10

DEFINING CRITICAL THINKING IN HIGHER EDUCATION

DETERMINING ASSESSMENT FIT

Martha L. A. Stassen, Anne Herrington, Laura Henderson,
University of Massachusetts Amherst

Critical thinking is an important learning outcome for higher education, yet the definitions used on campuses and national assessment instruments vary. This article describes a mapping technique that faculty and administrators can use to evaluate the similarities and differences across these definitions. Results demonstrate that the definitions reflected by standardized tests are more narrowly construed than those of the campus and leave dimensions of critical thinking unassessed. This mapping process not only helps campuses make better-informed decisions regarding their responses to accountability pressures; it also provides a stimulus for rich, evidence-based discussions about teaching and learning priorities related to critical thinking.

Critical thinking has emerged as an essential higher education learning outcome for both external audiences focused on issues of accountability and for colleges and universities themselves. One of the most recent national efforts to respond to accountability pressures, the Voluntary System of Accountability (VSA), requires campuses to use one of three standardized tests to measure and report student learning gains on critical thinking and written communication (Voluntary System, 2010, para. 17). In its survey of employers, the Association of American Colleges and Universities (AAC&U, 2008) found that 73 percent of employers wanted colleges to "place more emphasis on critical thinking and analytic reasoning" (p. 16). In a recent survey of AAC&U member colleges and universities, 74 percent of respondents indicated that critical thinking was a core learning objective for the campus's general education program (AAC&U, 2009, p. 4).

While there is general agreement that critical thinking is important, there is less consensus, and often lack of clarity, about what exactly constitutes critical thinking. For example, in a California study, only 19 percent of faculty could give a clear explanation of critical thinking even though the vast majority (89 percent) indicated that they emphasize it (Paul, Elder, & Bartell, 1997). In their interviews with faculty at a private liberal arts college, Halx and Reybold (2005) explored instructors' perspectives of undergraduate thinking. While participants were "eager to promote critical thinking" (p. 300), the authors note that none had been specifically trained to do so. As a result, these instructors each developed their own distinct definition of critical thinking.

Perhaps this variability in critical thinking definitions is to be expected given the range of definitions available in the literature. 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 (Bean, 1996; Beyer, Gillmore, & Fisher, 2007; Brookfield, 1987; Donald, 2002; Facione, 1990; Foundation for Critical Thinking, 2009; Halx & Reybold, 2005; Kurfiss, 1988; Paul, Binker, Jensen, & Kreklau, 1990). Academic discipline can also shape critical thinking definitions, playing an important role in both the forms of critical thinking that faculty emphasize and the preferred teaching strategies used to support students' development of critical thinking capacities (Beyer et al., 2007; Huber & Morreale, 2002; Lattuca & Stark, 2009; Pace & Middendorf, 2004).

The Dilemma for Student Learning Outcomes Assessment

External accountability pressures increasingly focus on using standardized measures of student learning outcomes as comparable indicators of institutional effectiveness, and students' critical thinking performance is among the outcomes most often mentioned (see VSA, 2010, as an example). The range of critical thinking dimensions and the lack of one agreed-on definition pose a challenge for campuses working to align their course, program, and institution-wide priorities for critical thinking with appropriate national or standardized assessment methods. Among the questions facing these institutions are these three: (1) What dimensions of critical thinking do national and standardized methods emphasize? (2) To what extent do these dimensions reflect campus-based critical thinking instructional and curricular priorities? (3) What gaps in understanding students' critical thinking performance will we encounter when we use national or standardized tools?

Answers to these questions are important to any campus that wants to develop assessment strategies that accurately reflect teaching and learning priorities and practices on campus. A focus on the alignment of assessment tools with campus priorities is also essential for engaging faculty in the assessment decision-making process. It is unlikely that faculty will use evidence to inform changes in instructional practices and curricular design unless they have been involved in the assessment design and believe the tools and results accurately represent instructional priorities and practices.

Methods

To determine the alignment of current assessment tools with institutional instructional priorities, we conducted a qualitative content analysis of five representations of the critical thinking construct and identified the common and distinct dimensions across the five sources. The five sources used for this study represent two different contexts for defining critical thinking: an internal definition developed by a group of general education instructors on our campus and a number of external sources representing the primary tools currently under discussion for national assessments of critical thinking in higher education.

