, are heavily nuanced, and develop slowly, such that the benefits of positive development may not even be realized until well after graduation. These circumstances necessitate the need for assessment approaches that can be used for both programmatic review and for engineering education research that address the development of professional skills in engineers. Toward meeting this need, this paper builds upon previous work to provide further evidence of validity for the Engineering Professional Responsibility Assessment (EPRA) as a tool for assessing elements of professional development in engineering students.

The Engineering Professional Responsibility Assessment

EPRA was developed to assess engineering students' attitudes toward personal and professional social responsibility, operationalizing the Professional Social Responsibility Development Model (PSRDM) [3]. In this context, social responsibility is seen as feelings of desire or obligation to help others who are in need through one's professional abilities, with particular emphasis on underserved and marginalized groups. The framework consists of eight dimensions, summarized in Table 1. The survey instrument, EPRA, consists of 50 Likert-items on a 7-point metric (from 'Strongly Disagree' to 'Strongly Agree') that are intended to assess these eight dimensions. The number of items attributed to each dimension are also shown in Table 1. In addition to the Likert-items, the tool also includes several open-ended questions about social responsibility attitudes and possible influences on those attitudes, volunteer experiences, expectations of future career attributes, and demographic information. This tool has been developed through multiple iterations; the process and results are described in detail in [4].

Table 1. Descriptions of the eight dimensions of the PSRDM [5]

Dimension	Abbreviation	Definition	No. of Items on EPRA
Awareness	*Aware*	An awareness that others are in need	5
Ability	*Ability*	A recognition that one has the ability to help others	4
Connectedness	*Conn*	A feeling of moral obligation, responsibility, or social requirement to help others	4
Base Skills	*Base*	With an expectation that all engineers value the technical skills, this dimension focuses on views of professional skills (i.e. communication, lifelong learning, teamwork, management, ethics, professional responsibility, understanding social and global impacts of engineering, etc.) and the role that they play for a professional engineer.	5
Professional Ability	*ProfAb*	A recognition that engineers or the engineering profession has the ability to help others and/or solve social issues	4
Analyze	*Analyze*	A recognition of the importance of including social aspects in the engineering process, including community feedback, a broad sense of stakeholders, etc.	5
Professional Connectedness	*ProfCon*	Addresses issues of responsibility or obligation that an engineer or the engineering profession may have to help solve social problems or help others	19
Costs/Benefits	*CB*	Discussion of the costs and/or benefits associated with engaging in socially responsible behavior, such as service.	4

Evidence of Validity

Key in the development of any assessment instrument is the examination of evidence of validity. "Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of test" (p. 11)[6]. There are many different ways in which evidence of validity may be explored. The most recent version of the *Standards*[6] separates sources of validity evidence into the following categories:

- Evidence based on test content
- Evidence based on response processes
- Evidence based on internal structure
- Evidence based on relations to other variables
- Evidence based on consequences of testing

The thorough development of a survey instrument should include the exploration of several different sources of validity evidence in order to build mounting confidence that an instrument actually measures what it claims to and to fully understand the circumstances in which these assumptions are accurate and when they are not. In the development of the EPRA tool, validity evidence based on response processes was examined through cognitive interviews (a subset of the interviews used in this validity exploration) that addressed the content of questions and student understanding of various items. Validity evidence based on internal structure was examined through structural equation modeling and Multidimensional Item Response Theory, specifically Rasch modeling; both showing strong evidence of validity. Convergent validity evidence based on relations to other variables was explored through correlations between respondent Likert-item responses and previous volunteer activities – believed to positively influence attitudes toward social responsibility – and desired attributes of respondents' future career, specifically around desires to help others through engineering. Both of these sources also provided strong evidence of validity that the EPRA tool does in fact measure attitudes of social responsibility in engineering students as defined by the eight dimensions of the PSRDM. In addition to validity evidence, reliability evidence was also explored through the use of the Ordinal Alpha, supporting strong reliability as well. All of these results are described in more detail, including the description of sample population and methods in [4].

