Preliminary Analyses of Survey and Student Outcome Data using the Global Real-Time Tool for Teaching Enhancement (G-RATE)

Ms. Nikitha Sambamurthy, Purdue University, West Lafayette

Nikitha Sambamurthy is pursuing a Ph.D. in Engineering Education at Purdue University. Her research interests include: blended-learning in science, technology, engineering, and mathematics (STEM) education, and the implementation and assessment of games for engineering knowledge transfer.

Dr. Monica Farmer Cox, Purdue University, West Lafayette

Monica F. Cox, Ph.D. is an Associate Professor in the School of Engineering Education and is the Inaugural Director of the College of Engineering's Leadership Minor at Purdue University. She also serves as the Executive Director of the International Institute for Engineering Education Assessment (i2e2a). She obtained a B.S. in mathematics from Spelman College, a M.S. in industrial engineering from the University of Alabama, and a Ph.D. in Leadership and Policy Studies from Peabody College of Vanderbilt University. Her teaching interests relate to the professional development of graduate engineering students and to leadership, policy, and change in STEM education. Primary research projects explore the preparation of graduate students for diverse careers and the development of reliable and valid engineering education assessment tools. She is a NSF Faculty Early Career (CAREER) and Presidential Early Career Award for Scientists and Engineers (PECASE) recipient.

Dr. Yukiko Maeda, Purdue University, West Lafayette

Yukiko Maeda is an assistant professor of Educational Psychology in the area of research methodology in the Department Educational Studies. She has expertise in educational measurement and statistics including the application of multilevel modeling in educational research.

Kimberly Joy Perram
Mr. Vivian Gregory Alexander

Preliminary Analyses of Survey and Student Outcome Data Using the Global Real-Time Assessment Tool for Teaching Enhancement (G-RATE)

Abstract

This paper presents a brief overview of a pilot study conducted with a tool called the Global Real-time Assessment Tool for Teaching Enhancement (G-RATE) and the development of instructor profiles. The purpose of the pilot study is to investigate how student perceptions of teaching practices on key aspects of the "How People Learn" (HPL) framework and students' end of the course grades differ among three conditions (i.e., (1) instructors were observed once and received no feedback on their instruction during the semester (C1), (2) instructors were observed once and received e-mail feedback about their instruction mid-semester (C2), and (3) instructors were observed once, received e-mail feedback about their teaching, and engaged in a face-to-face discussion about their instruction (T)). Multivariate Analysis of Variance (MANOVA) indicated students' perceptions of teaching practices differ among the three conditions (C1, C2, and T). Further investigation with post-hoc analyses and descriptive statistics indicated that C1 and C2 do not differ significantly except in a learning-centered area, and C2 and T showed significant differences in all key areas. Interestingly, students under C2 reported higher average rating scores on these areas than students taught by instructors under T. This suggests that providing feedback to instructors is likely to help them to reflect on their own instruction. Comparison of achievement among groups indicated that there are statistically significant differences among groups. No students scored very low for their final grades (less than 60% in achievement in T condition), and it seems that students in the T group tend to do better than students in C1 or C2 groups, although effect size is relatively small. However, the nature of the review session at an individual meeting needs to be refined for highlighting the utility of feedback provided by G-RATE. This paper concludes with insight about future work using the G-RATE.

Introduction and G-RATE Background

The Global Real-time Assessment Tool for Teaching Enhancement (G-RATE) was developed in 2010 to provide multidimensional feedback to instructors about classroom instruction based upon aspects of Bransford, Brown, and Cocking's "How People Learn" (HPL) framework.[1] This framework explores the extent to which instruction represents knowledge-centeredness (i.e. how well students learn knowledge to support understanding and expertise of a subject), learner-centeredness (i.e. attention paid to the knowledge and beliefs learners bring to an educational environment), assessment-centeredness (i.e. opportunities for feedback on learner understanding), and community-centeredness (i.e. supporting learners ability to learn from each other).

The G-RATE is comprised of five functions that serve various aspects within the tool.[3] The *administrator* function is used prior to the start of a class or lab and modifies the observation parameters of the G-RATE. The *observer* function records direct, real-time instructional data in the form of code strings during a class or lab. The *student* function provides assesses students via Likert scale survey items that may be used formatively or summatively for a class or lab. The *instructor* function provides instructors with reflective items that may be used by instructors to

explore their pedagogy after a class or lab session. Finally, the researcher function compiles the data collected by the other G-RATE functions.

