Paper ID #26524

The Development of a Coding Scheme Analyzing Formative Assessment in Undergraduate Engineering Science Courses

Dr. Aaron W. Johnson, University of Michigan

Aaron W. Johnson is a postdoctoral research fellow at the University of Michigan. He received his Ph.D. in Aeronautics and Astronautics from the Massachusetts Institute of Technology in 2014, after which he served as a postdoctoral research fellow at the Tufts University Center for Engineering Education and Outreach. Aaron also obtained a master's degree from MIT in 2010 and a bachelor's degree from the University of Michigan in 2008, both in aerospace engineering.

Dr. Jessica E. S. Swenson, University of Michigan

Jessica Swenson is a post doctoral fellow at the University of Michigan. She was awarded her doctorate and masters from Tufts University in mechanical engineering and STEM education respectively. Her current research involves examining different types of homework problems in undergraduate engineering science courses, flexible classroom spaces, active learning, responsive teaching, and elementary school engineering teachers.

Mr. Max William Blackburn, University of Michigan

Max Blackburn is a fifth year undergraduate Electrical Engineering student at the University of Michigan, focusing in Power systems and Energy. He is currently assisting Dr. Cynthia Finelli with research concerning the effects of flexible learning spaces and formative assessment techniques.

Dr. Cynthia J. Finelli, University of Michigan

Dr. Cynthia Finelli is Associate Professor of Electrical Engineering and Computer Science, Associate Professor of Education, and Director and Graduate Chair for Engineering Education Research Programs at University of Michigan (U-M). Dr. Finelli is a fellow in the American Society of Engineering Education, a Deputy Editor of the Journal for Engineering Education, an Associate Editor of the IEEE Transactions on Education, and past chair of the Educational Research and Methods Division of ASEE. She founded the Center for Research on Learning and Teaching in Engineering at U-M in 2003 and served as its Director for 12 years. Prior to joining U-M, Dr. Finelli was the Richard L. Terrell Professor of Excellence in Teaching, founding director of the Center for Excellence in Teaching and Learning, and Associate Professor of Electrical Engineering at Kettering University.

Dr. Finelli's current research interests include student resistance to active learning, faculty adoption of evidence-based teaching practices, the use of technology and innovative pedagogies on student learning and success, and the impact of a flexible classroom space on faculty teaching and student learning. She also led a project to develop a taxonomy for the field of engineering education research, and she was part of a team that studied ethical decision-making in engineering students.

The Development of a Coding Scheme Analyzing Formative Assessment in Undergraduate Engineering Science Courses

Abstract

This research paper addresses responsive teaching, which is a particular form of formative assessment that instructors use to understand and respond to the “disciplinary engagement” of students in real-time during instruction While ideas about what constitutes “disciplinary engagement” are established in science education where responsive teaching has received considerably study, the field of engineering has not yet established a clear idea of what “disciplinary engagement” means—particularly in *engineering science courses,* which we define as the technical courses at the sophomore or junior level that are non-lab and non-design courses. In this paper we make progress toward an understanding of disciplinary engagement—and therefore responsive teaching—in engineering science courses by focusing on the content that an instructor elicits, notices, and responds to during in-class formative assessment. We first discuss the development and final version of a coding scheme that classifies the content that an instructor elicits, notices, and responds to during question-initiated dialogue. The main focus of this coding scheme is a set of four codes that describe different content being invited by a question or offered as a statement: 1) a number of definition, 2) mathematical reasoning or process, and scientific or engineering reasoning related to either 3) a specific example being discussed in class or 4) the course content outside of a specific example. We conclude the paper by presenting an example of how coded question-initiated dialogue can be analyzed to investigate the type of content being discussed by the instructor and the students. Analyzing two different engineering science courses, we find that one course featured more students and instructor questions addressing science and engineering reasoning, which maps to our expectations of the course based on our observations of them.

Introduction

Responsive teaching is a particular form of formative assessment that instructors use to understand and respond to the “disciplinary engagement” of students in real-time during instruction [1]. While the specific definition of disciplinary engagement varies by subject, the process of responsive teaching is constant. In responsive teaching, instructors 1) *elicit* student thinking, 2) *notice* aspects of students’ disciplinary engagement in their responses, and 3) *respond* in real-time to support students’ disciplinary engagement [2], [3]. There is a specific focus on students’ *ideas* rather than “content,” which Coffey et al. define as “a body of correct information, centered on terminology, and selected in advance as lesson objectives” [1].
Prior empirical studies of responsive teaching have shown that the practice enhances students’ conceptual understanding [4], [5] and supports students’ engagement in disciplinary practices [1], [6], [7]. And while most prior studies of responsive teaching have focused on pre-college

mathematics or science education, we anticipate that responsive teaching will also benefit students in undergraduate engineering education.

In science education, where responsive teaching has received considerably study, “disciplinary engagement” in the practice of science involves arguing about concepts and disciplinary ideas using evidence [8]. As such, one example of responsive teaching in science is asking students to hypothesize explanations for natural phenomena they have observed, and then to design, conduct, and interpret the results of experiments that test these explanations [9]. However, the field of engineering has not yet established a clear idea of what “disciplinary engagement” means.

Engineering at its core is about creating solutions to problems using mathematics, science, and creativity through a design process. The engineering curriculum reflects this by containing different types of courses that teach the mathematical models of natural phenomena (i.e. engineering science courses, or technical core courses), laboratory and experimental techniques and processes (i.e. lab courses), and fundamentals of engineering design (i.e. design courses). These courses all ask students to engage disciplinarily in different ways, all in support of the overall practice of engineering to create new solutions. Prior research has offered definitions of disciplinary engagement in design by proposing frameworks (e.g. [10]) or by outlining design practices that are typical of experts (e.g. [11], [12]). However, the field has not yet established a clear idea of what “disciplinary engagement” looks like in *engineering science courses*, which we define as technical courses at the sophomore or junior level that are non-lab and non-design courses. These courses make up a significant portion of students’ engineering education, but they have been studied less in engineering education research than design courses [13].