The use of interviews as validity evidence has been used in other fields such as psychology, health and education work[7,8]. In most of these studies, however, cognitive interviews were used which focus specifically on the user's interpretation of survey language and the user's thought processes in selecting their answers. This technique is used to identify unintended or mistaken interpretations, not in alignment with the survey purpose. One of the interview methods used in this study took this same approach, but, instead of focusing on language interpretation, it used survey questions to foster deeper conversations about social responsibility. The other two methods (described later) took semi-structured approaches, more similar to traditional qualitative interviews. So, while surveys have been used for validity evidence with respect to respondent interpretation, the authors could not find examples in engineering education where interviews were used to explore the underlying dispositions that the survey is intended to measure.

The purpose of this paper is to extend the collection of validity evidence for the EPRA tool by comparing survey responses to coded interviews from 24 engineering students. Like the comparison of Likert-item scores with volunteer activities and career attributes done previously, this examination will provide evidence based on relations to other variables. Because the interviews also focused on attitudes toward social responsibility, this provides convergent evidence.

EPRA Intended Uses and Interpretations

Before presenting evidence of validity for the EPRA tool, it is important to first be explicit about the intended uses of the EPRA tool and the intended interpretations of the data that is derived from the EPRA tool. The EPRA tool is intended as a measure of attitudes toward personal and professional social responsibility, as defined by the PSRDM, in engineering students. The core of the EPRA tool, the 50 Likert-items, focus on student attitudes at the time in which the student is taking the survey, not asking about previous views or future views. The intended uses for this tool are to assess students' current attitudes with respect to social responsibility and, possibly, to assess changes in student views surrounding a specific course or intervention, or over the duration of their college career. It is believed that attitudes toward social responsibility develop and change slowly, so the intention is that this tool would be more successful at assessing changes over longer periods of time[9].

Results from this tool may be interpreted in different ways. In a pre-post design, the results may support or dissuade the use of a pedagogical intervention, such as service-learning, to positively affect student views of social responsibility. This tool may also be used at a programmatic level to assess global changes in student views of social responsibility as an element of professional responsibility toward accreditation. In all of these cases, the results are interpreted through analysis of the Likert-item data, either through means, medians, distributions, or other statistical methods.

Samples and Settings Used in Validation

For the validity evidence explored in this paper, interviews with engineering students were conducted in the spring 2012 semester. The researchers approached professors from different departments at a large public university, asking them to recommend students to talk to who would represent, in their opinion, a wide range of beliefs and experiences regarding the development of social responsibility. Using these recommendations, 33 students were emailed and asked to participate in an interview; 25 students agreed. Before each interview, students read and signed an informed consent form, consistent with IRB protocols, and were asked to take the EPRA survey. After completing the survey, the interview commenced, and would last another 20 to 45 minutes. Interviews were recorded and later transcribed for analysis. No incentives were provided to the students in exchange for their participation in the interview.

Three different interview methods were used to elicit conversation from these students. The formats were changed as the researcher moved through the interviews in attempts to find approaches that would create a more comfortable atmosphere for students to open up and explore deeper beliefs about social responsibility. The first method was a semi-structured format where students were asked questions regarding why they chose engineering as a major, their ideal future career, how they saw engineering contributing to society, their views on pro bono work, their definitions of social responsibility and life experiences that had influenced their views of social responsibility. Eleven students participated in this style of interview.

The second method used student responses to the EPRA survey questions as a guide for conversation, similar to cognitive interviews. Students were asked to lead the researcher through their survey responses, describing their general views regarding certain questions, what examples or experiences influenced their responses, or to explain why they selected a given response over others. This format was chosen because it provided a guide to help stimulate conversation in a more tangible way than the semi-structured method; this seemed to help the engineering students to open up and share more. Eight students participated in this style of interview.

The final interview method used a variation on Rappaport timelines[10] to guide conversation. Students were given a piece of paper with three lines drawn on it. The top line represented a time continuum leading up to their coming to college, the second line represented the beginning of college to the present, and the third line represented the future. Students were asked to write down at least three events on each line which influenced their choosing engineering, their view of engineering, and what they hoped to do as an engineer, respectively. The events that students wrote on their timelines then became the focus of conversation, specifically around influential events or people in their lives that helped shape their views of social responsibility. Through this autobiographical approach, students would generally end up defining their views along the way. Six students were interviewed using this method.

