AC 2010-576: GRADUATE TEACHING ASSISTANTS' ASSESSMENT OF STUDENTS' PROBLEM FORMULATION WITHIN MODEL-ELICITING ACTIVITIES

Amani Salim, Purdue University

Amani Salim is a Post-Doctoral Researcher in the School of Engineering Education at Purdue University. She received her B.Sc. and M.Sc. in Electrical Engineering from University of Minnesota Twin Cities, and her Ph.D. in BioMEMS and Microelectronics from Weldon School of Biomedical Engineering at Purdue University. Her research focuses on problem formulation within Model-Eliciting-Activities (MEAs) with realistic engineering context.

Heidi Diefes-Dux, Purdue University

Heidi A. Diefes-Dux is an Associate Professor in the School of Engineering Education at Purdue University. She received her B.S. and M.S. in Food Science from Cornell University and her Ph.D. in Food Process Engineering from the Department of Agricultural and Biological Engineering at Purdue University. Since 1999, she has been a faculty member within the First-Year Engineering Program at Purdue, the gateway for all first-year students entering the College of Engineering. She coordinated (2000-2006, 2010) and continues to teach in the required first-year engineering problem solving and computer tools course, which engages students in open-ended problem solving and design. Her research focuses on the development, implementation, and assessment of model-eliciting activities with realistic engineering contexts. She is currently the Director of Teacher Professional Development for the Institute for P-12 Engineering Research and Learning (INSPIRE).

Graduate Teaching Assistants' Assessment of Students' Problem Formulation within Model-Eliciting Activities

Abstract

Model-Eliciting-Activities (MEAs) are open ended engineering problems set in realistic contexts, which requires teams of students to create a mathematical model for solving a client's problem. At the beginning of each MEA, students are required to answer three questions: Q1) "Who is the client?", Q2) "In one of two sentences, what does the client need?" and Q3) "Describe at least two issues that need to be considered when developing a solution for the client". These questions are designed to guide students' problem formulation. As graduate teaching assistants (GTAs) are responsible for assessing these student responses, it is anticipated that GTAs contribute to students' ability to formulate problems. However, a cursory review of GTAs assessment of student work indicated that some GTAs struggle to properly assess students' responses. To guide future GTA professional development with MEAs, and problem formulation in particular, this paper seeks to explore these questions in more detail: "How are the GTAs' assessing students' responses to the MEA individual questions?" and "Do students' ability to answer these questions improve across the three MEAs implemented in a single semester?" Three distinct MEAs were implemented in a required first-year engineering problem-solving course in Fall 2007. Open coding and content analysis of ~500 (out of ~1500) student responses per MEA was performed to establish an expert assessment of the student responses. The expert assessment of the student responses were used to evaluate the GTAs assessments. Results verify that the GTAs' assessments of student responses were very weak. It is clear that GTA professional development with problem formulation is needed. Recommendations for such professional development are put forth.

I. Introduction

Problem formulation often occurs at the early stages of a problem solving process when it has the greatest potential for affecting the direction and success of all succeeding stages[1-3]. If the problem formulation step is not successful, the result is often an incomplete, over-simplified and unclear problem solution, or even worse, an incorrect problem solution[4-5]. An important emphasis in problem formulation is the resolution of the open-ended engineering problem in its context with its multiple dimensions[6-9]. By nature, open-ended problems are ambiguous, and they potentially have multiple solutions. Open-ended engineering problems require understanding clients and their needs. In the business world, clients do not come for services with infinite resources at their disposal; they operate within boundaries, which includes financial or design constraints[10]. This requires an ability to recognize the relevant context and content of a problem. Problem formulation within open-ended problems is also an iterative process, where it requires accepting priori constraints and utilizing creative and multi-disciplinary thinking[9]. As engineering students' learn how to formulate problems, assessment strategies should accommodate these aspects of problem formulation.

Good assessment practice has a number of features. First, assessment should be aligned with institution and course objectives/goals[10,11]. Problem formulation is a relevant aspect of engineering practice. For ABET accreditation, it is a required outcome: "Engineering programs must demonstrate that their students attain the following outcomes: (e) an ability to identify, formulate, and solve engineering problems"[12]. As such, it is anticipated that engineering courses with a problem solving aspect will have course objectives centered on an ability to effectively formulate problems. In a first-year engineering course, one could envision assessment in the context of an open-ended design problem, which requires knowledge of the client, knowledge of the client's needs, and identification of constraints[10]. So the development of assessment strategies around problem formulation is aligned with the needs of engineering and student learning within engineering.

Second, good assessment practice should be a resource for understanding student learning[11]. To understand students' problem formulation strategies, assessment should provide answers that confirm and challenge educators' and instructors' assumptions about their students' abilities[13]. The assessment strategy should provide a window on students' understandings so that constructive feedback can be provided and instructional strategies can be adjusted as necessary. That is, assessment can influence the learning outcomes and instruction strategy used in a course, where assessments conducted at the beginning of a course or topic could change an instructor's initial assumptions of students' learning. Based on early assessment findings, an instructor can, for example, provide constructive feedback that better addresses deficiencies in students' learning[13]. Fostering problem formulation abilities in first-year engineering students can be challenging because students have had little to no experience with problem formulation. This may be their first time trying to sort out the context and content of a problem. Appropriate assessments are needed to understand their entry point into a problem, where as stated earlier, incorrect problem formulation leads to deficiencies in subsequent problem solving activities.

Finally, one of the hallmarks of good assessment includes ongoing efforts to evaluate and improve the assessment process itself[11]. Assessment, like problem formulation itself, should be ongoing and iterative[11]. Assessment should be conducted with knowledge of student learning and teaching, and is should be updated to address the needs of instructors and students. Assessment should be used in preparing instructors to implement the assessment itself and to effectively use the outcome of assessment[13]. This means rethinking of the assessment strategy and professional development of instructors are both part of a good assessment strategy.

Model-Eliciting Activities (MEAs) are a special class of open-ended problems set in realistic engineering contexts, where students have the ability to engage in problem formulation[14]. The design of robust assessment strategies for MEAs has been a challenge in a large (N ~ 1500) first-year engineering course where students engage in individual and team-work efforts to produce written solutions that result in products with a (sometimes high) degree of variation[15]. To assess and further students' abilities to formulate problems within MEAs, and within open-ended problems more generally, there is a need for high quality assessment strategies. A good assessment strategy would not only correct students' interpretation of their initial reading of the MEA problem text, but also could clarify concepts and direct successful framing of the problem within the context.

In this first-year engineering course, Graduate Teaching Assistants (GTAs) are responsible for providing assessment of both students' individual and team work on the MEAs. This study focuses on GTAs' assessment of students' individual responses to a series of individual questions that launch each MEA. These questions engage students in individual problem formulation, requiring students write out their initial interpretation of the content and context of the problem statement before they engage in team problem solving[16]. Posing short open-ended questions at the beginning of the problem solving process is a strategy often used in mathematics to encourage reasoning and thinking[17].

The GTAs assess students' responses to these individual questions using a Likert point scale[14]. The GTAs need to be able to apply this scale in a meaningful way as their assessment can influence the success and direction of the teams' subsequent problem solving stages, the learning outcomes for MEAs, and the successful achievement of MEAs objectives - such as the development of students' problem formulation abilities. A cursory look at GTAs' assessment of students' responses seemed to indicate that GTAs may not be consistently applying the scale and this may be impacting students' work and learning. In this study we seek to explore these questions in more detail: "How are the GTAs' assessing students' responses to the MEA individual questions?" and "Do students' ability to answer these questions improve across the three MEAs implemented in a single semester?" Answers to these questions may indicate a need for refinement in the scale and changes in the GTA professional development. The point here is not to test the reliability of a rubric but rather to understand how the GTAs are assessing students' responses to the individual questions so that better assessment strategies can be developed, where assessment strategies encompasses (1) professional development with problem formulation, (2) professional development with assessment, and (3) assessment rubrics and associated training materials.

