Alternative Approaches in Evaluation of Critical Thinking and Collaborative Testing

Dorababu Tadepalli
Associate Professor, GAE, School of Social Sciences, Jawaharlal Nehru University
New Delhi – 67, India
Currently serving as: Dean, School of Education Rusangu University Monze, Zambia

Abstract

Despite almost universal agreement that critical thinking needs to be taught in college, now perhaps more than ever before, there is much less agreement on definitions and dimensions. “Critical thinking can include the thinker’s dispositions and orientations; a range of specific analytical, evaluative, and problem-solving skills; contextual influences; use of multiple perspectives; awareness of one’s own assumptions; capacities for metacognition; or a specific set of thinking processes or tasks . Critical thinking is assessed in a variety of ways by individual teachers, but unlike many other college-level learning skills, it is also regularly assessed via a battery of standardized tests such as ACT’s Collegiate Assessment of Academic Proficiency (CAAP), the Collegiate Learning Assessment (CLA), ETS’ Proficiency Profile (PP), and a set of scoring rubrics known as the Valid Assessment of Learning in Undergraduate Education (VALUE). Although letting students work together on exam questions is still not a common instructional practice, it has been used more than might be expected and in a variety of ways. Sometimes students work together in groups; other times with a partner. Sometimes those groups are assembled by the instructor and sometimes students are allowed to select their partners or group members. Sometimes the groups share multiple exam experiences; other times they work collaboratively only once. Sometimes the group submits one exam with everyone in the group receiving that grade; other times students may talk about exam questions and answers but submit exams individually.

KEYWORDS: Assessment, Critical Thinking, Metacognition

Introduction:

Despite almost universal agreement that critical thinking needs to be taught in college, now perhaps more than ever before, there is much less agreement on definitions and dimensions. “Critical thinking can include the thinker’s dispositions and orientations; a range of specific analytical, evaluative, and problem-solving skills; contextual influences; use of multiple perspectives; awareness of one’s own assumptions; capacities for metacognition; or a specific set of thinking processes or tasks.” (p. 127)

Critical thinking is assessed in a variety of ways by individual teachers, but unlike many other college-level learning skills, it is also regularly assessed via a battery of standardized tests such as ACT’s Collegiate Assessment of Academic Proficiency (CAAP), the Collegiate Learning Assessment (CLA), ETS’ Proficiency Profile (PP), and a set of scoring rubrics known as the Valid Assessment of Learning in Undergraduate Education (VALUE).

Stassen, Herrington, and Henderson report on an interesting activity undertaken to answer several questions regarding critical thinking definitions. They wondered what dimensions of critical thinking were emphasized by these standardized tests and measures and whether those dimensions reflected how faculty at their institution defined critical thinking. "This exploratory analysis was intended to help us understand the relevance (or fit) of each of these tools to our faculty's priorities for students' critical thinking development." (p. 135)

They began by having a group of general education instructors generate an operational definition of critical thinking. The definition grew out of faculty responses to the following question and prompt: "What learning behaviors (skills, values, attitudes) do students exhibit that reflect critical thinking? Students demonstrate critical thinking when they …" (p. 128) Analysis of the instructors' responses resulted in 12 dimensions of critical thinking: judgment/argument, synthesizing, problem solving, evidence-based thinking, drawing inferences, perspective taking, suspend judgment, application, metacognition, questioning/skepticism, knowledge/understanding, and discipline-based thinking.

Next they looked at how the four standardized tests defined critical thinking. "To understand the commonalities between the four external sources and our campus's own critical thinking definition, we used our internal definition as the anchor definition and coded the external sources in relation to the categories present in that internal definition." (p. 130) A table in the article presents this comparison.

Their analysis shows that "judgment/argument is the predominant component of critical thinking reflected in all of the external assessment options (accounting for between one-half to over three-quarters of all the descriptors associated with critical thinking)." (p. 133) They found "substantial emphasis" on drawing inferences and evidence-based thinking and lesser emphasis on synthesizing, problem solving, and perspective taking. But some aspects of their definition of critical thinking, such as application, suspending judgment, metacognition, and questioning/skepticism, received no emphasis in the standardized assessments. "The results suggest that all three standardized tests address a narrow set of constructs present in the campus definition, with the primary focus on judgment/argument, evidence-based thinking, and drawing inferences." (p. 135)

This analysis was not a study of the validity of the items on the standardized assessments, but rather an exploration of how the basic construct of critical thinking was defined by the assessment tool. Furthermore, their campus definition was not assumed to be the "correct" definition. The authors note that it wasn't systematically vetted or compared with the responses of other groups of faculty on their campus or elsewhere, although the list of dimensions identified by these general education instructors is not notably unusual.