<u>Participants</u> – Twenty five students from Civil (13), Environmental (4), Mechanical (7) and Aerospace (1) Engineering were interviewed. Because the recruitment method relied on recommendations from professors, all of the students were upperclassmen (One junior, 14 senior, and 14 graduate students) and were generally very active students in curricular and/or extracurricular activities. This was expected because the students who faculty would most likely know would be the most active or outgoing students. Ten of the interviewees were women and 15 were men.

<u>Analysis</u> –To relate the interviews to the survey, a rubric was developed to help identify both when a person was talking about a certain dimension of the PSRDM and different degrees of each dimension. A preliminary rubric was created by the research team, then given to a panel of seven experts in engineering and engineering education and discussed in a focus group. Feedback from the focus group was used to develop the final rubric. Definitions used for each of the eight dimensions were based on the PSDRM framework[3]. The full rubrics used for each dimension are given in the Appendix, though a sample for the *Analyze* dimension is provided in Table 2. From early versions of the survey it became clear that item responses were generally skewed towards more positive answers. Therefore the rubric was also developed with a skew such that a "2" from the interviews correlated with a 'neutral' response (4) on the 7-point Likert scale of the survey. All three levels of disagreement were captured within a "1" for the rubric and degrees 3, 4, and 5 corresponded to a 5, 6, or 7 on the Likert-items of the survey. The "NE" (or null) designation was used if there was no evidence for a given dimension in that interview.

Table 2. Partial Rubric for Interview Scoring – *Ability* Dimension

Degrees	Corresponding Likert Score	*Ability*
NE	-	No Evidence
1	1-3	Speaks about an **inability** to help others in a meaningful way, or are averse to helping others.
2	4	Impersonal or distance acknowledgement of one's ability to help, speaks about possibly 'yes', possibly 'no' that they have the ability to help, or speaks about not ever having opportunities to help.
3	5	Expresses that they have the ability to help others, but limiting themselves to small acts, helping individuals more than causes.
4	6	Expresses a strong belief in their ability to help people on systemic levels. They may also tend to recognize how the complexity of these systems may limit their ability to help.
5	7	Superman – they surely have an ability to help on any front – perhaps a seemingly naïve perspective that they can do anything.

Two reviewers used the definitions and rubric to independently code 24 interviews using deductive coding techniques, identifying evidence of each of the eight dimensions and assigning degrees of alignment to each. One interview was omitted because the participant spoke directly about what score she gave on which questions (a traditional cognitive interview), making it difficult for the reviewers to remain objective in assigning degrees for each dimension. Then, based upon evidence from the coding, each participant was given a rating for each dimension. From the independent review, 80 of the 192 items (42%) matched perfectly between the two reviewers, and another 56 (29%) were within one degree, with an item being one dimension for one interviewee. There were 38 instances (20%) where one reviewer saw evidence of a given dimension and the other did not. The two reviewers discussed each item where there was disagreement, examining the evidence in order to come to consensus on an appropriate degree. After consensus was reached there remained 31 items (16%) where there was no evidence.

The degrees determined through consensus were converted to the equivalent Likert-score and compared to median scores for each dimension and interviewee from their survey. Spearman's rank correlation coefficients and Wilcoxon signed rank test values were used to examine correlation and difference, respectively, between the dimension averages from the EPRA Likert-items and the interview degrees as evidence of validity. IBM's SPSS software was used to make these comparisons.

Results

Results used for the examination of validity, including the EPRA median scores and the interview degrees for each interviewee, are given in Table 3. Using radial plots, differences and similarities between the survey scores and interview ratings can be qualitatively examined, as seen in Figure 1. The radial axis represents the 7-point Likert-item scale used in the survey and the recoded degrees from the interviews. Students are represented by "INT" and the order in which they were interviewed, such as INT10, INT13, and INT14. Visually, INT10 and INT13 seem to have good agreement between the EPRA survey responses and the recoded rubric scores from the interviews, whereas the survey and interview results for INT14 are clearly different. Both INT10 and INT13 had an item where there was no evidence to support a rating from the interviews, *ability* and *base skills*, respectively.