II. MEAs Descriptions

Model-Eliciting Activities (MEAs) are client-driven, open-ended problems that require students to use one or more mathematical or engineering concepts that are unspecified by the problem, make sense of new knowledge and understandings delivered through the text of a given MEA problem statement, and create a generalizable (share-able, re-usable, modifiable) procedure that can be used by the client to solve the given problem and similar problems. MEAs are based on six principles as outlined by Lesh, et al.[18] that have been modified to the engineering context[14, 19]. MEAs are "model-eliciting" and "thoughts-revealing". Of particular interest here is the window provided by MEAs on: (1) students' initial understandings and conceptualization during the problem formulation stage and (2) GTAs' understandings of problem formulation within the context of MEAs as seen through their assessment of students' problem formulations using a simple rubric.

III. Methods

A. Setting & MEA Implementation

The setting for this study was the Fall 2007 offering of a required first-year engineering problem solving and computer tools course with an enrollment of approximately 1500 students. This course was a combination of introduction to engineering problem solving, computational tools, and teamwork as described by Diefes-Dux and Imbrie[15]. Course meetings included two 50-

minute faculty-led lectures and one 110-minute graduate teaching assistants (GTAs) led lab period (N~28) per week.

In Fall of 2007, three MEAs were implemented: *Theft Prevention with Laser Detection System, Just-in-Time Manufacturing,* and *Nano Roughness*. The *Theft Prevention with Laser Detection System* MEA is set in the context of establishing a laser security system to prevent the theft of artwork in a display room[14]. The *Just-in-Time Manufacturing* MEA is set in the context of ranking shipping companies for Devon Dalton Technologies according to their ability to deliver materials on time between two company divisions[14]. The *Nano Roughness* MEA is set in the context of manufacturing surface coatings for biomedical implants[14]. For each MEA, student teams are required to construct a mathematical model in the form of a written procedure intended for the use by the client in the same or similar situations. At the start of each MEA, students are asked to individually respond to three questions: Q1 - "Who is the client – the direct user of your procedure?", Q2 - "In one or two sentences, what does the client need?", and Q3 - "Describe at least two issues that need to be considered when developing a solution for the client." These questions focus on the needs of the client and function to guide students to articulate the problem in their own words.

To assess students' responses to each individual question, the GTAs applied a 3-point Likert scale. The scale was used to indicate the level of correctness of the response. For this scale, 0 corresponded to "No" or incorrect, 1 corresponded to "Sort of" - meaning not-far-off or some elements of the response are correct, and 2 corresponded to "Yes" – correct or most of the elements of the response are correct.

B. Participants & GTA Professional Development

Eleven (11) experienced and six (6) inexperienced GTAs were employed in Fall 2007. Experienced GTAs had been assigned a first-year engineering laboratory section and graded nearly all students' work, including students' work on MEAs, in at least one prior semester. Inexperienced GTAs had no prior experience with the first-year engineering course. All GTAs received four hours of professional development (PD) training prior to the start of the Fall 2007 semester. The PD focused on several aspects: connecting engineering practice to teaching, the MEA pedagogy, audience information (first-year engineering students), and practical issues of MEA implementation and assessment[14]. GTAs were trained to understand the open-ended and realistic nature of MEAs, the MEA modeling and design process, and MEAs fit with the goals of the first-year engineering course. The GTAs solved the first MEA they would implement in the Fall 2007 laboratory setting. They read the MEA problem statement and answered the individual questions. After acquiring consensus on the three individual questions, the GTAs worked in teams of three or four to complete a first draft of a solution to the MEA. By answering the individual questions, the GTAs experienced what the students typically encounter during the problem formulation stage of a MEA activity.

In terms of GTAs' PD in assessment of student work, GTAs were first taught about the role of feedback and assessment in teaching with open-ended realistic problems. Much of the assessment portion of the GTAs' PD focused on the grading of the student teams' solutions to the MEA. Assessment of students' responses to the individual questions was only briefly discussed and the discussion focused on the GTAs answers to the individual questions.

C. Data Collection & Analysis

Students answered the three individual questions through a web-based interface connected to a database. The web-based system manages the organization of MEAs, facilitating students' interactions throughout the MEA activities. For this study, ~500 (out of ~1500) students' responses to the individual questions for three MEAs: *Theft Prevention with Laser Detection System, Just-In-Time Manufacturing* and *Nano Roughness* were examined.

Open coding and content analysis of the student responses was performed to establish an expert assessment of the student responses. The expert's assessment of the student responses were then used to evaluate the GTAs' Likert point scale assessments of the student work. For Q2 and Q3, due to the complexity of the student responses, the quality of the student responses had to be assessed by the expert using the coding scheme developed during open coding and content analysis. A 5-point quality level scale was developed for each question so that an evaluation of the GTAs assessment of student work could be made. The scale was developed through discussions between the authors - a doctoral student with background in engineering, experience teaching the class for one semester, and two years experience working with MEAs and a faculty member in engineering education with extensive MEA expertise. The Q2 and Q3 quality scales capture core elements of a students' responses, and often reveal the level of complexity (or lack thereof) of a given MEA. The experts' quality level descriptions are provided in the next section.

Students' responses to Q2 and Q3 could consist of several issues. The quality level of each issue was determined independently, and then an average was taken to determine the quality level of the response as a whole.

IV. Results

A. Q1 – "Who is the client – the direct user of your procedure"

Expert Coding of Student Responses for Q1

For Q1 - "Who is the client?", students are required to identify the direct client. This is the person(s) who will utilize the written procedure or solution developed by the team. Descriptions of the types of clients identified by students and sample student responses are shown in Tables 1-3. For each MEA, the types of clients are divided into four categories: 1) the direct client, 2) the indirect client, 3) non-clients, and 4) multiple clients - a combination of 1-3. For MEA 1 (Table 1), the direct client is the technical installation team of TLP (a security system company) who will install the security laser system and ultimately use the written procedure developed by the team. The indirect clients include the CEO of TLP and TLP itself. The non-clients include Polk University and Polk University's Chancellor who has contracted TLP to install a security system. Other non-clients include clients that are outside the scope of the problem (i.e. Polk University installation team – a made-up entity by the students, a building, or an object). There are three multiple client categories; these are combinations of direct and indirect clients, direct and non-clients, and several different non-clients.

Table 1. Q1 – "Who is the client?" client types found in student responses for MEA 1.

Client Type	Description	Student Example
Direct Client	Technical installation team at TLP Products & Installations (the company)	*"TLP's technical installation team."*
Indirect Client	Werner Miller, CEO of TLP TLP Products & Installations (the company)	*"The client is Werner Miller of TLP Products and Installation. He is the person who was chosen to install the equipment in the art galleries."*
Multiple Clients	A combination of direct client and indirect clients	*"Theft and Loss Prevention (TLP) is the client, but more specifically the Technical Installation Team needs my procedure."*
	A combination of direct client and non-clients	*"The client of the scenario is the museum at Polk University, and the direct user of the procedure is the Product & Installations team."*
	A combination of non-clients	*"The client is the curator, the chancellor, and the Polk University art department."*
Non-Client	Polk University Polk University Chancellor Polk University technical installation team A building, and an object	*"The client is an art gallery or a museum."* *"The store owners that believe they have a problem with theft. The stores, art galleries, museums and any other place that needs to protect their stuff."*

For MEA 2 (Table 2), there is no direct client. The company Devon Dalton Technologies is an indirect client and needs a procedure to rank potential shipping companies in ability to operate in a just-in-time manufacturing fashion. Other indirect clients include the CEO of Devon Dalton Technologies and the non-clients are the company sub-divisions of DDT: Ceramica or Alphalon. The multiple client category combinations are similar to MEA 1.

For MEA 3 (Table 3), direct clients are the scientists at Liguore Laboratories; they need a procedure to quantify the roughness of a material surface coatings at the nanoscale using AFM images. Similar to MEA 1 and MEA 2, the indirect client is the director or CEO of Liguore Laboratories, and the non-clients are person(s) outside the problem scope. In MEA 3, the multiple client categories only consist of two combinations - a combination of direct and indirect clients and a combination of several indirect clients.