Internal Source

To represent our campus's operational definition of critical thinking, we use a definition developed by a group of general education instructors and administrators at a large public research university. The definition was developed as a part of a campuswide workshop on teaching critical thinking in general education and was generated by collecting the responses of groups of participants 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 . . ." Participant responses were then clustered by researchers in the campus's Office of Academic Planning and Assessment into twelve dimensions of critical thinking, listed in Table 10.1 in the "Results" section. A post-hoc confirmation of these dimensions was done by comparing the categories to the definitions of critical thinking present in the literature (see Office of Academic Planning, 2007, for the full set of responses and the links to the literature).

External Context

Critical thinking definitions from four external sources were used, which include three national standardized tests of critical thinking currently being used as a part of the VSA.

STANDARDIZED TESTS

ACT's Collegiate Assessment of Academic Proficiency (CAAP) comprises six independent test modules, of which the writing essays and critical thinking are relevant to this study. The critical thinking assessment is a forty-minute, thirty-two-item, multiple-choice test that, according to ACT, measures "students' skills in clarifying, analyzing, evaluating, and extending arguments" (ACT, 2011). The writing essays consist of two twenty-minute writing tasks, which include a short prompt that provides the test taker with a hypothetical situation and an audience.

The Collegiate Learning Assessment (CLA) is the Council for Aid to Education (CAE)'s testing instrument. Varying in length between ninety minutes (for the performance task) and seventy-five minutes (for the make-an-argument and critique-an-argument tasks, taken together), these written tests require students to work with realistic problems and analyze diverse written materials. CLA measures students' critical thinking skills with respect to analytic reasoning, problem solving, and effectiveness in writing. CLA is unique among the three standardized tests in its view of writing as integral to critical thinking.

Educational Testing Service (ETS) offers the Proficiency Profile (PP), a test of four skills, including reading and critical thinking. The PP is available in a standard form (two hours, 108 questions) and an abbreviated form accepted by VSA (forty minutes, 36 questions). Reading and critical thinking are measured together on a single proficiency scale.

NATIONAL ASSESSMENT TOOL

The fourth external source, the Valid Assessment of Learning in Undergraduate Education (VALUE) rubrics, is a set of scoring rubrics faculty or other reviewers can use to assess student work. The rubrics provide specific criteria for each of fifteen learning outcomes, two of which are relevant to this study: critical thinking, and inquiry and analysis (AAC&U, 2010a).

Three-Phase Content Analysis

Using these five sources, our research team conducted a three-phase content analysis.

PHASE ONE: IDENTIFYING THE DEFINITIONS

In order to compare our internal definitions with those of the external sources, we had to identify what aspects of critical thinking serve as the focus of each external assessment tool. We used a number of approaches to gather this information for the three standardized tests. To ascertain

each testing agency's working definition of critical thinking, we used the most detailed descriptions available, drawing from promotional materials, information on their websites, and communication with company representatives.

ACT (2010) describes the skills tested within each of three content categories: analysis of elements of an argument (seventeen to twenty-one questions, 53 to 66 percent of the test), evaluation of an argument (five to nine questions, 16 to 28 percent of the test), and extension of an argument (six questions, 9 percent of the test). Since it is not accessible through ACT's website, we obtained this document through a representative of ACT.

For ETS's PP, we selected passages from the User's Guide (Educational Testing Service, 2010): an introductory section that describes the abilities that the critical thinking questions measure and a more detailed description of the skills measured in the area of reading and critical thinking at the intermediate and high proficiency levels.

For the CLA, we began with the skills contained in the CLA Common Scoring Rubric (Council for Aid to Education, 2008). This rubric is divided into two categories: (1) critical thinking, analytic reasoning, and problem solving and (2) written communication. In spring 2010, we learned that CAE was in the process of implementing new critical thinking rubrics: analytic reasoning and evaluation, problem solving, and writing effectiveness. We analyzed these new descriptions (CLA II) alongside the older rubric (CLA I). In fall 2010, after the research described here was completed, CAE published a more detailed version of the critical thinking scoring rubric that is now available on its website. While differently formatted, the descriptors we used for this analysis are similar to the categories in this new rubric.