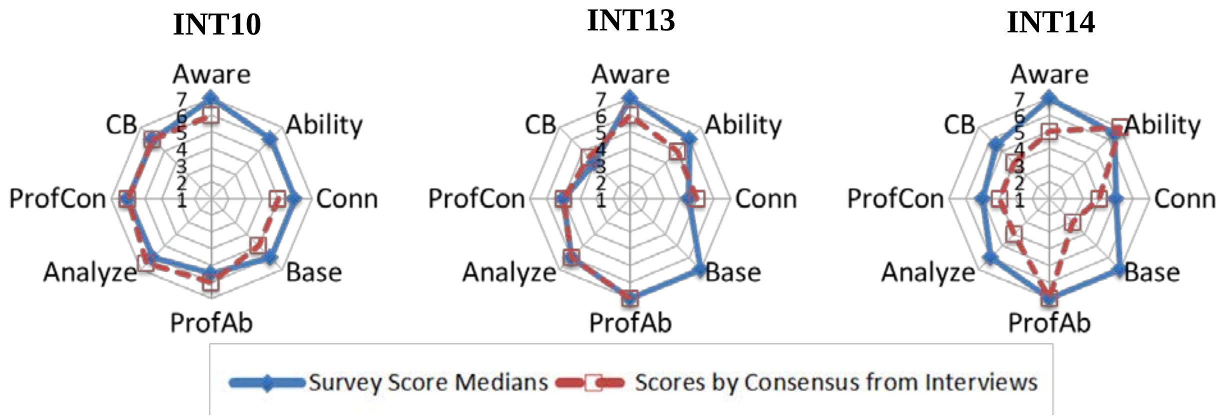

Figure 1. Sample Interviewee Degrees of Social Responsibility and Survey Dimension Median Scores

Quantitatively, the correlation and difference between these two methods were examined using the Spearman's rank correlation coefficient and the Wilcoxon signed-rank test, respectively. The data were examined from two perspectives, one focusing on the overall correlation for each individual, similar to the plots shown in Figure 1, and the other looking at each of the eight dimensions using all 24 interviewees. Spearman's rho, p-values and Wilcoxon p-values for each interviewee are given in Table 3. Examining the p-values for the Spearman test showed that four of the interviewees had statistically significant ($p<0.05$) correlation between the interview and EPRA construct values. Two more individuals had suggestive correlation ($p<0.10$). Examining the rho-squared valued helps to inform the importance of the correlation, and, of the six interviewees with significant or suggestive correlation, three were 'very important' ($R^2>0.74$ for n=8) and three had fair to low importance ($R^2>0.25$)[11]. One of those interviewees with a 'very important' correlation was INT13, supporting what was observed qualitatively using the radar plots. INT10, however, had no correlation per Spearman's rank correlation test because the relationship between the two methods switches, where one is higher for one dimension, but lower for the next, and then back. This 'crossing over' leads to low Spearman values.

The Wilcoxon test is used to support a hypothesis of difference. Eighteen samples rejected the hypothesis of difference ($p>0.05$), including five of the six interviewees with correlation from the Spearman test. Therefore, five of the 24 data pairs provided supportive evidence of both correlation and a lack of difference. It is worth noting that the qualitative examination of INT14 from the radar plots was supported by the statistical methods shown here, where the Spearman test showed no correlation and the Wilcoxon test showed difference.

Table 3. Interviewee EPRA median values, interview ratings, and correlation and difference measures