Table 2. Q1 – "Who is the client?" client types found in student responses for MEA 2.

Client Type	Description	Student Example
Direct Client	***Not Applicable***	
Indirect Client	Devon Dalton Technologies (the company) CEO of Devon Dalton Technologies	*"The client is D. Dalton Technologies."* *"The client is Devon Dalton, the CEO of Deveon Dalton Technologies."*
Multiple Clients	A combination of indirect clients	*"The client is Devon Dalton and his company, DDT."*
	A combination of indirect client and non-clients	*"D.Dalton Technologies, Ceramics and Alphalon."*
	A combination of non-clients	*"The Alphalon and Ceramica Divisions of D. Dalton Technologies."*
Non-Client	Random clients Workers at Ceramica and Alphalon Ceramica	*"The client is the Applications Engineering Team."* *"The clients are the workers at Ceramica and Alphalon."* *"The client is Ceramica. They are the direct users of the new shipping company that our team will determine."*

Table 3. Q1 – "Who is the client?" client types found in student responses for MEA 3.

Client Types	Description	Student Example
Direct Client	Scientist at Liguore Lab.	*"The client of my team's procedure is the scientists at the Liguore Labs."*
Indirect Client	Kerry Prior at Liguore Lab Liguore Lab	*"Kerry Prior, Vice President of Research, is the client."*
Multiple Clients	A combination of direct client and indirect clients	*"The client would be the scientists who would use this procedure directly. The person who would want this for his business would be Kerry Prior, VP of research."*
	A combination of indirect clients	*"The client in this case is Kerry Prior or Liguore Laboratoies."*
Non-Client	Clients from other MEAs Generic clients	*"The many global manufacturers who use Liguore Laboratories to provide coating techniques for the orthopedic and biomedical implants they make."*

Evaluation of GTA Assessments for Q1

Summaries of the GTAs' and expert's assessments of the students' responses to Q1 are presented in Tables 4-6. The students' responses are divided into the four client types identified in the student responses. Assessments made by experienced and inexperienced GTAs are provided separately.

As can be seen from Table 4, for MEA 1, most student responses were of the indirect client type (54.9% of all responses). Many of these responses were given full credit (2 = "Yes") by the experienced and inexperienced GTAs (48.8% and 34.8%, respectively); though a larger number of these responses were also given no credit by experienced and inexperienced GTAs (41.7% and 43.8%, respectively). There were an equal number of student responses of both the direct and non-client types (12.6% each). For the direct client responses, experienced GTAs gave full credit to only 48.8% of the responses, while inexperienced GTAs gave 72.2% of these responses full credit. For non-clients responses, most experienced and inexperienced GTAs gave no credit (0 = "No") (65.2 % and 53.3%, respectively). Most of the multiple-client responses were of the non-clients type. Experienced GTAs gave 69.7% of these responses full credit, while inexperienced GTAs gave both partial (1 = "sort of") and full credit for this type of response. A much lower number of student responses were of the multiple client type with combinations of direct client and non-clients, and direct client and indirect clients.

In Table 5, for MEA 2, there is no direct client type. The indirect client type has the most response (87.0% from total response). Most of the experienced and inexperienced GTAs gave full credit for the indirect client response (83.2% and 93.1% from the experienced and inexperienced GTAs, respectively). Similar to MEA1, the non-client type has the lowest number of response (2.8% from total response). For the non-client category, 40.0% and 66.7% of experienced and inexperienced GTAs, respectively, gave no credit. In MEA 2, most multiple-client type responses are a combination of indirect clients (5.3% of total response and of 52.1% of total response of the multiple client category), which both experienced and inexperienced GTAs categories mostly gave full credit (100.1% and 66.7% of experienced and inexperienced GTAs, respectively).

For MEA 3 (Table 6), most student responses were of the indirect client type (64.2%). While the experienced GTAs tended to give full credit for this response (84.8%), the inexperienced GTAs did not (18.9%). For the direct-client type, experienced GTAs gave full credit (90.0%) and inexperienced GTAs gave partial credit (50.0%) and full credit (41.7%). For the non-clients type, 58.1% and 60.0% of experienced and inexperienced GTAs, respectively, gave full credit. Only 8.3% of the student responses were of the multiple client type. Both experienced and inexperienced GTAs were inclined to give full credit for such responses at least 60% of the time.

B. Q2 – "What does the client need"

Expert Coding of Student Responses for Q2

For Q2 - "In one or two sentences, what does the client need?", students are required to identify the needs of the client that must be met by the team. The quality level assigned to a students' response by the experts is based on the inclusion of the following elements in the students' response: the deliverable, the deliverable description, the function of the deliverable, the function description, and the constraints. The deliverable is the thing that the client wants the student team to develop. One of two codes is typically assigned for the deliverable; students identify either a written procedure or a physical system as the deliverable.

Table 4. GTAs' assessments of students' responses to Q1 – "Who is the client?" for MEA 1.

Client Types		MEA 1 Student Responses (N = 483)	Experienced GTAs (N=11)				In-Experienced GTAs (N=6)			
			Graded Student Reponses (R=321)	0	1	2	Graded Student Reponses (R=162)	0	1	2
Direct Client		61 (12.6%)	43	27.9%	23.3%	48.8%	18	27.8%	0.0%	72.2%
Indirect Client		265 (54.9%)	176	41.7%	10.3%	48.0%	89	43.8%	21.4%	34.8%
Multiple Clients	**Non-Clients**	74 (15.3%)	43	25.6%	4.7%	69.7%	31	22.6%	38.7%	38.7%
	Direct client and non-client	15 (3.1%)	10	54.5%	27.3%	18.2%	5	25.0%	25.0%	50.0%
	Direct client and indirect client	7 (1.4%)	2	50.0%	0.0%	50.0%	5	60.0%	20.0%	20.0%
Non-Clients		61 (12.6%)	47	65.2%	3.3%	31.5%	14	53.3%	33.4%	13.3%

Table 5. GTAs' assessments of students' responses to Q1 – "Who is the client?" for MEA 2.

Client Types		MEA 2 Student Responses (N = 468)	Experienced GTAs (N=11)				In-Experienced GTAs (N=6)			
			Graded Student Responses (R=315)	0	1	2	Graded Student Responses (R=153)	0	1	2
Direct Client		NA	NA	NA	NA	NA	NA	NA	NA	NA
Indirect Client		407 (87.0%)	277	12.8%	4%	83.2%	130	3%	3.8%	93.1%
Multiple Clients	**Non-clients**	13 (2.8%)	11	72.7%	0.0%	27.3%	2	0.0%	50.0%	50.0%
	Indirect client and non-client	10 (2.1%)	7	14.3%	0.0%	85.7%	3	0.0%	0.0%	100.0%
	Indirect clients	25 (5.3%)	10	0.0%	0.0%	100.0%	15	0.0%	33.3%	66.7%
Non-Client		13 (2.8%)	10	40.0%	20.0%	40.0%	3	66.7%	33.3%	0.0%

Table 6. GTAs' assessments of students' responses to Q1 – "Who is the client?" for MEA 3.

Client Types		MEA 3 Student Responses (N = 445)	Experienced GTAs (N=11)				In-Experienced GTAs (N=6)			
			Graded Student Responses (R=295)	0	1	2	Graded Student Responses (R=150)	0	1	2
Direct Client		52 (10.8%)	40	2.5%	7.5%	90.0%	12	8.3%	50.0%	41.7%
Indirect Client		310 (64.2%)	204	6.4%	8.8%	84.8%	106	62.2%	18.9%	18.9%
Multiple Clients	**Direct client and non-client**	20 (4.5%)	10	0.0%	20.0%	80.0%	10	20.0%	20.0%	60.0%
	Direct client and indirect client	17 (3.8%)	10	10.0%	10.0%	80.0%	7	14.3%	14.3%	71.4%
Non-Client		46 (9.5%)	31	25.8%	16.2%	58.1%	15	33.3%	6.7%	60.0%

The student responses then receive codes for the other elements in their response that provide details about the deliverable. Codes are associated with the deliverable descriptions (i.e. the characteristic of the deliverable), the function (what the deliverable is intended to do), the function description (i.e. how the function behaves or acts), and the constraints (the given artifacts needed to develop the deliverable). Tables 7-9 describe the quality levels that apply to student responses for MEAs 1, 2, and 3.