We were able to use the actual measures in the VALUE rubrics because they are the components of the rubric used to review and assess students' work (AAC&U, 2010b). We incorporated both the critical thinking and the inquiry and analysis rubrics in our analysis. We chose to include them because this category seemed particularly relevant to the conceptualization of critical thinking emerging from our campus discussions.

PHASE TWO: CODING FOR COMMONALITIES WITH CAMPUS CRITICAL THINKING DEFINITION

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. The research team reviewed each descriptor of the four external source definitions and coded each for its alignment with one or more of the twelve dimensions of our internal definition. For example, the CLA listed "constructing cogent arguments rooted in data/information rather than speculation/opinion" (Council for Aid to Education, 2008) as one descriptor of their critical thinking/writing effectiveness definition. In our analysis, we coded this descriptor as falling into the judgment/argument dimension of the campus-based definition. In conducting this coding, we used two approaches. First, to develop common understandings of the process, we worked as a team (three coders) to code two of the external sources (CAAP and PP). We then individually coded the CLA and VALUE sources and met to confirm our coding. In both approaches, we identified areas of disagreement and worked together for clarity in our standards, coming to mutually agreed-on final codes.

Once the coding was completed, we sorted the individual descriptors by dimension and reviewed them again for consistency. For example, we checked to see if the items we had coded as evidence-based thinking all reflected our deepening understanding of the construct. This stage helped us further clarify distinctions among the dimensions.

PHASE THREE: ANALYSIS OF PATTERNS

Once the coding and checking were complete, we arrayed the results in a table to facilitate a comparative analysis. We calculated how many of each tool's descriptors referenced each of the twelve dimensions in our campus definition and, to get a sense of the relative emphasis each tool gave to each of the twelve dimensions, we calculated the proportion of all descriptors listed that reflect each dimension. In this way, we denote what proportion of each tool's definition reflects each of the twelve campus-based critical thinking dimensions.

Results

Table 10.1 summarizes the commonalities and gaps among the various definitions. This table indicates how many of the critical thinking dimensions listed in each of the external assessment tools reflect each of the twelve campus critical thinking dimensions. To provide a very rough estimate of the relative emphasis or importance of these dimensions in our campus definition, we counted how many descriptors emerged in the workshop for each dimension and calculated the proportion of all descriptors that this dimension represents (under the assumption that the number of descriptors of a dimension generated by a group of faculty

Table 10.1 Relationships Between Campus Critical Thinking
Definitions and Four External Sources

Campus-Based Definition	Campus		Collegiate Learning Assessment I		Collegiate Learning Assessment II		Proficiency Profile		Collegiate Assessment of Academic Proficiency		Value	
	N	%	N	%	N	%	N	%	N	%	N	%
Judgment/ argument	8	15	24	80	6	55	10	56	8	73	6	55
Synthesizing	6	12	2	7	2	18	2	11	0	0	4	36
Problem solving	2	4	1	3	1	9	1	6	0	0	0	0
Evidence-based thinking	3	6	7	23	3	27	5	28	6	55	1	9
Drawing inferences	2	4	1	3	2	18	3	17	3	27	2	18
Perspective taking	7	14	3	10	1	9	0	0	1	9	3	27
Suspend judgment	1	2	0	0	0	0	0	0	0	0	0	0
Application	10	19	0	0	0	0	0	0	0	0	0	0
Metacognition	5	10	0	0	0	0	0	0	0	0	1	9
Questioning/ skepticism	4	8	0	0	0	0	0	0	0	0	1	9
Knowledge/ understanding	3	6	0	0	0	0	1	6	0	0	1	9
Discipline-based thinking	1	2	0	0	0	0	0	0	1	9	2	18
Total items in definition	52	100	30	100	11	100	18	100	11	100	11	100

Note. Percentages provided as an indicator of relative importance or emphasis of each construct across sources. Numbers represent duplicate counts across categories—one item in a list can reflect more than one campus-related CT category. The "total" numbers in the bottom row of the table reflect the unduplicated count of the descriptors.