Collaborative Testing:

Although letting students work together on exam questions is still not a common instructional practice, it has been used more than might be expected and in a variety of

ways. Sometimes students work together in groups; other times with a partner. Sometimes those groups are assembled by the instructor and sometimes students are allowed to select their partners or group members. Sometimes the groups share multiple exam experiences; other times they work collaboratively only once. Sometimes the group submits one exam with everyone in the group receiving that grade; other times students may talk about exam questions and answers but submit exams individually.

Why let students work on exams collaboratively? There are a number of reasons, most of which have been explored empirically. When students discuss questions and possible answers, they intensely engage with the content, which increases the learning potential of an exam experience. The activity develops cooperation and communication skills. But the reason most often given is that working with other students decreases exam anxiety, particularly for those students whose levels of anxiety compromise their ability to perform on exams.

When asked why they don't use collaborative testing, most faculty report being afraid that students who have not prepared for the exam may inappropriately benefit from the knowledge of students who have studied. Grades should be measures of individual learning.

Based on a review of the literature on collaborative tests, these authors decided to explore four questions in their study: 1) How does collaborative testing affect test scores? 2) Is anxiety related to the effectiveness of collaborative testing? 3) Which students benefit from collaborative testing? and 4) How does the quality of interaction within the groups affect test performance? (p. 165)

The study included some unique design features. To deal with the potential problem of students coming to the exam unprepared and thinking the group would pull them through, students were told to prepare for the exam as if they were taking it individually. Those who would be taking the exam collaboratively would be randomly selected at the beginning of the period. On test days, those selected to take the exam collaboratively were moved to another room. Their group interactions were observed by a proctor who evaluated the quantity of interaction, the level of enthusiasm, and the degree of give-and-take displayed by the participants. (p. 166) The groups did not have to agree on answers, although they could change their answers based on discussions that occurred within the group. Tests were still submitted and graded individually.

As for how the collaboration affected test scores, the results were positive. "Collaborative testing was more successful for a significant majority of students than was traditional, individual testing, although the advantage (3.83%) was smaller than found in some previous studies." (p. 172) Based on standardized measures of test anxiety, "students with higher initial test anxiety scores were most likely to benefit from the collaborative testing procedure and to show a decrease in test anxiety when taking tests collaboratively." (p. 172) And finally, "high interaction scores, both proctor and student-rated, were related to better performance under the collaborative condition." (p. 172)

Consistent in the research on collaborative testing mechanisms is the favorable response they generate from students. Students regularly report enjoying the experience and indicate they would choose it in the future, if given the opportunity. In this study, students said the testing mechanism was beneficial because it helped them develop good working relationships with classmates and helped them remember material they had forgotten. There were students in this study who did perform better on the tests they took individually, and a third of those students thought they did better on the individual exams because they found the discussions with other students confusing.

The researchers conclude that collaborative testing is a "worthwhile technique." "For, even if overall test performance is not greatly improved by collaborative testing, the positive experiences of increased camaraderie and anxiety reduction it engenders could constitute substantial long-term benefits in the form of increased confidence, motivation, and willingness to continue one's education." (p. 173)

Conclusion: Despite these limitations, other benefits derive from this kind analysis. Most notably it generates rich conversations about critical thinking. It helps individual faculty, collections of faculty teaching related courses (in this case general education), and institutions clarify what they mean when they say they are teaching critical-thinking skills.

Reference:

Katherine Robertson is an associate professor of biology at Westminster College, New Wilmington, Pa.

Katz, S. N. (2010) Beyond Crude Measurement and Consumerism. Academe, 96 (5). Available online at www.aaup.org

Robertson, K. (2012) A Journal Club Workshop that Teaches Undergraduates a Systematic Method for Reading, Interpreting and Presenting Primary Literature. Journal of College Science Teaching, 41 (6) 20-26.

Stassen, M. L., Herrington, A., and Henderson, L. Defining Thinking in Higher Education. In Miller, J. E. and Groccia, J. E., eds. To Improve the Academy, 30. San Francisco: Jossey-Bass, 2011.

Wagner, T. The Global Achievement Gap: Why Even Our Best Schools Don't Teach the New Survival Skills Our Children Need &ndash and What We Can Do About It. New York: Basic Books, 2010.

Weimer, M. (2013) The Assessment Movement: Revisiting Faculty Resistance. Faculty Focus, January 18th. Available online at http://www.facultyfocus.com/articles/educational-assessment/the-assessment-movement-revisiting-faculty-resistance/