One conceptualization of disciplinary engagement in engineering science courses is encouraging students to make sense of the relationship between “canonical” mathematical models that engineers apply to physical scenarios in service of analysis or design (the “engineering”) and the underlying physical phenomena (the “science”). This conceptualization takes inspiration from science education literature. To participate in the discourse of science and build science knowledge, science education researchers call for students to construct explanations and arguments to make sense of phenomena and disciplinary ideas ([14], [15]). Sense-making conversations are “seen as the primary mechanism for promoting deep understanding of complex concepts and robust reasoning” [14]. Studies of undergraduate physics courses have examined how students approach solving physics problems and have found some students believe they must exclusively use equations and math to solve physics problems instead of also making sense of physical scenarios using common knowledge [16]–[18]. The authors of this previous work argue it would be more productive for these students to approach problem solving as opportunities for sense-making. Sense-making has also been studied in undergraduate engineering. A study of students solving statics problems found students who produced more self-explanations were stronger problem solvers [19].

This question of what constitutes disciplinary engagement in engineering science courses is critical to the study of responsive teaching in these courses, but it is a complex question that we cannot fully answer in this single paper. Instead, to make progress toward an understanding of disciplinary engagement—and therefore responsive teaching—in engineering science courses, in this paper we focus on the content that an instructor elicits, notices, and responds to during in-class formative assessment. Many engineering science instructors already use some type of formative assessment in their courses to elicit student feedback in real-time during a class meeting. Examples include both formal methods, such as "clicker" questions where students have to respond to a multiple-choice question using a personal student response system, and informal methods, such as when instructors pose questions to the class or when students ask questions of the instructor. Our prior research has suggested that dialogue initiated by instructor or student questions (what we call "*question-initiated dialogue*") allows instructors to elicit, notice, and respond to unexpected avenues of student thinking whenever they occur during the lesson [20].

In this paper we discuss the development and final version of a coding scheme that classifies the content that an instructor elicits, notices, and responds to during question-initiated dialogue. This coding scheme is called the **T**eacher **E**liciting, **NO**ticing, and **R**esponding (TENOR) Protocol. By applying the TENOR Protocol to in-class dialogue, we produce data that helps us to understand and argue about 1) the disciplinary substance of engineering science courses, 2) the amount and types of reasoning that instructors are encouraging from their students, 3) how instructors can and should enact responsive teaching. Specifically, the two research questions of this paper ask:

1. What type of instrument allows us to capture the different content that an instructor elicits, notices, and responds to during question-initiated dialogue?
2. Can this instrument identify differences between courses in the types of content discussed during question-initiated dialogue?

Development of the TENOR Protocol

We developed the **T**eacher **E**liciting, **NO**ticing, and **R**esponding (TENOR) Protocol by iteratively coding transcripts of question-initiated dialogue in four engineering science courses and evaluating the coded data to ensure that the coding scheme fit the data and fully captured the content being discussed in an engineering science course**.** We created these transcripts from the lecture capture videos of three class meetings of four courses, for a total of twelve class meetings. We transcribed sections of dialogue that began when the instructor or a student asked a question (called an "invitation" in the TENOR Protocol) and contained all of the resulting discussion by the instructor or students. This discussion could include statements from the instructor or students (called "offers" in the TENOR Protocol) or additional questions. We then parsed the question-initiated dialogue into *utterances*—continuous sets of dialogue spoken by one person, which may be as short as a single word or number or as long as a few sentences.

First, two of us individually coded one, two, or three of the twelve total video recordings, applying the coding scheme to categorize every utterance of the question-initiated dialogue. When we disagreed on the coding of an utterance, we discussed it with each other, and with a third researcher who had been using the TENOR Protocol during real-time in-class observations, until a consensus was reached. At three separate times in the process, we modified the coding scheme to ensure that it fully captured the content being discussed in class. This resulted in the initial coding protocol (TENOR v0), two intermediate versions (TENOR v.1 and TENOR v.2), and our final coding scheme (TENOR v.3), as illustrated in Figure 1. Appendix A shows how each version of TENOR was applied to a segment of question-initiated dialogue from a small section of a 200-level chemical engineering course.

Version 0

We first began by using an existing coding scheme for characterizing productive forms of educational dialogue—the Scheme for Educational Dialogue Analysis (SEDA) [21]. The SEDA codes categorize dialogue into eight clusters: 1) invite elaboration or reasoning, 2) make reasoning explicit, 3) build on ideas, 4) express or invite ideas, 5) positioning and coordination of ideas, 6) connect, 7) guide direction of dialogue or activity, and 8) reflect on dialogue or activity. Each of these clusters contains sub-cluster codes that describe the dialogue in more detail. For example, three of the six codes within the "Invite elaboration or reasoning" cluster include "invite building on / elaboration / (dis)agreement / evaluation of another's contribution or view," "ask for explanation or justification," and "ask for elaboration or clarification."

While SEDA was developed to apply to a wide range of educational contexts, after using SEDA to code our transcripts we found that the eight clusters did not map well to the engineering science courses in our data. This is not surprising, as SEDA was developed using pre-college courses in topics such as natural science and history, which involve more argumentation-based discussion than undergraduate engineering science courses. For example, SEDA has many sub-cluster codes focused on addressing another's contribution to the discussion. The "invite building on / elaboration / (dis)agreement / evaluation of another's contribution or view" code is further described in SEDA as "Inviting participants to take up others' contribution(s)/ideas in order to promote the extension, elaboration, or deepening of ideas" [21]. While an instructor could certainly encourage students to build on their peers' ideas, we rarely saw this happening in engineering science courses. Instead, we more often saw students asking questions in response to something the instructor had previously said. In order to fit this behavior into a SEDA code, we considered the student to be "building off" of the instructor's contribution.