For MEA 1 (Table 7), responses focusing on a written procedure are assessed at quality levels 3-5, with higher quality levels being assigned to responses with more detail and thus more assigned codes. Responses at quality level 2 focus on a "system" deliverable (where the "system" refers to the laser detection hardware), or responses that indicate no deliverable but have the function code "to secure the art", both with other associated codes. A quality level 1 corresponds to responses with the deliverable "system" with no other codes, or the function code "to secure art" with no other codes.

For MEA 2 (Table 8), responses focusing on a written procedure are assessed at quality levels 3-5. Responses focusing on a written procedure with the specific function to "rank shipping company" with other codes were given a quality level 5, while responses with the function to "choose shipping company", to "analyze shipping data", and to "analyze shipping company", with other codes, were given a quality level of 4. Quality level 3 responses also focus on the "procedure" deliverable but with no specific function, but with at least one code. Quality level 2 responses focus on quantitative results as the deliverable - so the deliverable codes were "analysis of shipping company", "analysis of shipping data", or "ranking of shipping company" with at least one other coded segment. Quality level 1 responses includes the physical object deliverables, in this case the "shipping company", with or without other codes. Quality 1 responses also includes "professional opinions and advice", deliverables outside the problem scope, or responses that are simply a reiteration of the problem.

For MEA 3 (Table 9), responses focusing on a written procedure are assessed at quality levels 3-5. Quality level 5 consists of responses where the deliverable is a "procedure" with the function "to quantify roughness", along with a combination of two or more of other codes. Quality level 4 responses state that the deliverable is a "procedure" with the function "to quantify roughness" with a combination of at least one other codes. Quality level 3 responses state that the deliverable is a written procedure to "determine smoothness/roughness", to "measure smoothness/roughness", or with no function to quantify roughness along with at least one other coded segment. Quality levels 2 and 1 consist of responses referencing the deliverable as a "material". Level 2 responses state that the deliverable is a "material" with a code for function description or deliverable description. Level 1 responses also focus on a "material" deliverable but with no other codes; level 1 responses also include deliverables that are outside the scope of the problem.

Evaluation of GTA Assessments for Q2

Summaries of the GTAs' and experts' assessments of the students' responses to Q2 are presented in Tables 10-12. The students' responses are sorted into the five quality levels. Assessments made by experienced and inexperienced GTAs are again provided separately.

Table 7. Q2 – "In one or two sentences, what does the client need?" quality level of student responses for MEA 1.

Quality Level	Description	Student Example
5 – Very Good	Responses include: (1) Deliverable = "procedure" . (2) Deliverable Function = "to find stop positions" (3) At least two coded segments associated with the deliverable description, the function description, or the constraints.	*"The client needs a* ***procedure*** [deliverable] *detailing how* ***to find specific stop positions*** [function] *for a laser transmitter* ***given its location in the room and the other locations of the laser receiver devices*** [constraints]. *This procedure must also be* ***re-useable, share-able, and adaptable*** [deliverable description] *to other situations."*
4	Responses include: (1) Deliverable = "procedure". (2) Deliverable Function = "to install the security system" or "to secure art" (3) At least one coded segments associated with the deliverable description, the function description or the constraints.	"*A* ***procedure*** [deliverable] ***to set up a laser theft detection system*** [function] *that is* ***subtle (to not detract from the art)*** [function description] *for the art."*
3	Responses include: (1) Deliverable = "procedure". (2) At least one coded segment associated with the deliverable description, the function, the function description, or the constraints. Responses that are outside of the problem scope but pertain to a procedure.	*"The client needs a* ***procedure*** [deliverable] *to* ***protect the valuable art from theft*** [function]*."*
2	Responses include: (1) Deliverable = "system". (2) At least one coded segment associated with the deliverable description, the function, the function description, or the constraints. Responses with the function "to secure art" and at least one coded segment associated with the function, description or the constraints	*"The client needs a* ***surveillance system*** [deliverable] *that is* ***re-usable and share-able*** **[**deliverable description], *that will* ***effectively*** [deliverable description] ***moniter*** [function] *the priceless pieces of art work in the room."* *"My client needs* ***efficiently affective*** [deliverable description] ***installation*** [function] *of Theft and Loss Prevention system in one of the exhibition rooms. So that, it* ***does not detract from the art*** [function description] *and* ***secure the art*** [function] *objects as well."*
1 – Very Poor	Responses include the deliverable "system" with other coded segments. Responses with the function "to secure art" with other coded segments.	*"The client wishes to safely* ***secure valuable art pieces*** [function] *in a gallery room, and* ***prevent them from being damaged or stolen*** [function] *in any way."*

Table 8. Q2 – "In one or two sentences, what does the client need?" quality level of student responses for MEA 2.

Quality Level	Description	Student Example
5- Very Good	Responses include: (1) Deliverable = "procedure". (2) Deliverable Function to "rank shipping company". (3) At least one coded segment associated with the deliverable description, the function, the function description, or the constraints.	*"A* ***procedure*** [deliverable] ***to rank potential shipping companies*** [function]*."* *"The client needs a* ***procedure*** [deliverable] ***to rank potential shipping companies*** [function] ***to meet the shipping needs of DDT*** [function description]*."*
4	Responses include: (1) Deliverable = "procedure". (2)The function to "choose shipping company", "analyze shipping data", and "analyze shipping company". (3) At least one coded segment associated with the deliverable description, the function, the function description, or the constraints.	*"They need a* ***procedure*** [deliverable] ***to find*** [function] ***which shipping company is best suited to meet their ever growing need for on time shipments*** [function description]*."*
3	Responses include: (1) Deliverable = "procedure". (2) At least one coded segment associated with the deliverable description, the function, the function description, or the constraints.	*"A* ***procedure*** [deliverable] ***for choosing a shipping company*** [function]*."*
2	Responses include: (1) Deliverable = "analysis of shipping company", "analysis of shipping data", or "ranking of shipping company". (2) At least one coded segment associated with the deliverable description, the function, the function description, or the constraints.	*"The client needs* ***a ranking of the shipping companies*** [deliverable] *so that they can* ***decide*** [function] ***which one is best for them to use*** [function description]. *This ranking should include, among other things,* ***the number of minutes that the shipments are late*** [constraints]. *One of their biggest goals is* ***to minimize the late time of their shipments*** [function description]."
1– Very Poor	Responses include: (1) Deliverable ="shipping company". (2) At least one coded segment associated with the deliverable description, the function, function descriptions, or the constraints. Help to choose/find company with other codes. Professional opinions or advice with other codes. Deliverables outside the problem scope with other codes.	*"The company needs a* ***reliable*** [deliverable description] ***shipping company*** [deliverable] *so that they can keep their production schedule and* ***deliver the final product*** [function] *to their clients* ***on time*** [function description]*."* *"The* ***client needs help to choose the best company*** [deliverable]*."* *"Client wants* ***to find the best shipping company*** [deliverable] *for D. Dalton Technologies."* *"The client needs a* ***professional analytic opinion*** [deliverable] *as* ***to find a suitable shipping company that will not delay in shipping products*** [function description]*."*

Table 9. Q2 – "In one or two sentences, what does the client need?" quality level of student responses for MEA 3.