reflects greater centrality or emphasis for this dimension). Note that looking at the campus definition this way highlights the relative emphasis (10 percent or more of the descriptors) placed on five dimensions of critical thinking: judgment/argument, synthesis, perspective taking, application (representing the most emphasis with 19 percent of the descriptors reflecting this particular dimension), and metacognition. We followed the same method to determine the relative emphasis of each dimension in the external assessment tools. Looking at the CLA I column as an example, we found that twenty-four of the thirty descriptors listed in the CLA I definition of critical thinking reflect our campus's construct of judgment/argument. These twenty-four occurrences represent 80 percent of the dimensions in the CLA I list.

As the results in Table 10.1 illustrate, 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). For the three standardized tests and VALUE, there is also a substantial emphasis on drawing inferences. Evidence-based thinking is emphasized in all three standardized tests. To varying degrees, synthesizing, problem solving, and perspective taking also receive some attention from the external sources.

In our analysis, a number of the campus dimensions receive no attention from any of the standardized tests: application, suspending judgment, metacognition, and questioning/skepticism. Of those that are missing from the standardized tests, the VALUE rubrics do reflect metacognition and questioning/skepticism.

The results suggest differences among the four external sources. The CAAP appears the most focused or limited in scope, with primary emphasis on judgment/argument, use of evidence, and drawing inferences. The VALUE rubrics are the most expansive, with references to nine of the twelve dimensions from the campus-based definition. Two of the three dimensions that are not included, problem solving and integrative and applied learning, are actually present as separate VALUE rubrics (AAC&U, 2010b), so their absence from the rubrics used in this analysis is not surprising.

In addition to providing us with one perspective on the relationship between the four external assessment tools and our campus's critical thinking definition, this analysis also provided us with the opportunity to revisit the campus definition. Our analysis helped us clarify a number of our dimensions in relationship to the four external sources. For example, our category of multiple perspectives/perspective taking emerged as

the dimension where we coded all external source descriptions that referenced "dealing with complexity" in addition to items that indicated "addressing various perspectives." As we coded, we also noted that our campus descriptions of perspective taking tended toward the positive dimension of multiple perspectives (that is, taking into account these perspectives) but did not include more critical aspects of this dimension (that is, critiquing or refuting a perspective that is weak or uninformed, "considering *and possibly refuting* [italics added] alternative viewpoints" [CLA II]). We also were made aware of dimensions of critical thinking present in the external sources that are not present in the campus definition.

Limitations

It is important to acknowledge that this analysis is not a study of test item validity. Instead, it focuses on how the basic construct of critical thinking is defined, and the dimensions emphasized, within both contexts and across the five sources. Obviously these definitions and emphases drive item development and have important implications for the appropriateness of each assessment tool as an indicator of institutional effectiveness as measured by students' critical thinking performance. However, the technique we used to determine definitional emphases is limited.

The limitations fall into two categories: those having to do with the campus definition and those having to do with the sources used for the external definitions. First, to represent the campus definition, we used the results of a collaborative brainstorming session conducted as part of a campuswide workshop on critical thinking in general education. The definition that emerged is multidimensional, and the elements correspond to common elements of critical thinking as defined in various sources in the literature. However, the campus definition has not been systematically vetted or tested against the responses of other groups of faculty, so it is still very much an emerging document on our campus. Still, many of the dimensions of this definition are identified in other faculty-generated statements of general education learning objectives, including the learning objectives for the campus's junior-year writing requirement (an upper-division writing requirement that addresses the writing conventions of the student's major) and the results of a survey where instructors report emphasizing these objectives in their general education courses (Office of Academic Planning and Assessment, 2008). The definition also does not definitively reflect the faculty's beliefs about the relative importance of each of these constructs. We used the number

of references as a rough indicator of importance, but this is certainly not a systematically tested assumption.