ID	Method	Dimension Medians and Interview Degrees per Interviewee								Spearman		Wilcoxon
		Aware	*Ability*	*Conn*	*Base*	*ProfAb*	*Analyze*	*ProfCon*	*CB*	R^2	ρ	ρ
INT01	EPRA	6	5.5	5.5	5.9	6	6	5	5.5	0.171	0.309	1
	Interview	6	6	7	4	6	5	6.5	5			
INT02	EPRA	6	5	5	6.0	6	5	4.5	5.5	0.240	0.264	0.705
	Interview	7	5	4	NE	6	5	5	4			
INT03	EPRA	7	4.5	6	6	7	5	5	4.5	0.947	0.005	0.180
	Interview	7	5	6	NE	7	NE	6	NE			
INT04	EPRA	6	6	6	6	5.5	6	6	6	0	1	0.581
	Interview	6	5	6	7	6	7	5.5	NE			
INT05	EPRA	7	6	5.5	4	6	5	6	6	0.194	0.383	0.276
	Interview	6	6	6	NE	5	5	6	NE			
INT06	EPRA	7	6	6	5.7	6	6	5	5	0.031	0.675	0.131
	Interview	5	5	5	5	7	5	5	5			
INT07	EPRA	7	7	6	7	7	7	6	7	0.466	0.091	0.317
	Interview	6	6	7	NE	6	5	7	7			
INT08	EPRA	6	4.5	4.5	5.6	6.5	5	5	6	0.045	0.650	0.786
	Interview	6	NE	6	6	6	7	5	4			
INT09	EPRA	6	6	5.5	5.6	6	5	6	6.5	0.045	0.648	0.244
	Interview	5	NE	5	7	5	5	5	5			
INT10	EPRA	7	6	6	6	5.5	6	6	6	0	1	0.214
	Interview	6	NE	5	5	6	6.5	6	6			
INT11	EPRA	6	6	6	7	6	7	6	6.5	0.178	0.345	0.357
	Interview	7	6	6	NE	6	5	5	6			
INT13	EPRA	7	6	4.5	7	7	6	5	4	0.803	0.006	0.458
	Interview	6	5	5	NE	7	6	5	4.5			
INT14	EPRA	7	6.5	5	7	7	6	5	5.5	0.100	0.446	0.027
	Interview	5	7	4	3	7	4	4	4			
INT15	EPRA	7	4	5	6	6.5	6	4	4.5	0.650	0.029	0.141
	Interview	6	4	3	6	7	5	4	NE			
INT16	EPRA	7	7	6	6	7	6	6	6.5	0.774	0.021	0.034
	Interview	7	NE	5	5	6	4.5	5	NE			
INT17	EPRA	7	6	6.5	6	7	7	6	7	0.331	0.136	0.041
	Interview	7	5	6	6	6	5	4	7			
INT18	EPRA	6	5	5	7	6.5	5	4	5.5	0.132	0.548	0.063
	Interview	NE	4	4	NE	4	NE	4	3			
INT19	EPRA	6	6	6	6	7	6	5	6	0.250	0.253	0.020
	Interview	5	5	4	5	6	NE	5	5			
INT20	EPRA	7	7	7	7	5.5	7	7	6	0.009	0.837	0.045
	Interview	7	6	5	NE	6	5	6	5			
INT21	EPRA	6	4	4	7	6	5	4	4	0.033	0.726	1
	Interview	6	5	5	NE	4	5	4	NE			
INT22	EPRA	6	6	5.5	7	7	5	5	5	0.100	0.488	0.168
	Interview	5	5	6	5.5	5	5	6	NE			
INT23	EPRA	6	6	5	6	6	6	5	5	0.500	0.293	0.414
	Interview	NE	4.5	6	NE	5	NE	5	NE			

ID	Method	Dimension Medians and Interview Degrees per Interviewee								Spearman		Wilcoxon
		Aware	*Ability*	*Conn*	*Base*	*ProfAb*	*Analyze*	*ProfCon*	*CB*	R^2	ρ	ρ
INT24	EPRA	6	4	3	7	6.5	6	3.5	5	0.410	0.088	0.443
	Interview	5.5	4	4	6	5	5	5	3			
INT25	EPRA	6	6	6	6	6.5	6	6	6.5	0	1	0.102
	Interview	6	NE	5	NE	6	NE	6	5.5			

Shaded cells denote importance and statistical significance from Spearman and rejection of hypothesis of difference from Wilcoxon.

Correlation and difference for each dimension was also examined, similar to above. The relevant values are given in Table 4, as well as construct averages from the EPRA tool across all interviewees. Three of the eight dimensions showed statistically significant ($p<0.05$) correlation from Spearman, two of which were also 'very important' ($R^2>0.406$ for n=24). Three dimensions rejected the hypothesis of difference based upon the Wilcoxon test ($p>0.05$). No single dimension met all three requirements.