Quality Level	Description	Student's Example
5 – Very Good	Responses include: (1) Deliverable = "procedure". (2) Deliverable Function = "to quantify roughness". (3) At least two coded segments associated with the deliverable descriptions, the function descriptions, or the constraints.	*"They need the* ***procedure*** [deliverable] *that will* ***quantify the roughness*** [function] *of the nanosale* ***diamond coatings*** [function description] *produced in the lab, and the* ***procedure*** *should be able to quickly* ***apply to the three AFM gold images****.*[function description]*"* *"The Client needs a* ***feasible*** [deliverable description – not coded], ***reusable****,* ***and simple*** [deliverable description] ***procedure*** [deliverable] *that will help them* ***quantify the smoothness*** [function] *of the* ***diamond coating*** [function description]*."*
4	Responses include: (1) Deliverable = "procedure". (2) Deliverable Function = "to quantify roughness". (3) At least one coded segment associated with the deliverable description, the function, the function description, or the constraints.	*"The client needs a* ***procedure*** [deliverable] *to* ***quantifying roughness*** [function] *of* ***nanoscale*** **material** [function description]. *"A* ***procedure*** [deliverable] *for* ***quantifying roughness*** [function] *using* ***AFM images*** [function description]*."*
3	Responses include: (1) Deliverable = "procedure". (2) Deliverable Function = to "determine smoothness/roughness", to "measure smoothness/roughness", or no function. (3) At least one coded segments associated with the function, the function description, or the constraints.	*"They need a* ***quick easy*** [deliverable description] ***procedure*** [deliverable] ***to use to determine weather their smooth*** [function] ***diamond experiments are working or not***[function description]*"* *"The research team needs a* ***procedure*** [deliverable] *using* ***AFM images*** [function description] ***that will let them know whether or not the production process of the gold coatings is working*** [function]. *"The client needs a* ***procedure*** [deliverable] *by which they can* ***measure the roughness*** [function] *of* ***the diamonds and the melted gold*** [function description] ***using the AFM graphs*** [function description]*."*
2	Responses include: (1) The deliverable "material" (2) Any of the coded segments associated with of the function description and the deliverable description.	*"The client needs us to create* ***medical device coatings*** [deliverable] *that provide better* ***performance and durability*** [deliverable description] *in medical implants."*
1 - Very Poor	Responses include deliverable "materials" with no other coded segments. Deliverables outside the problem scope.	*The client needs an* ***exhibit*** [deliverable] *that correctly shows to middle schoolers what engineering is in an interactive and fun way.* *The client needs* ***biomedical supplies with special surface coatings*** [deliverable].

When multiple deliverables are included in a student response, the quality level assigned was an average assessment. For each MEA, approximately 6% of all students provided multiple deliverable types, with all of these responses receiving a quality level 3 or lower assessment. For example, see this student response for MEA 1 Q2:

> *"The client needs a* ***laser security system*** [deliverable] ***to protect 5 rare collections*** [function] *in glass cases during an art show* (level 2)*. The client also needs the* ***procedure*** [deliverable] *to be created so that way it may be understood and* ***modified for future use*** [deliverable description] (level 3)*."* Average level = 2.5 = level 3.

For MEA 1 (Table 10), most students' responses were of poor quality, with 41.4% assessed at quality level 2. Only 48 responses (9.9%) were considered very good (level 5). Both experienced and inexperienced GTAs often gave full credit for very good responses (80.0% and 66.7 %, respectively). For quality level 2-4 responses, both experienced and inexperienced GTAs gave no, partial and full credit anywhere from 21.9% to 44.4% of the time.

For MEA 2 (Table 11), many students provided responses that were of a level 5 quality (34.4%), and more than 85% of these responses received full credit from both experienced and inexperienced GTAs (93.0% and 87.5%, respectively). Quality level 4 responses received partial and full credit from the experienced GTAs (50.6% and 47.2%) and full credit from the inexperienced GTAs (71.1%). Level 3 quality responses received partial credit and full credit in equal measure (42.3%) from experienced GTAs; again the inexperienced GTAs tended to assign full credit (66.6%). Level 2 quality responses (15.4%) received partial credit at least 50% of the time from both experienced and inexperienced GTAs (57.5% and 45.8%, respectively). Level 1 quality responses mainly received partial credit from experienced and inexperienced GTAs (43.6% and 48.6% respectively).

In MEA 3 (Table 12), most students' responses were of quality level 4 followed by quality level 5 quality (63.1% and 18.0% respectively). The vast majority of the GTAs gave full credit for level 4 and level 5 responses. Experienced GTAs provide no credit for level 1 response (82.8%) while inexperienced GTAs tended to give full credit (87.9%).

C. Q3 – "Describe at least two issues…"

Expert Coding of Student Responses for Q3

For Q3 - "Describe at least two issues that need to be considered when developing a solution for the client"; students are asked to provide issues that are important and relevant for creating the deliverable that the client wants. This question guides students to think about the relevant constraints and how they might impact their solution. Descriptions of the quality levels assigned by the experts and sample student responses which correspond to each level are provided in Tables 13-15. Issues that pertain to the mathematical model (procedure) and its development were considered to be within the scope of the problem and were therefore good responses. For MEA 2 (Table 14), there are no quality levels 2 and 4. For MEA 3 (Table 15), there is no quality level 4. The presence (or absence) of a quality level is a result of the complexity (or lack of complexity) of the problem statement. The existence of quality level 2 and 4 responses in Table

Table 10. Level Assigned by Expert versus GTAs to Q2 – "In one or two sentences, what does the client need?" for MEA 1.

Quality Level	MEA 1 Student Responses (N = 483)	Experienced GTAs (N = 11)				Inexperienced GTAs (N = 6)			
		Graded Student Reponses (R=321)	0	1	2	Graded Student Reponses (R=162)	0	1	2
5 – Very Good	48 (9.9%)	32	20.0%	0.0%	80.0%	16	0.0%	33.3%	66.7%
4	14 (2.9%)	5	40.6%	21.9%	37.5%	9	43.8%	31.3%	25.0%
3	98 (20.4%)	74	27.1%	28.4%	44.4%	24	35.3%	23.5%	41.2%
2	200 (41.4%)	138	32.6%	31.2%	36.2%	62	33.9%	40.3%	25.8%
1– VeryPoor	123 (25.4%)	72	27.8%	19.4%	52.8%	51	11.8%	35.3%	52.9%

Table 11. Level Assigned by Expert versus GTAs to Q2 – "In one or two sentences, what does the client need?" for MEA 2.

Quality Level	MEA 2 Student Responses (N = 468)	Experienced GTAs (N = 11)				Inexperienced GTAs (N = 6)			
		Graded Student Reponses (R=315)	0	1	2	Graded Student Reponses (R=153)	0	1	2
5– Very Good	161 (34.4%)	113	0.0%	7.1%	93.0%	48	0.0%	12.5%	87.5%
4	127 (27.2%	89	2.2% %	50.6%	47.2%	38	5.3%	23.7%	71.1%
3	32 (6.8%)	26	15.4%4	42.3%	42.3%	6	16.7%	16.7%	66.6%
2	72 (15.4%)	48	18.7%	56.3%	25.0%	24	29.2%	45.8%	24.9%
1– VeryPoor	76 (16.2%)	39	38.5%	43.6%	17.9%	37	24.3%	48.6%	27.1%

Table 12. Level Assigned by Expert versus GTAs to Q2 – "In one or two sentences, what does the client need?" for MEA 3.

Quality Level	MEA 3 Student Responses (N = 445)	Experienced GTAs (N = 11)				Inexperienced GTAs (N = 6)			
		Graded Student Reponses (R=295)	0	1	2	Graded Student Reponses (R=150)	0	1	2
5 – Very Good	80 (18.0%)	49	0.0%	0.0%	100.%	31	0.0%	0.0%	100.0%
4	281 (63.1%)	197	2.0%	9.6%	90.6%	84	17.9%	10.7%	71.4%
3	33 (7.4%)	19	0.0%	10.5%	89.5%	14	28.6%	0.0%	71.4%
2	7 (1.6%)	2	100.0%	0.0%	0.0%	5	60.0%	0.0%	40.0%
1 – Very Poor	44 (9.9%)	28	82.8%	6.9%	10.3%	16	9.1%	3.0%	87.9%

17 and a level 4 responses in Table18 are due averaging of the quality levels assigned to individual issues as described earlier in the “Data Collection & Analysis” section.