With respect to the external sources, the characteristics we used for the three standardized tools (CLA I and II, PP, and CAAP) come from each test company's description of the critical thinking components covered in their test. We took these descriptors and coded each against our campus-based definition. Because we do not know the relative emphasis on each component in the test itself (that is, the number of test items, or scoring weights, for each item), we considered each descriptor of equal importance and looked at the number of them that reflect each of our campus categories. Once they were coded, we then looked to see what proportion of the items reflect each campus construct. While we believe this was the most appropriate step to take given the available information, it may misrepresent the actual emphasis of the test. Conducting a more finely tuned analysis would require us to look at the actual tests and, for those with open-response items, the evaluative rubrics and weights used. This, of course, is an analysis of even greater complexity, requiring us to address proprietary constraints with the testing companies. The public relations representation of the test substance is the information most academics would use to make such determinations, so we felt it was a relevant source to use and dissect.

Discussion

We set out to understand the relationship between our campus's emerging definition of critical thinking and the definitions used by four external tools for assessing students' 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. The analysis process also ended up challenging us to clarify our own expectations for student performance and assessment. Finally, this research offers an analytical and evidence-based process for engaging faculty in reviewing teaching and learning priorities within the context of responding to external accountability demands.

Focusing first on the issue of fit between the four external sources and our campus definition, 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. The VALUE rubrics provide more comprehensive coverage of the campus definitions, touching on nine of the twelve dimensions. Two that are not included (application and problem solving) are

referenced in separate VALUE rubrics, which could be used to address the fuller range of campus dimensions.

These results help inform the campus discussion of which assessment options would be most appropriate. For example, if the faculty on our campus determine that judgment/argument is appropriate as the focus of our externally driven assessment, then any of the standardized tests might be acceptable. But if we decide we want our assessment strategy to reflect more of the dimensions of critical thinking present in the campus definition, the VALUE rubrics might be a better choice but would not necessarily reflect the same relative importance of these constructs as emerged from our faculty workshop results. This discrepancy could be remedied in part by including the integrative and applied learning VALUE rubric to the assessment since it would address the dimension that received the most attention from faculty application.

It should be noted, however, that selecting the VALUE rubric tool would not be sufficient for fulfilling the current VSA requirements for a standardized assessment method. VALUE rubrics also require more faculty time and expertise than standardized tests since rubrics require raters to be trained and then to assess samples of student work. The standardized tests have other costs (testing fees, incentives for the students, and staff effort in recruiting respondents) that, if used for VALUE analysis instead, would defray the costs described above. Clearly, associated costs also need to be a part of the campus's decision-making process.

Our analysis has raised another essential question that the faculty need to address: What sort of evidence of students' critical thinking is appropriate? The various descriptors of critical thinking used in these five sources (both the internal and the external sources) suggest the different kinds of performance tasks being used. The PP and CAAP rely on multiple-choice tasks—and their descriptors reflect identifying and recognizing aspects of an argument—for example, "identify accurate summaries of a passage" (Educational Testing Service, 2010) and "distinguish between rhetoric and argumentation" (ACT, 2010). The CLA, on the other hand, requires students to craft an argument. The CLA definition uses descriptors that reference creating an argument—for example, "constructing organized and logically cohesive arguments," "considering the implications of decisions and suggesting additional research when appropriate" (Council for Aid to Education, 2010). In this test, however, the parameters of student-generated responses are limited in scope. Students write answers to a set of narrowly focused prompts that address specific elements of the task and evidence presented.

The VALUE rubrics were designed specifically to assess portfolios of students' work from their courses—tasks that would be varied in focus, content, and types of writing contexts. The items in these rubrics reflect the comprehensiveness of these types of student work, referencing contextual analyses, identifying and describing a problem, and articulating the limits of one's position. Students' responses in this case would be unconstrained, reflecting the variety of ways one demonstrates a range of critical thinking dimensions across an array of courses and assignments.

Finally, our campus definition came from the discussions of a diverse group of instructors who responded to the prompt they were given by, quite naturally, thinking about the evidence of critical thinking they see in the assignments and tasks they ask of their students. Therefore, their responses focus to a larger degree on the doing: the creation of arguments, the application of theory to new settings, and the identification of evidence to support those arguments or assertions. The focus of these faculty-derived definitions, based as they are on what students are actually asked to do in the classroom, seems particularly distant from the tasks associated with the standardized multiple-choice tests that focus more on identifying and selecting over creating and constructing.