While aspects of the G-RATE have been written, this paper represents the first time that G-RATE data across multiple classrooms have been linked to students' academic outcomes. Prior papers have presented a general overview of the tool,[2,3] have explored the reliability and validation of the survey component of the student survey,[4,5] and have presented details about the revision of the observer function of the G-RATE.[6] Informed from the VaNTH Observation System, a classroom observation tool used to assess instruction in bioengineering classrooms at Vanderbilt, Northwestern, Texas, and Harvard/MIT,[7] the G-RATE has the potential to assess instruction at the K-12 and higher education levels in a variety of contexts (e.g., classrooms and laboratories).

Methodology

Sampling and Samples. Twelve sections of first-year engineering classes with approximately 120 students each were recruited for a pilot study in spring 2013. Students in each section were provided an overview of the research study and were given options to participate or not participate in the study. Students who completed consent forms within the study gave the research team permission to obtain de-identified copies of all course deliverables grades in the course. Total numbers of students per group are presented in Table 1.

Across all sections, 868 first-year engineering students and 10 instructors participated in this pilot study. A total of 12 sections (4 sections per condition) were randomly selected and assigned into three conditions: (1) instructors were observed once and received no feedback on their instruction during the semester (C1), (2) instructors were observed once and received e-mail feedback about their instruction mid-semester (C1), and (3) instructors were observed once, received e-mail feedback about their teaching, and engaged in a face-to-face discussion about their instruction (T). Ten instructors agreed to participate in the study, resulting in 3 C1 sections, 4 C2 sections, and, 3 T sections.

Research Design and Procedures. Table 1 depicts the summary of the research design employed for the study. Details about the observation, treatment, and outcome assessments also follow.

Group	Feedback Description	Outcome Assessment
Control 1 (C1) (3 instructors, 333 students completed surveys)	No pedagogical profile presented to instructor; No post-observation feedback given	Student survey of teaching practices framed within the context of the HPL framework Percent of achievement as determined by students outcomes across sections
Control 2 (C2) (4 instructors, 314 students completed surveys)	Pedagogical profile e-mailed to instructor No post-observation feedback given	
Treatment (T) (3 instructors, 221 students completed surveys)	Pedagogical profile e-mailed to instructor Post-observation feedback given	

Table 1. Summary of the Research Design

Observation. Researchers recorded one 2-hour class of each section at the beginning of the semester and used the G-RATE to observe 40 minutes of each class.

Treatment. After the observation, teaching profiles reflecting the instructor's classroom interactions were then created and distributed to the C2 group via email. These profiles presented information about the extent to which instructors demonstrated elements of the HPL framework along with information about when these occurrences happened within observed sections.[5] Profiles were presented to the T group through one-on-one sessions, in which a researcher would present a participant with a teaching profile and answer any questions regarding the profile. No profiles were presented to the C1 group.

Outcome assessments. Two outcomes were measured to evaluate the effect of feedback/reflection provided using G-RATE. First, all students in the course responded to a validated end-of-course G-RATE survey that was assigned as a homework within the first-year engineering course. Using final validated questions[5] representing the HPL framework, survey responses were calculated across sections and used to explore the extent to which students' perceived their instructors to demonstrate the main HPL elements of the survey- knowledge-centeredness, assessment-centeredness, and learner-centeredness. Second, students' grades were grouped and evaluated across the C1, C2, and T groups.

Data analysis. A variety of analyses were conducted to explore if the course outcomes (represented as %) are different across three research conditions (C1, C2, and T). First, descriptive analyses were conducted to examine if sub-factors in students' perception on teaching practice (i.e., Knowledge/Community-centeredness; learner-centeredness; Assessment centeredness) were correlated by group (C1, C2, and T) and differs by group. Next, using Multivariate Analysis of Variance with post-hoc comparison, differences in teaching practices among three research conditions were explored. Finally, Kruskal-Wallis test was performed to test differences in the percent of achievement by group.

Results and Discussion

Table 2 reports the descriptive statistics of students' perception of teaching practice by group.

	Knowledge-Centered Items				Assessment-Centered Items				Learner-Centered Items			
	Group				Group				Group			
	C1	C2	T	Total	C1	C2	T	Total	C1	C2	T	Total
Mean	3.18	3.39	3.27	3.28	3.10	3.34	3.20	3.21	2.40	2.60	2.49	2.50
Std. Dev.	0.55	0.60	0.59	0.59	0.69	0.72	0.70	0.71	0.47	0.45	0.47	0.47
N	333	314	221	868	333	314	221	868	336	315	221	872

Table 2. Students' ratings about the extent to which HPL elements were present in their courses.

The preliminary analyses of survey data and MANOVA by group indicated that the factor scores (represented as the average rating[1]) significantly differed by the type of groups (C1, C2, or T) (Wilks'= 0.96, F (2, 865)=5,203, p<.001). Post-hoc analyses and descriptive analyses also indicated that C1 and T seems not to differ significantly except for the learning-centered factor,

[1] Listwise deletion method was used for handling missing data.

but C2 and T showed significant differences in all three factors. C2 tended to have a larger average rating than T. However, the effect size was very small. In addition, the comparison of achievement among groups indicated that there are statistically significant differences in achievement among groups ($p=0.03$). Students in T groups do better than students in other two groups. Specially, there is no student who scored very low (i.e., less than 60%) in the T group.