For MEA 1 (Table 13), issues pertaining to the procedure, or elements that are essential for procedure development (i.e. transmitter placements and angles of laser beams), are considered quality level 5 responses. Issues required in procedure development but are considered less important (because they had no information on these issues) for procedure development were considered level 4 responses. These issues include the height of objects, ranges that are detectable by lasers and obstacles between laser and objects. Quality level 3 corresponds to procedure issues that are outside the scope of the problem (i.e. solutions produced need to be fool proof), restatement of the problem, systems issues within the scope of the problem, or room requirements. Quality level 2 corresponds to issues that pertain to the system that are outside the problem scope, and quality level 1 issues are just physical issues of a security system, cost issues, or various randomness.

Table 13. Q3 – “Describe at least two issues that need to be considered when developing a solution for the client?” quality level of student responses for MEA 1.

Quality Level	Issues Description	Student’s Example
5-Very Good	• Procedure requirements • Mirror placements • Laser beam angles	*“Two issues that need to be addressed are transmitter placement and the angles of the laser beams.”*
4	• Height of objects to be detected • Range capability of the laser transmitter/receiver system • Obstacles between the transmitter and receiver (or objects)	*“Also, height will be a problem. These receivers are only mounted in one place and we have to ensure that the laser is reaching the receiver. If either the receiver or the laser is at the wrong height then the system will not work. Thus, we must have accuracy and attempt to keep all of the beams in the same plane.”*
3	Procedure requirements that are outside the scope of the problem. • Restatement of the problem • Laser detection system requirements within scope of problem • Room requirements System issues that are within the problem scope.	*“Secondly, the solution* [previously noted as procedure] *that we come up with should be fool proof and should meet all the conditions that have been set foorth by the Chancellor of Polk University.”* *“Also, procedure should effectively influence the outcome without creating other problems to solve.”* *“One issue is that the laser needs to quickly and very accurately rotate repeatedly to shoot a beam at every receiver that is placed on top of a case containing art.”*
2	• Laser detection system requirements not within the scope of the problem	*“What are the limits as to how strong the laser is when pointed at glass? This may affect the effectiveness of how strong the laser is, potentially weakening the signal and allowing a possible theft.”*
1- Very Poor	• Physical security issues • Cost of system issues	*“Size and obviousness of the security system is the first issue that needs to be dealt with because it is a client requirement.”* *“1 the solution should not be too costly. “*

For MEA 2 (Table 14), issues pertaining to the procedure, characteristics of procedure, and time issues are considered quality level 5. Similar to MEA 1, restatement of the problem to be solved is considered quality level 3, such as "reasons for shipments to be late" or "ways to analyze data". Issues pertaining to cost, safety and random issues are considered quality level 1.

Table 15 shows the quality levels for MEA 3. Issues pertaining to the procedure and characteristics of the procedure are considered quality level 5. Similar to MEA 2 and MEA 1, issues that are a restatement of the problem to be solved, and the quality of AFM images are considered quality level 3. Quality level 2 consists of issues pertaining to the characteristics of the material in AFM images, and quality level 1 issues pertain to the characteristic of materials not applicable to AFM images and random issues that are outside the problem scope, which includes cost and applications of coating material to implants.

Table 14. Q3 – "Describe at least two issues that need to be considered when developing a solution for the client?" quality level of student responses for MEA 2.

Quality Level	**Issues Description**	**Student's Example**
5- Very Good	• Procedure requirements (i.e. address tie breaks, provides ranking in results) • Characteristics of procedure • Time	*"1.include our team's reasoning for the each step 2.to develop a list of other factors besides timing 3. include our team's procedure and the results in rank order generated by applying out procedure to the sampling of data."* *"The procedure must be simple enough to use. Another issue to be considered is the need to make sure that there aren't any ties between companies."* *"Average delay for each company -max delay for each company -variation in arrival times -consistency in arrival times."*
4		
3	• Reasons for late shipments, Problems in rank development • Ways to analyze the data • Restatement of the problem	*"1. How long has the data been collected? How long a data is needed to conclude a potential shipping company? (eg. 2 months or 2 weeks)? 2. How about reasons for being late?"*
2		
1- Very Poor	• Cost • Safety • Random issues which are not in the scope of the problem	*"1) One of the issues to be considered is the safety record of the shipping companies, since the cargo is highly fragile material. 2) The other issue to be taken into consideration is the comparitive cost effectiveness of the companies, and the overall efficiency of its processes as a clue of its ability to deliver."*

Table 15. Q3 – "Describe at least two issues that need to be considered when developing a solution for the client?" quality level of student responses for MEA 3.

Quality Level	Issues Description	Student's Example
5- Very Good	• Procedure requirements (i.e. units in results, sampling methods, scale of images) • Characteristics of procedure	"*the scale of the images....*" *"Another thing to consider is units. There are many different units and they should all be kept similar."*
4		
3	• AFM images • Definition of roughness • Restatement of the problem	*"What are the quailties of AFM images? How are the three type of AFM image utilizied?."*
2	• Characteristics of material used for AFM images	*"One issue is the overall height of the smooth diamond. One could be smooth overall but could have one or two rough parts."*
1- Very Poor	• Material issues and random issues which are not in the scope of the problem	*"What will be a durable yet affordable coating? How will the coating be applied?"*

Evaluation of GTA Assessments for Q3

Summaries of the GTAs' and experts' assessments of the students' responses to Q3 are presented in Tables 16-18. Similar to Q2, the students' responses are sorted into the five quality levels, and the assessments made by experience and inexperienced GTAs are provided separately.

Table 16 shows the student responses sorted by quality level and the GTAs' assessment for Q3, MEA 1. Most of the students' responses were of quality level 3 or 5 (31.1% and 30.4%, respectively). Again, the quality level is an average assessment of the issues in the student response. For example, see this student response for MEA 1 Q3:

> "*The first issue needed to adress is* ***the strength of the laser beams*** [system issue-level 2]. *IE. Will the be able to* ***keep track of up to 5 pieces of art work*** [system issue- level 4]*? The* ***laser beams should be fast and effective transitioning from piece to piece so making sure the device can transmit across the room and adjust from piece to piece is crutial*** [system issue – level 3]. *The second issue would* ***be the angles needed to bounce the laser off the wall to reach the transmitters*** [angle issue – level 5]. ***The laser beam needs to be accuarate*** [system issue- level 3] *when transitioning to prevent false alarms and bringing unwanted disturbances to the party."* Average: =3.4 = level 3.

Regardless of quality level, experienced GTAs tended to give full credit for student responses. Inexperienced GTAs gave a mix of partial and full credit across the quality levels, but increasingly gave partial credit as the quality level decreases.

For MEA 2 (Table 17), students' responses were somewhat split between quality level 3 (34.4%), followed by level 1 and 5 quality (25.9and 23.5%, respectively). An example response for MEA 2 is as follows:

"Two issues that need to be considered when developing a solution for the client are the ***number of minutes late (average and range)*** [time – level 5] *and* ***the condition*** *of whatever is being shipped* [random issues – level 1]*."*Average: 3 = level 3

Again, the experienced GTAs tended to give full credit regardless of the quality of the response. Now too the inexperienced GTAs tended to give full credit regardless of the quality of the response, with one exception. Inexperienced GTAs divided their grades between partial and full credit for level 2 quality responses (42.1% and 57.9%, respectively).

Table 18 shows the evaluation of the student responses for MEA3. Most students' responses were of quality level 4 and 3 (38.2% and 31.1%, respectively). An example quality level 4 response for MEA 3 is as follows:

*"**the size of the images*** [size of image – level 5] *given &* ***how precisely the roughness can be measured*** [restatement of the problem – level 3]*"* Average: 4 = level 4

Within this quality level range, both experienced and inexperienced GTAs tended to give full credit for the responses. Few responses were of quality level 5 or 1 (3.8% and 7.6%, respectively). Both experienced and inexperienced GTAs gave no credit for very poor quality (level 1) responses. Experienced GTAs gave no or partial credit for very good quality (level 5) responses, and inexperienced GTAs gave partial credit. .