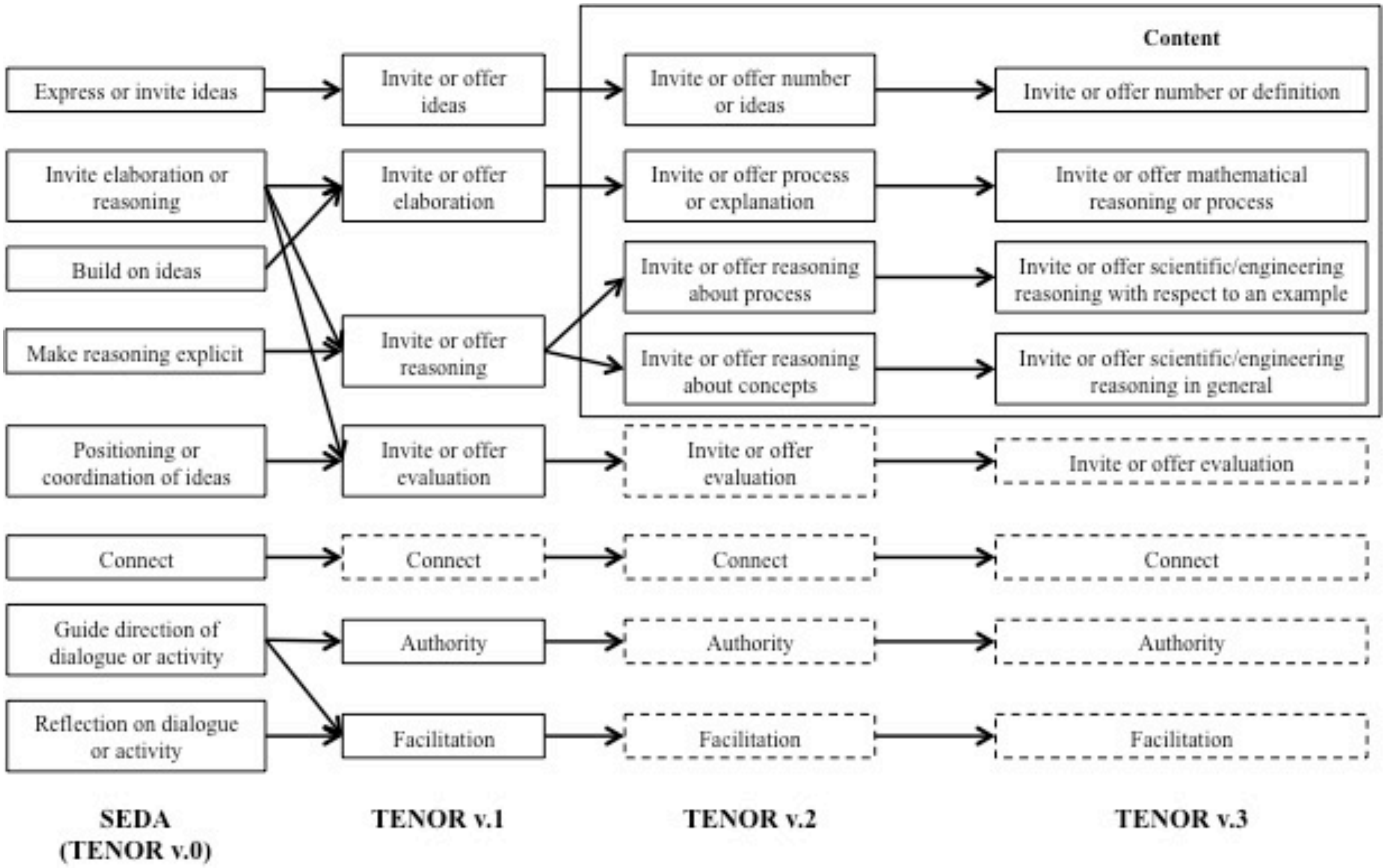

Figure 1. The evolution of the TENOR Protocol.

Version 1

Due to the shortcomings of SEDA when applied to our data, we rearranged and renamed the SEDA clusters and sub-cluster codes for TENOR v.1, as shown in Figure 1. When coding the initial data with SEDA, we frequently used certain sub-cluster codes of the *Invite elaboration or reasoning* cluster—particularly “ask for explanation or justification,” and “ask for elaboration or clarification.” We viewed the difference between these two codes as whether or not there was an invitation for reasoning. Therefore, we created two TENOR categories from the SEDA *Invite elaboration or reasoning* cluster: *Invite or offer elaboration* (without reasoning) and *Invite or offer reasoning*. The SEDA *Invite elaboration or reasoning* cluster also contained a sub-cluster code that was an invitation for evaluation of a student’s idea, which we turned into its own TENOR category: *Invite or offer evaluation*.

While a single SEDA cluster (*Invite elaboration or reasoning*) provided the invitation (i.e. the question) for multiple TENOR categories described in the previous paragraph, the offering of these categories (i.e. the answer) in the TENOR Protocol came from multiple different SEDA clusters. *Offer elaboration* was derived from the SEDA *Build on ideas* cluster, *Offer reasoning* was derived from the SEDA *Make reasoning explicit* cluster, and *Offer evaluation* was derived from the SEDA *Positioning or coordination of ideas* cluster.

We split the *Authority* sub-cluster code from the SEDA *Guide direction of dialogue or activity* cluster into its own TENOR category (*Authority*), and combined the rest of the codes in the *Guide direction of dialogue or activity* cluster with the SEDA *Reflection on dialogue or activity* cluster to create a *Facilitation* category with two codes. Lastly, we brought the SEDA *Express or invite ideas* cluster and the *Connect* clusters to the TENOR Protocol unmodified.

We applied TENOR v.1 to a subset of the video transcripts, once again coding question-initiated dialogue. After individually coding the transcripts and discussing our results as a group, we found the categories were too vague to consistently code the transcripts and did not exactly capture the content of engineering science courses. For example, it was difficult to consistently distinguish whether utterances should be coded as *Invite or offer elaboration* or *Invite or offer reasoning*. We determined that “elaboration” utterances were “How?” questions that dealt with clarifying the steps of mathematical processes, and “reasoning” utterances were “Why?” questions that dealt with further explaining scientific concepts.

