Lessons Learned From Using an Asynchronous Online Discussion Board to Facilitate Scientific Thinking in a Large Cognitive Psychology Lecture Class

Jordan Lippman, James Pellegrino, Renee Koziol, Emily Whitehair, University of Illinois at Chicago, Psychology Department (M/C 285), 1007 W. Harrison St., Chicago, IL 60657,
Email: lippmanj@hotmail.com, pellegjw@uic.edu, reneek2217@yahoo.com, Patchie6@aol.com

Abstract: Students' responses to scientific reasoning questions posted on an online discussion board were coded for completeness, coherence, accuracy, use of evidence, reasoning, and comprehensibility. Early performance was low for every dimension. By the end of the term, ratings of coherence and completeness increased but use of evidence decreased, and other dimensions stayed the same. Implications regarding the types of feedback students need to improve their understanding and use of scientific argumentation are discussed.

Introduction

As part of an ongoing instructional redesign effort, driven by theories of how people learn and the design of powerful learning environments (National Research Council, 2000), different constellations of course components have been implemented in a large lecture course on Cognition. Components were implemented and then modified, added, or removed based on student performance and reactions to an extended end-of-term evaluation. Thus, our approach to instructional design and assessment resembles the description of design research provided by Collins, Joseph, and Bielaczyc (2004).

One component was an asynchronous online discussion board (DB) designed to create smaller discourse communities (of roughly seven students) within the large lecture class. We wanted students to think deeply and make public their reasoning about course content, and especially the relationship between theory and evidence, as a means of formative evaluation of student learning and thinking (National Research Council, 2001). For each of seven questions spread over the term, students had to post initial responses before seeing those of their peers. Students were instructed to spend about 30 minutes writing a paragraph or two and to use (and reference) any sources they wished (e.g., the textbook and/or class PowerPoint slides). Once everyone had responded, each student chose a peer's initial post and provided constructive feedback. Throughout the term and in each of the questions, students were instructed to use empirical evidence to support the theoretical claims in their responses. Our objective was to help students improve their thinking and writing through multiple forms of feedback including assignment of grades to each post, graded peer feedback for each post, anonymous posting of best responses for the entire class to view, and emails from the teaching assistant about common misconceptions and mistakes in generating responses.

Method

The DB was implemented in an upper-level Introduction to Cognition lecture course at a large, urban, Midwestern University during the Spring 2004 semester. Students' initial responses to two questions – one from the beginning and one from the end of the term – were analyzed. The first question asked students to take a position (and support it with evidence) on whether knowledge is represented in the mind using multiple or a single representational system. The late semester question asked students to argue (using evidence as support) whether learning would always be best if a text is highly coherent or if incoherent texts could also facilitate learning.

The authors worked as a team to develop a coding scheme that captured the quality and variability in students' responses. The final coding scheme had six dimensions, which were all applied to each post on a scale of 0 (incorrect/incomplete) to 4 (scientific/excellent). The six dimensions were: *completeness* (answered all, some, or none of the parts of the question), *coherence* (quality of writing, structure, logical flow of ideas, brevity), *accuracy* (correctness of descriptions of theories and terms), use of *evidence* (amount, relevance, and/or quality of empirical or other type of less-scientific data used), *reasoning* (effectiveness of use of disconfirmatory, confirmatory, or less-scientific logic to present and connect evidence to their position), and *comprehensibility* (global assessment of level of understanding and integration of concepts and/or materials from class and text). Details of the coding scheme dimensions and definitions of the rating scale are available upon request. The initial posts of seventy-two students

(78% of the class) who responded to both questions (and who provided informed consent) were independently coded by two undergraduate raters who then reached agreement on ratings of each dimension.

Results

Qualitative descriptions of the modal response on the first question for each dimension illustrate the typical student's performance at the beginning of the term which turned out to be weak on multiple dimensions as expected. *Completeness* (mode = 2) - some aspects of the question were missing or left implicit. *Coherence* (mode = 2) - satisfactory writing quality: somewhat structured and coherent, but the flow of ideas is not highly logical. *Accuracy* (mode = 2) - fairly accurate: erroneous descriptions/use of some terms and/or minor conceptual contradictions. *Evidence* (mode = 0) - mostly anecdotal or irrelevant empirical evidence (if any) used. *Reasoning* (mode = 2) - failed to tie components of the answer together and/or ineffective/fallacious (i.e. study x *proved* theory y) reasoning. *Comprehensibility* (mode = 1) - response indicated faulty understanding, and/or no integration of concepts from class and text.

The Stuart-Maxwell test – a nonparametric statistic for matched-samples – was used to assess equivalence (i.e. nonsignificant results) of ratings on the early versus late semester questions. These analyses revealed nonsignificant changes for *accuracy, reasoning*, and *comprehension* but significant changes for *completeness*, *coherence*, and *evidence*. Students' responses were more likely to have higher *completeness* (mode = 4) and *coherence* (mode = 3) ratings, but lower ratings of their use of *evidence* (mode = 0) on the second question. Thus, despite some improvement over time, the quality of the students' responses remained weak in terms of scientific reasoning. Interestingly, using Spearman's rho (a nonparametric correlation), no performance dimension was significantly correlated with the TA's grading on the first question. For the second question, all of the dimensions except evidence were significantly and moderately correlated with TA evaluations; reasoning had the strongest correlation. This is likely due to the TA's leniency in grading the first question.

Conclusions and Implications for Instruction

At the beginning of the term, students had difficulty writing clear and concise arguments for a theoretical position and effectively using evidence to support their position. The DB activities and the various types of associated feedback appear to have helped them learn to write more coherent and complete answers by the end of the term. However, the objectives of helping students learn to more effectively use evidence and reason scientifically were not met. We think that the feedback provided was not explicit enough in these areas and that peer feedback may have focused on issues related to style (e.g., completeness and coherence). It may not be possible to use peer-feedback to accomplish our original learning goals if it is not mediated by a knowledgeable instructor. Additionally, students may need explicit feedback, guided by a rubric such as that derived for subsequent data coding. In the future we will be using a rubric to guide student writing and TA grading of responses. Finally, it may be difficult to achieve these learning aims in a truly low-stakes, formative manner without turning the activities into high-stakes assessments; when we subsequently made the DB voluntary rather than obligatory, it was rarely used at all despite students' perceptions of its potential learning value.

References

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design Research: Theoretical and Methodological Issues. *The Journal of the Learning Sciences, 13(1)*, p. 15-42.

National Research Council. (2000) *How People Learn: Brain, Mind, Experience, and School: Expanded Edition.* Bransford, J. D., Brown, A. L., Cocking, R. R., Donovan, S., & Pellegrino, J. W. (Eds.) Washington, D.C.: National Academy Press.

National Research Council (2001). *Knowing What Students Know: The Science and Design of Educational Assessment.* Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). Washington, DC: National Academy Press.