V. Discussion

In Fall of 2007, students were provided with three individual questions at the start of each MEA. These questions were designed to ensure that students read the MEA text and were familiar enough with the problem to engage in teamwork activities and to develop the requested written procedure. The questions required students sort out from the written MEA problem statement what is within and not within the scope of the problem their team had been asked to solve. In other words, they had to formulate the problem. The success of the problem formulation activities is highly dependent on the quality of assessment. Assessment here was provided by GTAs. In order to understand the GTAs assessment of the student work, an expert assessed a subset of the student responses and evaluated the GTAs assessments of that student work. For each individual MEA question, the research questions are addressed below.

A. Q1 – Who is the client – the direct user of your procedure?

It should be noted that during the analysis of MEA 2, it became apparent that there was no clearly identifiable direct client (or user of the procedure that the student teams were being asked to create) in the text of the problem. This may have impacted how the GTAs assessed the students' responses on both MEA 2 and MEA 3 and may have impacted students' responses to MEA 3 Q1. This is discussed below.

"How are the GTAs' assessing students' responses to the MEA individual questions?" The GTAs seem somewhat unable to identify who the direct client is in MEA 1 (Table 4). When the students identified the direct client in MEA 1, inexperienced GTAs seemed to more readily recognize this as a correct response than the experienced GTAs (though few students responded

Table 16. Level Assigned by Expert versus GTAs to Q3 – "Describe at least two issues that need to be developed when developing a solution to the client" for MEA 1.

Quality Level	MEA 1 Student Responses (N = 483)	Experienced GTAs (N = 11)				Inexperienced GTAs (N = 6)			
		Graded Student Reponses (R=321)	0	1	2	Graded Student Reponses (R=162)	0	1	2
5 – Very Good	147 (30.4%)	105	10.5%	4.8%	84.7%	42	31.0%	35.7%	33.3%
4	30 (6.2%)	17	5.9%	23.5%	70.6%	13	7.7%	30.8%	61.5%
3	150 (31.1%)	98	2.0%	10.9%	87.1%	52	22.4%	44.9%	32.7%
2	61 (12.6%)	41	4.9%	14.6%	80.5%	20	25.0%	50.0%	25.0%
1 – Very Poor	95 (19.7%)	60	5.0%	16.7%	78.3%	35	11.4%	57.2%	31.4%

Table 17. Level Assigned by Expert versus GTAs to Q3 – "Describe at least two issues that need to be developed when developing a solution to the client" for MEA 2.

Quality Level	MEA 2 Student Responses (N = 468)	Experienced GTAs (N = 11)				Inexperienced GTAs (N = 6)			
		Graded Student Reponses (R=315)	0	1	2	Graded Student Reponses (R=153)	0	1	2
5 – Very Good	110 (23.5%)	100	2.0%	8.2%	89.8%	10	1.6%	18.0%	80.3%
4	39 (8.3%)	22	0.0%	22.7%	77.3%	17	0.0%	11.8%	88.2%
3	161 (34.4%)	96	1.1%	12.4%	86.5%	65	1.4%	22.2%	76.4%
2	37 (7.9%)	18	0.0%	5.6%	94.4%	19	0.0%	42.1%	57.9%
1 – VeryPoor	121 (25.9%)	79	0.0%	7.6%	92.4%	42	7.1%	19.0%	73.8%

*Responses in level 2 and 4 exist due to averages from multiple answers with different levels.

Table 18. Level Assigned by Expert versus GTAs to Q3 – "Describe at least two issues that need to be developed when developing a solution to the client" for MEA 3.

Quality Level	MEA 3 Student Responses (N = 445)	Experienced GTAs (N = 11)				Inexperienced GTAs (N = 6)			
		Graded Student Reponses (R=295)	0	1	2	Graded Student Reponses (R=150)	0	1	2
5 – Very Good	17 (3.8%)	9	44.4%	55.6%	0.0%	8	0.0%	100.0%	0.0%
4	170 (38.2%)	131	6.9%	3.8%	89.3%	60	6.7%	18.3%	75.0%
3	138 (31.1%)	77	14.3%	7.8%	77.9%	40	23.0%	14.7%	62.3%
2	86 (19.3%)	56	7.1%	7.1%	85.8%	30	26.7%	10.0%	63.3%
1 – Very Poor	34 (7.6%)	22	100.0%	0.0%	0.0%	12	91.7.0%	8.3%	0.0%

*Responses in level 4 exist due to averages from multiple answers with different levels.

with the direct user). Both experienced and inexperienced GTAs gave bipolar assessments when students identified indirect clients – so many GTAs believe the CEO or TLP company is the direct user of the procedure. When multiple clients or non-clients were given in the student responses, the GTAs' assessments show no discernable patterns. For MEA 2 (Table 5), there was little variation in the student responses as the complexity of the problem context was much lower. GTAs accepted the indirect client response, alone or mixed with other client types, as correct. By MEA 3 (Table 6), it is difficult to tell whether the experienced GTAs are better at identifying the direct user as a correct response or not because they are accepting almost any response from the student as correct. The inexperienced GTAs seem more clear about the indirect client not being the correct response, but they are not committed to giving students who identify the direct client full credit. GTAs in general do not know how to assess non-client type responses.
We believe the GTAs are having trouble identifying the direct client due terminology issues. They may be associating the word "client" with persons hiring an engineering team, lab or company to do engineer work. This causes confusion between the users of the procedure being produced by the team (who we've called the direct clients) and persons in a supervisory or company role (who we've called indirect clients).

"Do students' ability to answer these questions improve across the three MEAs implemented in a single semester?" Across MEAs 1 to 3 (Table 4-6), students are focused on the indirect client. Again, we believe this is due to the "client" terminology in which students associate the word "client" with persons hiring an engineering team, lab or company to do engineer work. The lack of an identifiable direct client in the MEA 2 problem statement and the subsequent GTA acceptance of all indirect client responses as correct may have enabled that misconception to be held going in to MEA 3. Since Fall 2007, we have moved away from using the word "client" in favor of using the word "user" in conjunction with ideas of user-centered design.

By MEA 3, the students do not respond with multiple clients as frequently. However, despite a decline in non-client identification in MEA 2, more non-clients were identified again in MEA 3. We believe this may be due to students' lack of familiarity with the nanoscale technology and medical research contexts. This may make it more difficult for them to sort out which persons in the problem context will actually use with the procedure they are creating.

B. Q2 – What does the client need?

"How are the GTAs' assessing students' responses to the MEA individual questions?"
For MEA 1 (Table 10), the GTAs seem to know what they are looking for in a good response. When the response is good, they indicate the response is correct. When the response is of intermediate quality, the GTAs provide a mix of assessments. The GTAs also seem unable to decide how to assess a poor response. The GTAs are somewhat better able to discern a good response in MEA 2; better student responses received more credit. Experienced GTAs seem more likely to give partial credit for responses with less detail or detail that is not correct, while the inexperienced GTAs seem satisfied that a procedure has been identified. For MEA 3 (Table 12), the GTAs seem to be able to differentiate high quality response, although most GTAs provide full credit to a level 3 quality response. Inexperienced GTAs are not identifying very poor quality responses (though there were few of these).

"Do students' ability to answer these questions improve across the three MEAs implemented in a single semester?" For MEA 1 (Table 10), the students' mis-identify the deliverable. They do not understand that the client wants them to design a procedure not an object. By MEA 2 (Table 11), many more students' have the idea that they need to create a procedure and can identify its function. The number of students identifying that the deliverable is a procedure is greater in MEA 3 than that seen for MEA 2. Across the three MEAs there was growth in students' ability to identify the deliverable and describe that deliverable. So, there may be some preliminary evidence that the GTAs assessment may be having a positive impact on students' learning. This will need to be further explored. Since Fall 2007, we have spent more time in lecture talking to student about the anatomy of a good response to this question, where the anatomy consists of the deliverable, the criteria for success, and the constraints. It is hoped that this discussion will further accelerate students' ability to identify client's needs.