Version 2

The iterations we made in the TENOR Protocol v.2 primarily focused on further modifying the first three categories in v.1 around our ideas about the content of engineering science courses. We renamed the *Invite or offer ideas* category to *Invite or offer number or ideas* to better clarify that this category was to be used when a student or the instructor offers a numerical answer, which is common in engineering science courses. We also renamed *Invite or offer elaboration* to

Invite or offer process or explanation, a category to capture the students describing the mathematical process they followed to reach an answer. Lastly, we split *Invite or offer reasoning* into two separate categories to distinguish between dialogue justifying a particular mathematical process (*Invite or offer reasoning about process*) and dialogue about natural phenomena or engineering models (*Invite or offer reasoning about concepts*). These new category names followed our internal rules about distinguishing between "elaboration" and "reasoning" in v.1.

The other four categories (*Invite or offer evaluation, Connect, Authority*, and *Facilitation*) stayed the same from version 1, although we did add two sub-category codes to *Facilitation* to capture an instructor inviting questions from students and an instructor revoicing something that a student had said, which was often done to make sure that he or she understood the student's question or answer.

We applied the TENOR Protocol v.2 to the transcripts, and once again noted difficulty distinguishing between inviting or offering process and reasoning about process. We developed internal rules amongst ourselves to distinguish between the two; for example, if a student asked if they could use a particular process, we coded this utterance as *invite process*. One such example is when an electrical engineering student asked, "Could we just use KCL [Kirchhoff's Current Law] at this point?" However, if the instructor asked students *why* they could make a particular assumption when doing a problem, we coded this utterance as *inviting reasoning about process*. One such example is when a chemical engineering instructor said, "So if I say that P_A is fixed, the partial pressure of A in vapor on the gas phase is fixed. Why can I make that statement? Or why can I make that assumption?" With these rules, we had fewer discrepancies between our individual coding of dialogue. Finally, we presented this coding scheme to a group of engineering education students for expert validation and further refinement. In response to their feedback we modified the names of the four content categories.

Final TENOR Protocol (v.3)

In the final version of the TENOR Protocol, version 3 (Table I), there are five top-level categories: *Content*, *Evaluation of Ideas*, *Connect*, *Authority*, and *Facilitation*—each with a number of sub-category codes. When coding question-initiated dialogue within a transcript, a single code is applied to each utterance from an instructor or student. If a single utterance covers multiple codes, it may be divided into a set of utterances, each with its own code. When coding question-initiated dialogue, a researcher also indicates whether the speaker of each utterance is the instructor or a student. Furthermore, a researcher can indicate when a new student is speaking by numbering the students within each set of question-initiated dialogue. Two examples of dialogue from a small section of a 200-level chemical engineering course are presented in Appendix B.

Table I. Final TENOR Protocol Categories and Sub-Category Codes

Category	Sub-Category Code
Content	Invite number or definition
	Offer number or definition
	Invite mathematical reasoning or process
	Offer mathematical reasoning or process
	Invite scientific/engineering reasoning with respect to an example
	Offer scientific/engineering reasoning with respect to an example
	Invite scientific/engineering reasoning in general
	Offer scientific/engineering reasoning in general
Evaluation of Ideas	Invite evaluation of a number or idea
	Compare alternate views or methods
	Express agreement or confirmation
	Express disagreement
	Challenge a number or idea
	Acknowledge being incorrect
Connect	Refer back to any point in the semester
	Refer forward to any point in the semester
	Connect beyond class
	Invite inquiry beyond lesson
Authority	Invoking a Law
	Invoking a textbook, solutions posted by the instructor, etc.
Facilitation	Guide dialogue or give a hint
	Reflect
	Invite questions
	Revoicing without making agreement

There are four sub-category codes within the *Content* category, each of which can be invited (by a question) or offered (as a statement). In *Invite or Offer number or definition*, the dialogue is usually very brief and focuses on such content as a number, qualitative representation of a specific number (e.g. a variable name), a definition, or a vocabulary word. In *Invite or offer mathematical reasoning or process*, the dialogue focuses on the mathematical steps undertaken to reach a particular answer (rather than the underlying science or engineering concepts). The other two sub-category codes within the *Content* category focus on scientific or engineering reasoning. In both of these codes the dialogue centers on a natural phenomenon or a mathematical model of a natural phenomenon. The difference between the codes is whether they relate to a specific example or problem that is currently being discussed in class (*Invite or offer scientific/engineering reasoning with respect to an example*), or if they are a general question about the course content outside of a specific example (*Invite or offer scientific/engineering reasoning in general*).

The other four categories in the TENOR Protocol are more similar to the original SEDA clusters. The second category of dialogue is *Evaluation of Ideas*, in which the instructor and students *Invite evaluation of a number or idea* or offer an evaluation. When offering an evaluation of an idea or number, a speaker can compare an alternate method, express agreement, express disagreement, challenge, or acknowledge that he or she was incorrect. Students or instructors can also *Connect* (the third category) ideas to other content previously discussed, to be discussed in the future, or beyond the course. Speakers, primarily instructors, can also invite inquiry beyond the lesson. The fourth category, *Authority*, captures when a speaker refers to an authority outside of the lecture, which could be a specific Law, the textbook, or a solution set posted by the instructor. This category is <u>not</u> to be used to indicate that the instructor is speaking as an authority figure him- or herself. In the fifth category, the instructor can *Facilitate* the discussion in a number of ways. He or she can guide the dialogue, which includes giving a hint to a current question or a problem. He or she can also reflect on the discussion, invite questions, or revoice something said by a student. When revoicing something said by a student, the instructor is typically making sure that he or she understands the students' question, or making sure that the entire class has heard the student. If the instructor revoices something said by a student to signify that the student is correct, this should be coded with the *Express agreement or confirmation* sub-category code, which is a part of the *Evaluation of Ideas* category.