Table 4. Correlation statistics between interview and survey data

Dimension	Spearman		Wilcoxon	EPRA Scores		
	R^2	p	p	Averages	Range of Individual's Average	Standard Deviations
Awareness (n=22)	0.053	0.304	0.017	6.22	5.2-7.0	0.55
Ability (n=19)	0.430	0.002	0.013	5.54	3.75-7.0	0.88
Connectedness (n=24)	0.102	0.127	0.330	5.40	3.25-7.0	0.82
Base Skills (n=13)	0.044	0.491	0.088	6.30	4.67-7.0	0.50
Professional Ability (n=24)	0.067	0.224	0.007	6.17	4.50-7.0	0.62
Analyze (n=19)	0.014	0.624	0.033	5.87	4.8-6.8	0.57
Professional Connectedness (n=24)	0.233	0.017	0.885	5.10	3.27-6.23	0.76
Costs/Benefits (n=16)	0.543	0.001	0.003	5.59	4.0-6.75	0.74

Shaded cells denote importance and statistical significance from Spearman and rejection of hypothesis of difference from Wilcoxon.

Examining the dimensions with poor Spearman p-values and difference based upon the Wilcoxon test, these dimensions also tended to have higher average scores and lower standard deviations across all interviewees. Perhaps the saturation and narrow distribution on these dimensions, specifically *Awareness* and *Base Skills*, influenced the disparity between the interview ratings and survey responses. Additionally, the interviews were not conducted with this purpose in mind, and therefore the conversations were not focused directly on these base level perspectives. Most of the evidence for both of these dimensions came from peripheral comments, or from examples that the interviewees used in relation to some other topic. Few of the interviewees spoke directly about their awareness that others needed help. Exemplifying this disparity between the interview focus and some of the survey dimensions, there were 11 interviews where no evidence for *Base Skills* was seen by the reviewers. On the other hand, all 24 interviews had some evidence to support a rating for *Professional Connectedness*. Perhaps more focused interviews, with questions directed at perceptions of each dimension, would produce data that would fill in these gaps and provide stronger evidence of agreement across the dimensions.

Conclusions and Study Limitations

This study set out to use the analysis of in-depth interviews with engineering students regarding their attitudes toward social responsibility as evidence of validity for a survey instrument, the Engineering Professional Responsibility Assessment. The strength of qualitative methods, like interviews, is that they can more accurately capture many nuances and subtleties that are overly simplified in Likert-type items. This is particularly important when looking at complex beliefs, such as attitudes toward social responsibility. As evidence of validity, comparing responses from interviews related to the eight dimensions of the PSRDM could provide further confidence that scores from the Likert-items are assessing the dimensions they were intended to measure.

The results presented here showed that there was significant and important correlation and a lack of difference for five students (21% of those interviewed) across all dimensions, providing evidence that perceived degrees of social responsibility aligned well between the interviews and the Likert-items from EPRA for those individuals. This analysis also showed that three of the eight dimensions had significant correlation across all interviewees, and three rejected the hypothesis of difference. In majority, however, there was not strong alignment between coded interviews and survey results. This lack of alignment can most likely be attributed to the focus of the interviews and interview questions not originally being directed at obtaining validity evidence for the eight dimensions of the PSRDM. If this technique were to be used in future work as a source of validity evidence, it would be best to design the interviews ahead of time to address the specific constructs of the instrument – lending itself more to the semi-structured interview approach as opposed to the cognitive interviews or timeline exercises.

In addition to the focus of the interviews not being directly related to the survey dimensions, this study was also limited by the student population that was interviewed. All students were from the same institution and were found through recommendations from faculty. This likely produced a skewed interview population toward students who were active in some form (through research, as graders, or extracurricular activities) such that faculty members thought to recommend them. This sample could not be considered representative of the larger engineering student population. Future attempts to use this method for validity evidence should try to use a more representative sample to test the full range of perspectives and experiences.

Using the deductive coding of qualitative data as a source of validity evidence for the development of a survey instrument is a potentially powerful tool. This approach combines the strength of qualitative approaches to bolster the confidence in a quantitative tool, which is easier to use for large data collection. Future implementations of this method would likely be more successful using interviews that were more directed at the specific constructs measured by the instrument and would be strengthened by using a more representative sample population. That being said, these results still provided strong evidence of validity for the EPRA tool from some of the interview cases examined.