Another complexity emerges that is particularly relevant to assessment methods that use open-ended or constructed responses that are scored by sources outside the control of the faculty or the campus (like the CLA tasks and the CAAP and CLA essays). In these cases, it is important to make a distinction between what the assessment task is and what actually gets scored for performance assessment purposes. For example, the CLA task certainly seems to qualify as representing critical thinking application since it asks students to apply their analysis of various sources of information to a real-world question. It is therefore interesting that in our analysis, we did not find evidence of application in the CLA critical thinking definition—the elements of critical thinking they say their test addresses. Instead, their critical thinking descriptors focus primarily on judgment/argument, evidence-based thinking, synthesizing, and drawing inferences (CLA II).

Without more specific information about how the constructed responses are actually scored (that is, what elements of performance actually count) it is unclear whether application, for example, is actually a performance factor that is assessed or only the frame through which the performance of interest is stimulated. For example, is the student's capacity to judge the relevance of evidence to a particular context scored, or is the focus on being able to make the distinction between correlation and causation? Both would be a reflection of evidence-based thinking.

However, the first would also be a more complex or advanced form of critical thinking that reflects application. The second reflects a somewhat more basic but still important component of evidence-based thinking but would not reflect application as we have conceived it in our campus definition. This is an important point in reminding ourselves that the assessment task itself is only one component of the consideration of fit. When student performance is scored by parties removed from the campus context, it is also particularly important to be clear about what elements of student performance are included in the final score.

The importance of taking account of the types of tasks and the scoring criteria is illustrated in a recent study conducted by the University of Cincinnati and highlighted in an AAC&U publication (AAC&U, 2010a). Researchers compared first-year students' performance on the CLA with those students' performance on an e-portfolio assignment, assessed by faculty at the university using a slightly modified version of the VALUE rubrics. Researchers found no significant correlation between the two sets of assessment results, suggesting that the two assessment tools capture very different elements of students' critical thinking performance. These results raise an important question for campuses to consider: Does our assessment strategy capture the kind of student learning and performance we emphasize and value? Tools that do not effectively measure what matters to faculty are not appropriate sources of evidence for promoting change or for accurately reflecting instructional and curricular effectiveness.

Connecting Research and Practice: A Note to Faculty Developers

Finally, and perhaps most important, this method of inquiry leads to productive and engaging faculty discussions of critical thinking teaching, learning, and assessment. This project illustrates a way to address external accountability pressures while also generating joint faculty and administration discussions and insights into campus-based teaching and learning priorities. The first example of this productive inquiry was the workshop activity that produced the cross-disciplinary definition of critical thinking for our campus. Having this definition in place made it possible to pursue the line of inquiry described here, which served as an essential starting point for our campus's consideration of how to assess critical thinking in ways that are internally valid and externally legitimate.

The exercise of mapping our critical thinking dimensions against the definitions of the four assessment tools sparked a rich discussion among the coders. It was as we tried to code the external definitions using our internal critical thinking categories that we began to clarify the meaning of our own definition and see both the gaps and strengths of that definition. During this process, we also discovered the essential links between our definition and our faculty's pedagogical values in facilitating students' critical thinking. We believe that workshops that provide groups of faculty and administrators the opportunity to conduct this kind of analysis together can generate an important evidence-based dialogue about expectations for student learning, the assessment tools that most appropriately reflect those expectations, and the trade-offs inherent in making those kinds of decisions. The coding process opens up a conversation about what we mean when we use the term *critical thinking,* a process of clarification that informs one's own teaching as well as the larger campus conversation about critical thinking assessment.

REFERENCES

ACT. (2010). *CAAP critical thinking content categories and subskills.* Iowa City, IA: Author.