This final version of the TENOR Protocol answers the first research question of this paper: *What type of instrument allows us to capture the different content that an instructor elicits, notices, and responds to during question-initiated dialogue?* We envision a researcher applying the TENOR Protocol to instances of question-initiated dialogue in a given class meeting in order to examine how the instructor influences the discourse of the class. Looking at the instructor's questions, "invitations" to the students, shows the different content that he or she elicits from students. While asking any question to engage students is a positive behavior, we believe it is also important to identify what type of questions the instructor is asking. Analyzing the instructor's utterances may find that he or she primarily elicits brief numerical answers, or that the instructor also elicits scientific or engineering reasoning from students. Looking at the instructor's responses to students' questions and answers, "offers" of explanation, shows how the instructor notices and responds to student thinking. Analyzing the instructor's utterances may find that he or she primarily responds to numerical questions with brief numerical answers, or that the instructor “escalates” simple student numerical questions to responses or additional questions for students that involve mathematical or scientific/engineering reasoning.

Application of TENOR Protocol

In this section, we turn to the second research question of this paper: *Can this instrument* (the TENOR Protocol) *identify differences between courses in the types of content discussed during question-initiated dialogue?* To answer this research question, we present an example of how question-initiated dialogue coded with the TENOR Protocol can be analyzed to investigate the type of content being discussed by the instructor and the students. Specifically, we apply this coding scheme to question-initiated dialogue from half of the courses on which we developed TENOR v.3. Specifically, we study three meetings of two engineering science courses to investigate the following three questions:

a. How much students are asking and answering questions?
b. What content are students discussing when asking and answering questions?
c. What content are instructors discussing when asking and answering questions?

Although we have not computed inter-rater reliability for our protocol or focused on edge cases where the protocol may not be clearly applied, this example does demonstrate the feasibility using the TENOR protocol to study question-based dialogue in engineering science courses. This analysis also shows that the TENOR Protocol is capable of identifying differences in the degree to which instructors are encouraging student sense-making.

We chose two very different courses to highlight the abilities of the coding scheme: a 400-level mechanical engineering course taught to third- and fourth-year students in a traditional lecture style (Course 1), and a small section of a 200-level chemical engineering course designed for second- and third-year students which featured both lecture and active learning (Course 2) . The enrollment of the Course 1 was approximately 60 students, and it met for 80 minutes twice per week. The enrollment of Course 2 was approximately 25, and it met for 50 minutes three times per week.

In Course 1, students were seated in rows facing the front of the classroom and we perceived the professor to be mainly focused on delivering content. He spoke a majority of the time during his lecture, but he did invite and ask questions during this lecture. In Course 2, students sat with the tables arranged in small groups of six students. Instead of lecture, we perceived the course to be focused on conversation around the examples. We saw the lecture being led by students’ ideas and contributions more than a delivery of content. Since we observed these clear differences in Course 1 and 2, we knew that these would provide a good test for our TENOR protocol. The full data set for Course 1 and 2 can be found in Appendix C.