Limitations and Future Directions

This study was able identify significant differences in students' perceptions of teaching practices across the three conditions (C1, C2, and T) and shows that providing feedback to instructors is likely to help them to reflect on their own instruction. However, there are some limitations to the study. One limitation is that there was only observer for the recorded data. Therefore, the rater reliability may not be as high as it should be. Future studies involving the G-RATE will include multiple observers reviewing each recording so that no observer codes the same video twice. Corollary to this, multiple training sessions for observers will help to ensure the delivery of quality feedback to instructors in the treatment condition and coding consistency among observers. Additionally, a standardized protocol will be created for all observers to follow when administering the treatment condition to ensure consistent implementation of the treatment. Subsequent studies will utilize an A-B-A within subject experimental design to assess the effectiveness of the treatment across the control and treatment conditions. The A-B-A is type of single subject design used to observe changes in an individual as the result of a treatment condition. In this experimental design, the participant is first receives baseline or control condition (A) i.e., no treatment administered, next the participant is administered the treatment (B), then finally there is a return to the baseline condition (A). This design helps detect any changes demonstrated by the instructor due to the treatment condition.

Another limitation of the study is that instructors' demographic characteristics were not included in the analysis. Subsequent studies will investigate the extent to which instructor characteristics such as previous pedagogical training and teaching experience relate to student perceptions of instructors' pedagogical practices and students' end of course grades. Further iterations of the G-RATE student survey will express clearly the four aspects of the HPL framework. This will be accomplished through the revising the current HPL questionnaire and conducting a series of exploratory and confirmatory factor analyses to revalidate and confirm the factor structure of the HPL. The fully validated HPL scale will then be administered at other institutions to investigate the invariance of the HPL construct in different student population.

Conclusion

The result of the study would suggest that providing feedback via pedagogical profiles in control group 1 and in the treatment group make a difference in instructors' demonstration of How People Learn framework elements. Such demonstrations are important, since the HPL framework has been found to increase student achievement. Preliminary results suggest that reviewing pedagogical feedback with instructors does not necessarily correlate with students' perceptions of their instructors' teaching practice. It is plausible that the feedback moderates the relationship between instructor pedagogical practices and student perceptions of the instructors' pedagogical practices and thus may not be directly observable. Another possible reason to explain the apparent lack of relation between the treatment and students' perceptions may be due

to a development trend that is not easily detected after one single pilot study. This could imply the need for the improvement of protocol used to debrief instructors post-observation. Future iterations also might more closely align the pedagogical experiences of instructors across groups to ensure that factors do not skew data indirectly. In addition, further refinement of the treatment conditions may be needed due to the small effect size, indicating the small difference between C2 and T groups.

Bibliography

1. Brown, A. L., & Cocking, R. R. (2000). *How people learn*. J. D. Bransford (Ed.). Washington, DC: National Academy Press.
2. Cox, M. F., Hahn, J., McNeill, N., Cekic, O., Zhu, J., & London, J. (2011). Enhancing the Quality of Engineering Graduate Teaching Assistants through Multidimensional Feedback. Advances in Engineering Education, 2(3).
3. Cox, M.F., Hahn, J., McNeill, N., & Kulkarni, A. (2010). Developing a Global Real-time Assessment Tool for the Teaching Enhancement of Engineering Graduate Teaching Assistants. 2010 Proceedings of the American Society for Engineering Education (7 pages).
4. Cox, M.F., Zhu, J., London, J.S., Hahn, J., & Ahn, B. (2012). Feedback about Graduate Teaching Assistants' Pedagogical Practices: Content Validation of a Survey Informed from Principles of the "How People Learn" Framework. In Greta Gorsuch (Ed.), Working Theories for Teaching Assistant and International Teaching Assistant Development. Stillwater, OK: New Forums Press..
5. Zhu, J., Li, Y., Cox, M. F., London, J., Hahn, J., & Ahn, B. (2013). Validation of a Survey for Graduate Teaching Assistants: Translating Theory to Practice. Journal of Engineering Education, 102(3), 426-443.
6. Sambamurthy, N., London, J., Hahn, J., Zhu, J., & Cox, M.F. (2013). Reliability of the Global Real-time Assessment Tool for Teaching Enhancement (G-RATE). ASEE Annual Conference. ASEE.
7. Harris, A. H., & Cox, M. F. (2003). Developing an observation system to capture instructional differences in engineering classrooms. *Journal of Engineering Education*, *92*(4), 329–336.