Acknowledgements

This material is based on work supported by the National Science Foundation under Grant #1158863. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Bibliography

1 ABET, "Criteria for Accrediting Engineering Programs Effective for Evaluation During the 2015-2016 Accreditation Cycle," ABET Engineering Accreditation Commission, Baltimore, MD, 2014.

2 L. J. Shuman, M. Besterfield-Sacre and J. McGourty, "The ABET "Professional Skills" - Can They Be Taught? Can They Be Assessed?," *Journal of Engineering Education*, pp. 41-55, 2005.

3 N. Canney and A. Bielefeldt, "A Framework for the Development of Social Responsibility in Engineers," *International Journal of Engineering Education*, vol. 31, no. 5, pp. 5-11, 2015.

4 N. Canney and A. Bielefeldt, "Validity and reliability evidence for the engineering professional responsibility assessment tool to measure social responsibility in students," *Journal of Engineering Education*, 2016 (in press).

5 N. Canney, Assessing Engineering Students' Understanding of Personal and Professional Social Responsibility, Ph.D. Dissertation. University of Colorado, 2013.

6 American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, *Standards for Educational and Psychological Testing*, Washington, DC: American Educational Research Association, 2014.

7 L. M. Desimone and K. C. Le Floch, "Are We Asking the Right Questions? Using Cognitive Interviews to Improve Surveys in Education Research," *Educational Evolution and Policy Analysis*, vol. 26, no. 1, pp. 1-22, 2004.

8 J. A. Ouimet, J. C. Bunnage, R. M. Carini, G. D. Kuh and J. Kennedy, "Using Focus Groups, Expert Advice, and Cognitive Interviews to Establish the Validity of a College Student Survey," *Research in Higher Education*, vol. 45, no. 3, pp. 233-250, 2004.

9 A. R. Bielefeldt and N. E. Canney, "Changes in the Social Responsibility Attitudes of Engineering Students over Time," *Science and Engineering Ethics*, pp. 1-14, 2015.

10 H. Rappaport, K. Enrich and A. Wilson, "Relation between Ego Identity and Temporal Perspective," *Journal of Personality and Social Psychology*, vol. 48, no. 6, pp. 1609-1620, 1985.

11 S. Siegel and N. J. Castellan, No*nparametric Statistics for the Behavioral Sciences*, *2nd Ed.*, New York, NY: McGraw-Hill, Inc., 1988.

Appendix

Table 5. Rubric used for interview analysis describing varying degrees for each of the eight dimensions of the PSRDM[5]

	NE	1	2	3	4	5
Awareness	No Evidence	Use language that distances themselves from those in need, or recognizes some groups that may need help, but denies others (i.e. "Maybe there are people in Africa that need help, but certainly not here in the US.")	Express both positive and negative statements about people needing help. Seem to be waffling on the issue (i.e. "Maybe there are people who need help, but maybe there aren't")	No direct comments about people in need, but peripherally discusses issues of people in need. (i.e. they volunteer at a soup kitchen, but never talk about the needs of the people they volunteer for)	Gives specific examples of people or groups that need help, but speaks about social issues as isolated events. Does not speak about social issues as interconnected. (i.e. "there are hungry people in many parts of the world" but no discussion of causes)	Fully aware that people need help and speaks to the interconnection between social issues and how that affects people in need. Evidence of complexity include: discussing systemic roots of problems, or cross disciplinary issues.
Ability	No Evidence	Speaks about an inability to help others in a meaningful way, or are averse to helping others.	Impersonal or distance acknowledgement of one's ability to help, speak about possibly 'yes', possibly 'no' that they have the ability to help, or speak about not ever having opportunities to help.	Expresses that they have the ability to help others, but limiting themselves to small acts, helping individuals more than causes.	Expresses a strong belief in their ability to help people on systemic levels. They may also tend to recognize how the complexity of these systems may limit their ability to help.	Superman – they surely have an ability to help on any front – perhaps a seemingly naïve perspective that they can do anything.
Connectedness	No Evidence	Speaks about not feeling any moral/ethical responsibility or obligation to help others.	Impersonal, indirect, or vague acknowledgment that "people in general", or companies or countries should help others.	An acknowledgement that they, personally, should help, but with little or no further discussions about why. Motivations of past experiences are surficial, such as ability to travel, or simply 'fun'.	Talk about how they should help others, explaining why they feel that way either because of external or internal motivations (i.e. church, privilege, wealth, ability, morals, etc.).	Talks a sense of personal obligation to help others. It's more than a 'should' but a 'will' or a personal mission. They talk about taking action as a critical component.