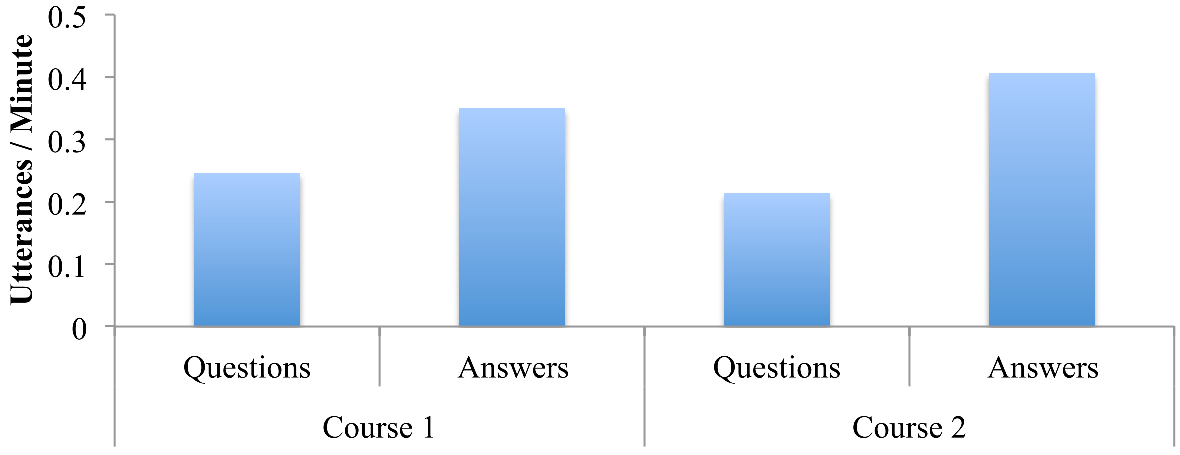

Figure 2. Number of student questions & answers per minute of class

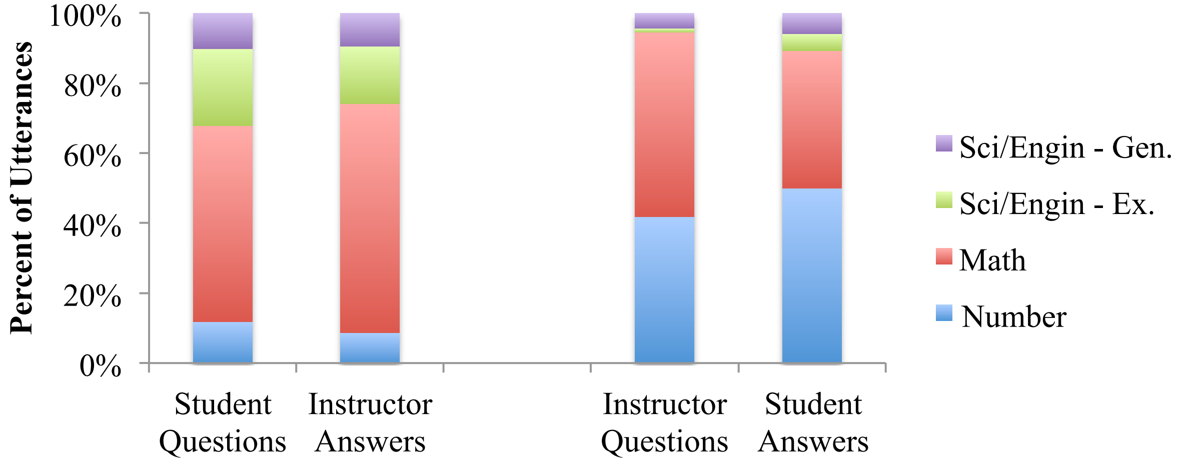

Figure 3. Content of student and instructor utterances in Course 1

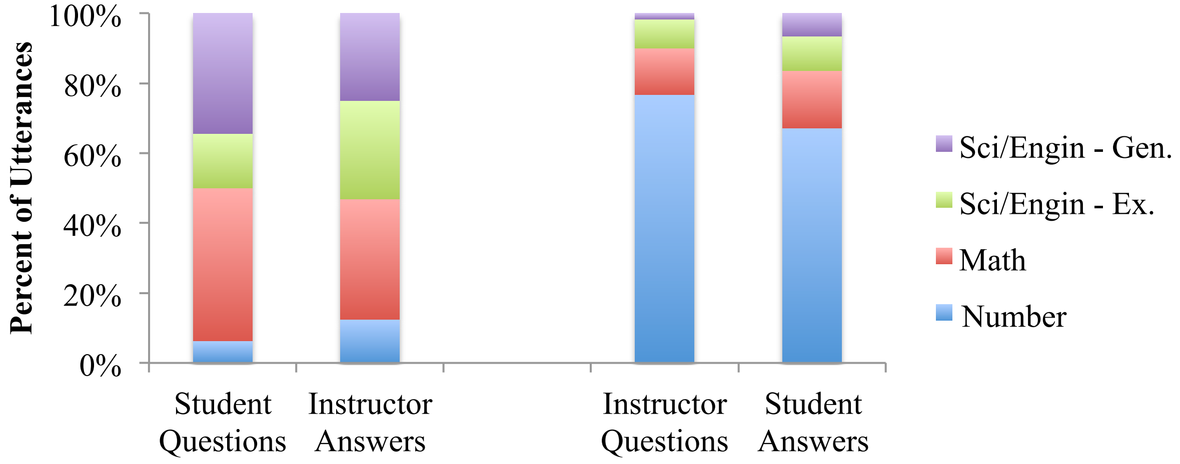

Figure 4. Content of student and instructor utterances in Course 2

Rate of Student Questions and Answers

First we present the rate of student utterances in the *Content* category, which is the number of times that students talked per minute of class time over the three course meetings (Figure 2). We separate this data into questions asked by students and answers given by students (to questions asked by the instructor or other students). It is necessary to normalize these data because the two courses met for different lengths of time (Course 1 - 80 minutes/meeting, Course 2 - 50 minutes/meeting). These data show that students in Course 1 did ask slightly more questions than Course 2 (1 question every 4 minutes vs. 1 question every 5 minutes). This may imply this instructor is open to more questions or the students are more confused. In contrast, students in Course 2 answer more questions than in Course 1 (8 answers every 20 minutes vs. 7 answers every 20 minutes). This may imply the instructor of Course 2 asks more questions than the instructor of Course 1 or the students in Course 2 are more willing to answer questions.

Substance of Student and Instructor Questions and Answers

After examining how much students talk during class, we turned to analyze the content of students' and instructors' questions and answers (Figures 3 and 4). As is expected, the data show a clear relationship between the content of questions and answers. The content of student questions is generally reflected in the content of instructor answers, and the content of instructor questions is generally reflected in the content of student answers. Furthermore, the data show that the majority questions from students and the instructor in both courses *did not* addressed science and engineering reasoning (student science and engineering questions were 32% in Course 1 and 50% in Course 2; instructor science and engineering questions were 5% in Course 1 and 10% in Course 2).

However, when looking closer at the data, we can see differences between the two courses. The students in Course 1 asked a larger percentage of questions addressing science and engineering reasoning than the students in Course 2, and the instructor of Course 1 asked a larger percentage questions addressing science and engineering reasoning than the instructor of Course 2. This maps to our expectations of the two courses based on our observations, as we saw Course 2 being led by students' ideas and contributions more than a delivery of content. Furthermore, the instructor of Course 1 asked a larger percentage of questions addressing mathematical reasoning or process than questions addressing numbers or definitions, whereas the instructor of Course 2 was the opposite. This is also unsurprising because the discipline of Course 1 was much more mathematical than that of Course 2.

Conclusion

Paper Summary

In this paper we have addressed two research questions:

1. What type of instrument allows us to capture the different content that an instructor elicits, notices, and responds to during question-initiated dialogue?
2. Can this instrument identify differences between courses in the types of content discussed during question-initiated dialogue?

We addressed the first research question by describing the development and final version of the **T**eacher **E**liciting, **NO**ticing, and **R**esponding (TENOR) Observation Protocol. The TENOR Protocol can be used to code question-initiated dialogue in engineering science courses, and this coded data can be analyzed to identify the content that an instructor elicits, notices, and responds to.

We have also shown examples of such an analysis applied to two different engineering science courses. The data shows that there was a higher frequency of sense-making questions from students and the instructor in Course 2 than in Course 1. This quantitative data from the TENOR Protocol reflects our qualitative observations of both courses: that Course 1 was oriented towards delivering content and Course 2 was oriented towards students' ideas and conversation about examples. Therefore, we can be confident that the TENOR Protocol identifies differences between courses in the amount of student sense-making and instructor responsiveness to this sense-making, answering our second research question.