C. Q3 – Describe at least two issues…

"How are the GTAs' assessing students' responses to the MEA individual questions?"
With the exception of the inexperienced GTAs on MEA 1, across all MEAs, GTAs typically gave full credit for all student responses. This was expected as GTAs were instructed to accept reasonable answers. For MEA 1, inexperienced GTAs did seem to show some evidence of recognizing the presence of irrelevant issues. In general, we believe that GTAs struggle to discern relevant from irrelevant issues. Or they have a tendency to give full credit when only one relevant issue is present.

"Do students' ability to answer these questions improve across the three MEAs implemented in a single semester?" The average quality of students' responses did not change across the MEAs (mean scores were 3.2, 2.9 and 3.1, respectively for MEA 1, 2 and 3). As of MEA 3 (Table 18), there are still a large number of responses of poor (level 1-2) quality (26.9%). Students are not able to differentiate relevant and irrelevant issues.

A quality level of 3 represents responses that are a mix of relevant and irrelevant issues. A quality level 3 is achieved by a student having multiple answers of mixed quality or two answers of extreme high and low quality. For instance, this student response to MEA 1 has multiple issues of varying quality leading to a quality level 3:

> "*The first issue needed to adress is* ***the strength of the laser beams*** [system issue - level 2]. *IE. Will the be able to* ***keep track of up to 5 pieces of art work*** [system issue - level 4]*? The* ***laser beams should be fast and effective transitioning from piece to piece so making sure the device can transmit across the room and adjust from piece to piece is crutial*** [system issue – level 3]. *The second issue would* ***be the angles needed to bounce the laser off the wall to reach the transmitters*** [angle issue - level 5]. ***The laser beam needs to be accuarate*** [system issue - level 3] *when transitioning to prevent false alarms and bringing unwanted disturbances to the party."* Average: =3.4 = level 3.

In this second instance, a student provides a relevant and irrelevant issue to MEA 2:

> *"Two issues that need to be considered when developing a solution for the client are the* ***number of minutes late (average and range)*** [time – level 5] *and* ***the condition*** *of whatever is being shipped* [random issues – level 1]*."* Average: 3 = level 3.

Again, students' tend to provide both relevant and irrelevant issues and GTAs tend to assign a full credit both kinds of issues. This could mean that both students and GTAs are not able to differentiate between issues that are and are not within the scope of the problem.

VI. Conclusions and Future Directions

Results of this work have implications for first-year engineering curriculum and instruction. It is clear that GTA PD is needed around problem formulation. It cannot be assumed that GTAs can formulate problems and assess students' formulations in open-ended problem solving situations, even when the problems are as scaffolded as MEAs. GTA PD needs to include the following elements:

- ***In-depth discussion of the importance of problem formulation skills in engineering problem solving and design***. One reason that the GTAs assessment of student work is somewhat scattered may be the perception of its relative importance. Students' grades on the individual questions do not weigh very much (~ 1% of the MEA grade) as it serves as a check for being in lab, but the impact on their solution could be great. The importance of problem formulation and the impact of the GTAs feedback on the students' success has not been properly conveyed.
- ***Practice with answering the individual questions for every MEA to be implemented in the classroom.*** The GTAs need to have the experience of teasing out this information from the problem so that they can provide more specific feedback to their students.
- ***In-depth discussion of the potential responses to the questions to understand the context of the problem.*** The GTAs also need to be familiar with the types of student responses they will likely encounter so that they can provide appropriate feedback.
- ***Practice assessment of student responses with a detailed rubric.*** Once the GTAs know what kinds of response they can expect, they need to know how these map the 3-point Likert scale. This will enable them to apply the scale in a reliable fashion.
- ***Expert feedback on student work.*** The GTAs need to see how an expert would review student work so that they can gauge the appropriateness of their written feedback.

Such PD has been implemented since Fall 2008. Future research will look at the impact of this new GTA PD on GTA performance and student learning.

Acknowledgements

This work was made possible by a grant from the National Science Foundation (EEC 0717508). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Bibliography

1. Ang, T.S.L., D.R. Arnott, and P.A. O'Donnell. 1994. Problem formulation and decision support systems development. In *Proceedings of the Australian Conference on Information Systems*. Monash University, Melbourne.
2. Mintzberg, H., D. Raisinghani, and A. Theoret. 1976 The structure of 'unstructured' decision processes. *Admin. Sci. Quart.* 21: 246-275.
3. Volkema, R.J. 1983 Problem formulation in planning and design. *Management Science* 29(6): 639-652.
4. Coman, D. 1986. Developing a process model of problem recognition. *Academy of Management Review* 11(14) : 763-776.
5. Jackman, J., S. Ryan, C. Ogilvie, and D. Niederhauser. 2008 Scaffolding to improve reasoning skills in problem formulation. In *Proceedings of the American Society of Engineering Education Annual Conference.* Pittsburgh, PA.
6. Atman, C., K. Yasuhara, R.S. Adams, T.J. Barker, J. Turns, and E. Rhone. Breadth in problem scoping: A comparison of freshman and senior engineering students. 2008. *International Journal of Engineering Education* 42(2): 234-245.
7. Atman, C., D. Kilgore, and A. Morozov. 2007. Breadth in design problem scoping: Using insights from experts to investigate student processes. In *Proceedings of the American Society of Engineering Education Annual Conference.* Honolulu, HI.
8. Neeley, K.A., D. Elzey, D. Bauer, and P. Marshall. 2004. Engineering in context: A multidisciplinary team capstone experience, incorporating real world constraints. In *Proceedings of the American Society for Engineering Education Annual Conference.* Salt Lake City, UT.
9. Elzey, D. 2006. Teaching intro to engineering in context - UVA engineering's new cornerstone. In *Proceedings of the American Society of Engineering Education Annual Conference.* Chicago, IL.
10. Rogers, G. 2005 Assessment: The ultimate open-ended design problems. *Communication Link*. Baltimore, MD.:ABET:14-15.
11. Banta, T.W,(Ed.). 2004. *Hallmarks of Effective Outcomes Assessment: Assessment Update Collections*. San Francisco, CA.: Jossey-Bass.
12. ABET (2009). *Criteria for Accrediting Programs in Engineering*. Baltimore, MD: ABET. http://www.abet.org/forms.shtml#For_Engineering_Programs_Only (last accessed, March 15 2010).
13. Angelo, T.A., and K.P. Cross. 1993. *Classroom Assessment Techniques: A Handbook for College Teachers*. 2nd Edition. San Francisco, CA.: Jossey Bass Higher and Adult Education Series.
14. Zawojewski, J.S., H. Diefes-Dux, and K.J. Bowman, (Eds). 2008. *Models and Modeling in Engineering Education*. Rotterdam, the Netherlands: Sense Publishers.
15. Diefes-Dux, H.A., and P.K. Imbrie. 2008. Modeling Activities in a First-Year Engineering Course. In *Models and modeling in engineering education: Designing experiences for all students, eds.* J.S. Zawojewski, H. Diefes-Dux, and K..J. Bowman, 55-92. Rotterdam, the Netherlands: Sense Publishers.
16. Diefes-Dux, H.A., and P.K. Imbrie. 2008. Model Development Sequence. In *Models and modeling in Engineering Education: Designing experiences for all student, eds.* J.S. Zawojewski, H. Diefes-Dux, and K.J. Bowman, 37-54. Rotterdam, the Netherlands: Sense Publishers.
17. Caroll, W.M. 1999. Using short questions to develop and assess reasoning. In *Developing mathematical reasoning in grades K-12*. Reston, VA.: NCTM.
18. Lesh, R., M. Hoover, B. Hole, A. Kelly, and T. Post. 2000. Principles for developing thought-revealing activities for students and teachers. In *Handbook of Research Design in Mathematics and Science Education eds*. A.E. Kelly and R.A. Lesh, 591-645. Mahwah, New Jersey: Lawrence Erlbaum.
19. Diefes-Dux, H.A., T. Moore, and J. Zawojewski. 2004. A Framework for Posing Open-Ended Engineering Problems: Model Eliciting Activities. In *Proceedings of Frontiers in Education Conference*. Savannah, GA.