	NE	1	2	3	4	5
Basic Skills	No Evidence	Speaks against professional skills (communication, cultural awareness, teamwork, etc.) as being important	No direct discussion of the importance of professional skills, but speaks peripherally about how these types of skills might be important or useful for an engineer.	Acknowledges the need for both technical and professional skills, but does not go into depth, nor give examples of why. Also, does not talk about degrees of importance.	Talks about the importance of a balanced range of skills for an engineer, including technical and professional skills, and/or gives examples of using professional skills in engineering applications.	Talks about how professional skills are central to the work of an engineer and also gives examples or previous experiences to support this.
Professional Ability	No Evidence	Talks about how engineering cannot help those who are in need	Talks about how engineering helps generally just by the projects that we do. Uses examples of roads and buildings (i.e. the status quo) as ways in which engineering helps	Talks more about the potential for engineering to help solve social/ environmental problems that face society. This is a step beyond "just doing what engineers do", making a mild connection between engineering projects and improving people's livelihood.	Talks about engineering as a crucial element towards finding solutions to social problems. May express that engineering could be highly effective in solving these problems, but recognizes that engineering may not be the entire solution.	Hands down, engineering is the central source of solutions for social problems and that human (social, political, personal) development has been possible because of engineers.
Analyze	No Evidence	Talks about how it is not important to consider any social elements of engineering design, but that an engineer only needs to focus on technical issues	Does not speak directly about how social elements should be tied into the engineering process, but they may peripherally talk about how social issues may sometimes play a part (i.e. a narrow view of project stakeholder to include boss and client.)	Agrees on case-by-case basis that social considerations are important in the engineering process, though not primary to the engineering design process (i.e. if the project causes pollution, it will affect the whole community). Includes a recognition of a broader group impacted by engineering, but does not include a wider group in the decision making process	Talks about how it is good to incorporate social elements, and gives examples of how social considerations were positive for successful projects. Hold a wider definition of stakeholders to include the community or potentially affected groups, and includes them in aspects of the decision making process. These stakeholders have input.	Talks about how it is critical to incorporate social issues into the engineering design process, throughout the entire process and that projects cannot be successful without it. Prioritizes consideration of social issues over technical issues in terms of importance to project success.

	NE	1	2	3	4	5
Professional Connectedness	No Evidence	Speaks against any ideas of responsibility or obligation in the engineering profession to help others	Talks solely about professional minimal expectations as sources or levels of responsibility, including ethics, public safety, and cost (i.e. “to avoid lawsuits, you should follow the engineering code of ethics.”)	Talks about the responsibility of an engineer extending beyond professional minimums, perhaps open to ideas of pro bono work, service, sustainability, or environmentalism	Talks about how engineers should do service, but are not required to. There is also a belief that it should be incorporated into their professional career, not just something that they individually do on the side – supported by the profession.	Expresses a strong connection between their personal moral obligation and having a professional responsibility to help others. They express their identity as an engineer being tied to service, more than just the profession in general.
Costs/Benefits	No Evidence	The costs of doing service seem to dominate the conversation, with few references to any benefits. Service is not worthwhile because the costs outweigh the benefits.	Equal discussion of costs and benefits, but no leaning one way or another. Conversation of costs/benefits is dominated by generalizations.	Acknowledges both costs and benefits, but emphasizes the benefits that are gained through doing service. Emphasis resides in vague or shallow examples of benefits, such as “it was fun”, “it makes me feel good”, “got to travel”, or “met new people.”	Talks positively about the benefits of doing service and draws from personal experiences or examples of how engaging in service has benefitted them and their personal development. Examples of benefits would be “opened my eyes”, expansion of cultural understanding	In spite of acknowledgements of the costs of doing service, they are willing to make personal or professional sacrifices to do engineering service long-term and with regularity. Benefits are worth the acknowledged costs.