Future Work

We are planning much future work that employs the TENOR Protocol to address the content of question-initiated dialogue and, eventually, responsive teaching in engineering science courses. First, we will apply the TENOR Protocol to a different set of classroom data than was used to develop the protocol. During the Fall 2018 semester we conducted 21 observations of 7 different engineering science courses. These courses range from 200- to 400-level across five departments. After transcribing question-initiated dialogue from the lecture capture video of these 21 class meetings, we will code these utterances and use this data set to compute inter-rater reliability. We are aiming for 80% agreement, an accepted level of the inter-rater reliability for an observation protocol [22]. After the inter-rater reliability of the transcript coding has been established and all edge-case utterances have been discussed, we will analyze this data to investigate the different ways in which these instructors notice, elicit, and respond to different content.

Another line of future work involves using these results to help better define the "disciplinary engagement" of engineering science courses. This will allow us to understand what responsive teaching looks like in engineering science courses, and outline evidence-based techniques that instructors can use to intentionally engage in responsive teaching in their courses and encourage student disciplinary engagement. These forthcoming evidence-based techniques will help to develop a community of practice around responsive teaching in engineering science courses.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. DUE-1711533. 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.

References

[1] Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching*, *48*(10), 1109-1136.

[2] Hammer, D., Goldberg, F., & Fargason, S. (2012). Responsive teaching and the beginnings of energy in a third- grade classroom. *Review of Science, Mathematics, and ICT Education*, 6(1), 51-72.

[3] Robertson, A. D., Scherr, R., & Hammer, D., Eds. (2015). *Responsive Teaching in Science and Mathematics*. Routledge: New York, NY.

[4] Empson, S. B., & Jacobs, V. R. (2008). Learning to listen to children's mathematics. In D. Tirosh & T. Wood (Eds.), *The International Handbook of Mathematics Teacher Education, Volume 2: Tools and Processes in Mathematics Teacher Education* (257-281). The Netherlands: Sense Publishers.

[5] Pierson, J. L. (2008). *The relationship between patterns of classroom discourse and mathematical learning* (Doctoral dissertation). Retrieved from Texas ScholarWorks.

[6] Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal*, *93*(4), 373-397.

[7] Richards, J. (2013). Exploring what stabilizes teachers' attention and responsiveness to the substance of students' scientific thinking in the classroom (Doctoral dissertation). Retrieved from Digital Repository at the University of Maryland.

[8] Berland, L. K., & Hammer, D. (2012). Framing for scientific argumentation. *Journal of Research in Science Teaching*, *49*(1), 68-94.

[9] Hammer, D. (1997). Discovery teaching, discovery learning. *Cognition and Instruction*, 15(4), 485–529.

[10] Wendell, K. B., Watkins, J., & Johnson, A. W. (2016). Noticing, assessing, and responding to students' engineering: Exploring a responsive teaching approach to engineering design. *Proceedings of the 123rd American Society for Engineering*

Education Annual Conference and Exposition. Washington, DC: American Society for Engineering Education.

[11] Crismond, D. P. & Adams, R. S. (2012). The Informed Design Teaching and Learning Matrix. *Journal of Engineering Education, 101*(4), 738-797.

[12] Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *Journal of engineering education*, *96*(4), 359-379.

[13] Lord, S. M., & Chen, J. C. (2014). Curriculum design in the middle years. *Cambridge Handbook of Engineering Education Research*, 181-200.

[14] Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in philosophy and education*, *27*(4), 283-297.

[15] Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26-55.

[16] Gupta, A, & Elby, A. (2011). Beyond Epistemological Deficits: Dynamic explanations of engineering students' difficulties with mathematical sense-making. *International Journal of Science Education,* 33(18), 2463-2488.

[17] Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction,* 12(2), 151-183.

[18] Lising, L., & Elby, A. (2005). The impact of epistmology on learning: A case study from introductory physics. *American Journal of Physics*, 73(4), 372-382.

[19] Litzinger, T. A., Meter, P. V., Firetto, C. M., Passmore, L. J., Masters, C. B., Turns, S. R., ... & Zappe, S. E. (2010). A cognitive study of problem solving in statics. *Journal of Engineering Education*, *99*(4), 337-353.

[20] Johnson, A. W., & Finelli, C. J. (2017). Responsive teaching in undergraduate engineering courses. *Proceedings of the 124th American Society for Engineering Education Annual Conference*. Washington, DC: American Society for Engineering Education.

[21] Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., ... & Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction*, *9*, 16-44.

[22] Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Appendix A. Versions of TENOR Protocol applied to the same question-initiated dialogue from Course 1

Speaker	**Utterance**	**SEDA (TENOR v.0)**	**TENOR v.1**	**TENOR v.2**	**TENOR v.3**
Instructor	If ΔH is negative, what does that say about the temperature?	Invite ideas	Invite ideas	Invite ideas	Invite number or definition
Student	*No response*				
Instructor	Temperature has to be lower at point 2.	Build off ideas	Offer ideas	Offer idea	Offer number or definition
Student	*Inaudible response*				
Instructor	So we simplify the energy balance equation to "ΔH = work shaft" because we know the system is doing work to turn the turbine, it's negative work.	Positioning and coordination	Offer elaboration	Offer reasoning about process	Offer scientific /engineering reasoning w.r.t. an example
Student	Why is the ΔH associated with the…	Invite elaboration or reasoning	Invite reasoning	Invite reasoning about process	Invite scientific /engineering reasoning w.r.t. an example
Instructor	ΔH is equal to internal energy, which is a function of temperature, plus PV, and we know that PV equals what…	Build off ideas	Offer reasoning	Offer reasoning about process	Offer scientific /engineering reasoning w.r.t. an example
Instructor	$PV = NRT$. So temperature goes up, PV goes up. So ΔH, or this should be H, H is higher when temperature is higher because H is made up of internal energy and PV.	Build off ideas	Offer reasoning	Offer reasoning about process	Offer scientific /engineering reasoning w.r.t. an example
Student	I just wanted to make it clear how [inaudible] work shaft	Invite elaboration or reasoning	Invite elaboration	Invite reasoning about concepts	Invite scientific /engineering reasoning in general
Instructor	So, work shaft is associated with a moving part, ok. So if you have a bucket and you're stirring it, its like you're cooking the spaghetti sauce. Your spatula is the shaft arm. In this case, you have a turnbuckle (and) it's moving. That's shaft work. In a car, you have pistons that move. That's shaft work.	Build off ideas Connect	Offer elaboration Connect beyond class	Offer reasoning about concepts	Offer scientific /engineering reasoning in general