ACT. (2011). *Critical Thinking Test.* Retrieved from www.act.org/caap/test_thinking.html

Association of American Colleges and Universities. (2008). *Our students' best work: A framework for accountability worthy of our mission* (2nd ed.). Retrieved from www.aacu.org/publications/pdfs/studentsbestreport.pdf

Association of American Colleges and Universities. (2009). *Learning and assessment: Trends in undergraduate education.* Retrieved from www.aacu.org/membership/documents/2009MemberSurvey_Part1.pdf

Association of American Colleges and Universities. (2010a). Assessing learning outcomes at the University of Cincinnati: Comparing rubric assessments to standardized tests. *AAC&U News.* Retrieved from www.aacu.org/aacu_news/AACUNews10/April10/

Association of American Colleges and Universities. (2010b). *VALUE: Valid assessment of learning in undergraduate education.* Retrieved from www.aacu.org/value/rubrics/index.cfm

Bean, J. C. (1996). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom.* San Francisco, CA: Jossey-Bass.

Beyer, C. H., Gillmore, G. M., & Fisher, A. T. (2007). *Inside the undergraduate experience: The University of Washington's study of undergraduate education.* San Francisco, CA: Jossey-Bass.

Brookfield, S. D. (1987). *Developing critical thinkers: Challenging adults to explore alternative ways of thinking and acting.* San Francisco, CA: Jossey-Bass.

Council for Aid to Education. (2008). *Common scoring rubric.* Retrieved from www.cae.org/content/pdf/CLA_Scoring_Criteria_%28Jan%202008%29.pdf

Council for Aid to Education. (2010). *CLA scoring criteria.* Retrieved from www.collegiatelearningassessment.org/files/CLAScoringCriteria.pdf

Donald, J. G. (2002). *Learning to think: Disciplinary perspectives.* San Francisco, CA: Jossey-Bass.

Educational Testing Service. (2010). *ETS Proficiency Profile user's guide.* Retrieved from www.ets.org/s/proficiencyprofile/pdf/Users_Guide.pdf

Facione, P. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction* [Executive summary]. Retrieved from ERIC database. (ED315423)

Foundation for Critical Thinking. (2009). *Our concept of critical thinking.* Retrieved from www.criticalthinking.org/aboutCT/ourconceptCT.cfm

Halx, M. D., & Reybold, L. E. (2005). A pedagogy of force: Faculty perspectives of critical thinking capacity in undergraduate students. *JGE: The Journal of General Education, 54*(4), 293–315. doi:10.1353/jge.2006.0009

Huber, M. T., & Morreale, S. P. (Eds.). (2002). *Disciplinary styles in the scholarship of teaching and learning: Exploring common ground.* Sterling, VA: Stylus.

Kurfiss, J. G. (1988). *Critical thinking: Theory, research, practice and possibilities* (ASHE-ERIC Higher Education Report No. 2). Retrieved from ERIC database. (ED304041)

Lattuca, L. R., & Stark, J. S. (2009). *Shaping the college curriculum: Academic plans in context* (2nd ed.). San Francisco, CA: Jossey-Bass.

Office of Academic Planning and Assessment, University of Massachusetts Amherst. (2007). *Defining critical thinking: Participant responses.* Retrieved from www.umass.edu/oapa/oapa/publications/gen_ed/critical_thinking_definitions.pdf

Office of Academic Planning and Assessment, University of Massachusetts Amherst. (2008). *Gen ed curriculum mapping: Learning objectives by gen ed course designations.* Retrieved from www.umass.edu/oapa/oapa/publications/gen_ed/instructor_survey_results.pdf

Pace, D., & Middendorf, J. (2004). Decoding the disciplines: A model for helping students learn disciplinary ways of thinking. In D. Pace & J. Middendorf (Eds.), *New directions for teaching and learning: No. 98. Decoding the disciplines: Helping students learn disciplinary ways of thinking* (pp. 1–12). San Francisco, CA: Jossey-Bass.

Paul, R., Binker, A., Jensen, K., & Kreklau, H. (1990). *Strategy list: 35 dimensions of critical thought.* Retrieved from www.criticalthinking.org/page.cfm?PageID=466&CategoryID=63

Paul, R., Elder, L., & Bartell, T. (1997). *Study of 38 public universities and 28 private universities to determine faculty emphasis on critical thinking in instruction.* Retrieved from www.criticalthinking.org/research/Abstract-RPAUL-38public.cfm

Voluntary System of Accountability Program. (2010). *Participation agreement.* Retrieved from www.voluntarysystem.org/docs/SignUp/VSAParticipationAgreement.pdf