Appendix B. Examples of *Content* sub-category codes from Course 1

Example 1				
Speaker	**Utterance**	**Sub-Category Code**		**Explanation of Code**
Instructor	So then, if I have both temperature change and pressure change I just have to say, well, Delta H, for example, is equal to 0 plus integral of cp dT.		----	*Introduction to instructor's question (next row) to provide context.*
Instructor	And what am I talking about here? Liquids, solids, or gas?	Invite	Scientific reasoning or process with respect to an example	*Instructor asks a question about the state of matter that is described by a specific equation*
Student 1	Gas	Offer	Scientific reasoning or process with respect to an example	*Student gives an answer. Because his or her answer follows an* Invite reasoning or process with respect to an example *utterance it is coded the same.*
Instructor	Gas. Ideal gas.		Express agreement or confirmation	*Instructor repeats students' answer, confirming that it is correct.*
Instructor	If it's a liquid or solid, this is vhat delta p plus integral of cp dT.	Offer	Scientific reasoning or process with respect to an example	*Instructor gives the equation for other states of matter.*
Instructor	Yes, [student name]?		----	*Instructor calls on a student, who has a question*
Student 2	So, um, you said in the video that the- the delta H equals the integral of cp dT but that works also for a non-ideal gas if…	Invite	Scientific reasoning or process in general	*Student asks about a video provided by the instructor. H or she asks for clarification about the assumptions one makes when using this particular equation. However, this question does not directly relate to the example, it is an extension to a general scientific / engineering concept.*
Instructor	You can- if, if all else fails and you have a non-ideal gas you can use it as an approximation, but you have to make sure that you do not have a volume change.	Offer	Scientific reasoning or process in general	*Instructor answers students' question*

Example 2

Speaker	**Utterance**	**Sub-Category Code**		**Explanation of Code**
Instructor	Okay. Do we have something that, uh, anybody, anybody has something they want to share? Okay, go ahead. [calls on student]	Invite	Mathematical reasoning or process	*Instructor asks for someone to share the mathematical process that their group has developed. (Note - it is not apparent from the instructor's question that she is asking for a mathematical process. However, because the students gives a process as an answer, this line is coded* Invite mathematical reasoning or process.*)*
Student	We said you could change it first [unintelligible] 75 degrees Celsius still at vapor. And then, um, constant temperature to a liquid. And then change it from 75 degrees Celsius liquid to the [final?]?	Offer	Mathematical reasoning or process	*Student gives their mathematical process.*
Instructor	Okay. Looks good to me. Okay?		Express agreement or confirmation	*Instructor confirms that students' process is correct*
Instructor	Uh, we know that we have the delta H vaporization which is negative delta H condensation at 75 degrees C. That's what we, uh, need. We can go from 400 to 75 because we know we can- say that delta H is equal to what? What is delta H for this step?	Offer	Mathematical reasoning or process	*Instructor explains the mathematical process that the student presented.*
Instructor	What is delta H for this step?	Invite	Number or definition	*Instructor asks a definitional question*
Student	The integral of cp dT.	Offer	Number or definition	*Student offers a brief, definitional answer.*
Instructor	The integral of cp dT.		Express agreement or confirmation	*Instructor revoices student's answer to confirm that it is correct*
Instructor	We're going from 400 to 75 degrees C. Mmmkay? And it's the same thing on the other side. Mmmkay?	Offer	Mathematical reasoning or process	*Instructor explains more of the mathematical process*

Appendix C. Number of utterances coded with each *Content* sub-category code in Course 1 and 2

Course 1

	Number or definition		Mathematical reasoning or process		Scientific/engineering reasoning with respect to an example		Scientific/engineering reasoning in general		
Instructor	**Invite**	**Offer**	**Invite**	**Offer**	**Invite**	**Offer**	**Invite**	**Offer**	**Total**
Class Meeting 1	20	3	8	11	1	1	4	1	49
Class Meeting 2	7	5	38	49	0	11	0	5	115
Class Meeting 3	11	2	2	16	0	7	0	5	43
Total	38	10	48	76	1	19	4	11	207

	Number or definition		Mathematical reasoning or process		Scientific/engineering reasoning with respect to an example		Scientific/engineering reasoning in general		
Students	**Invite**	**Offer**	**Invite**	**Offer**	**Invite**	**Offer**	**Invite**	**Offer**	**Total**
Class Meeting 1	0	19	3	7	0	0	0	5	34
Class Meeting 2	7	11	27	26	11	4	2	0	88
Class Meeting 3	0	12	3	0	2	0	4	0	21
Total	7	42	33	33	13	4	6	5	143

Course 2

Instructor	Number or definition		Mathematical reasoning or process		Scientific/engineering reasoning with respect to an example		Scientific/engineering reasoning in general		
	Invite	**Offer**	**Invite**	**Offer**	**Invite**	**Offer**	**Invite**	**Offer**	**Total**
Class Meeting 1	8	4	1	2	0	3	1	9	28
Class Meeting 2	15	1	2	3	1	3	0	3	28
Class Meeting 3	23	3	5	17	4	12	0	4	68
Total	46	8	8	22	5	18	1	16	124

Students	Number or definition		Mathematical reasoning or process		Scientific/engineering reasoning with respect to an example		Scientific/engineering reasoning in general		
	Invite	**Offer**	**Invite**	**Offer**	**Invite**	**Offer**	**Invite**	**Offer**	**Total**
Class Meeting 1	1	2	2	1	1	2	7	4	20
Class Meeting 2	0	18	0	1	1	1	2	0	23
Class Meeting 3	1	21	12	8	3	3	2	0	50
Total	2	41	14	10	5	6	11	4	93