GENERAL EDUCATION ASSESSMENT WITHIN THE DISCIPLINES

Douglas J. Eder

Most of the people I sit next to on airplanes are not college professors. Usually they are business people, and some of them are top executives with the interest, power, and skill to recruit and hire well. As the steward of my university’s assessment program, I enjoy inquiring from such people what they seek in their new employees—that is, I seek to assess their employment goals—mostly with the hope of helping to inform my own university about external expectations, and maybe even to improve its performance with respect to them. When my airplane seatmate discovers that I am not the expected business professional but, rather, an academic, s/he often launches the conversation with a variation of this loaded question: “So, what do you really think of higher education today?” I smile and reply, “Actually, I’m in favor of it.”

We pass lightly through such sallies and get to where we can explore this question: “What traits do you look for in a top-notch prospective employee?” One recent executive seatmate said that the excellent prospect:

- is self-motivated. High quality training is expensive and the self-motivated employee learns independently so I don’t have to spend excessive money getting him or her acquainted with our techniques;
- practices professional and personal etiquette and can represent my company with competence and style;
- demonstrates a capacity to analyze and solve problems effectively;
- writes well and can explain a technical or complex topic briefly and in jargon-free language so I can make a business decision;
speaks well in one-on-one conversations and in front of large groups;
• is ethical and dependable.

“What about competence in a major discipline?” I ask.

“Oh, that’s assumed. I thought you wanted to know what characterizes the top candidates.”

The lists that my various airplane acquaintances have given me are not new. Similar lists have been recited for years (Chickering, 1994; Hersh, 1994; Tucker, 1995). What is interesting is that when doing assessment seminar/workshops, I ask the professors early in the session what they would like their students to “look like” when they walk out the doors of their classes—or their universities upon graduation. I do this on the premise that we should construct an assessment plan around what matters most, and these professors are being asked, “What matters most to you for the learning of your students?” The professors invariably produce a list describing a high-quality graduate who:

• is a lifelong learner and has abiding curiosity;
• possesses self-confidence and a sense of personal achievement;
• exhibits and employs critical thinking;
• is competent in written communication;
• is competent in oral communication;
• has a sense of ethics and responsibility;
• knows something in some detail about a subject area.

What is remarkable about this list, generated by academic solons, is that it is roughly the same as the list created by corporate executives. That is, when professors get together outside the boundaries of academic departments and attempt to characterize the newly educated individual, they tend to recite from the heart those qualities known more-or-less as the “transferable baccalaureate skills” or the “goals of general education,” and these qualities are the same ones valued highly in corporate organizations and by society at large (Chickering, 1994; Jones, 1997; Association of American Colleges and Universities [AAC&U], 2002). Moreover, the importance of content knowledge in a discipline is often not mentioned in the
professors’ lists. When it does arise in either the corporate or the academic conversation, competence in a major discipline is mentioned as a necessary, but not sufficient, quality. Stated another way, a high grade point average in a major may get an application into the stack, but by itself it won’t get the applicant hired. In our hearts, we professors and our corporate colleagues value the properties of a general education.

Valuing something with the heart does not always translate into doing something with the head. Universities have generally arranged themselves into units that reflect academic disciplines, their attendant cultures, and their specialized ways of knowing. The resulting department-based structures serve the disciplines well, but they exhibit the Tragedy of the Commons when encountering an issue that involves them all: providing general education. To wit, everybody has an opinion about it but nobody owns it enough to take responsibility for it. Thus, effectively “assessing general education” has lagged behind assessing achievement in the disciplines. Moreover, some cross-cutting concepts embedded in general education, such as lifelong learning, aesthetic awareness, ethical behavior, and critical thinking, have acquired an aura of being ineffable and, correspondingly, unassessable.

Two Definitions

Before proceeding further, a couple of distinguishing definitions are in order. First is the distinction between direct and indirect assessments. Direct assessments acquire evidence about student learning and the learning environment. Examples are exams, projects, logs, portfolios, and direct observations. Indirect assessments acquire evidence about how students feel about learning and their learning environment. Examples include surveys, questionnaires, interviews, focus groups, and reflective essays. Relatively speaking, universities have made more progress in constructing direct assessments of student learning within the disciplines than they have in doing the same for general education. For the most part, faculties have left the business of direct assessment of general education to commercial enterprises. However, examples abound of educational institutions
invoking indirect assessments of general education in support of their statements about student learning.

The second definition involves deep learning, also called expert learning. The 1990s were known in neuroscience circles as “The Decade of the Brain.” Federal agencies dedicated extra money and scientists dedicated special effort to studying how the brain and mind operate. Within cognitive neuroscience, two major findings are relevant here: (1) The more of the brain—sensory, motor, and association pathways—involved during a learning process, the more effective is the learning; and (2) deep learning, which is the sort of learning that, through its cumulative processes and contents, transforms a novice into an expert, is a curriculum function, not a course function (National Research Council, 2000). The implications of these two findings have yet to really penetrate the teaching and learning of general education. Taken together, these two findings suggest strongly that deep learning—the kind of learning that faculty members have and wish their students to acquire—is effectively procured through active learning reinforced over an extended time.

Universities are systems of formally scheduled courses, residence halls, dining services, athletic events, libraries, informal conversations, and more. Deep learning occurs when students participate in a learning system. A full-time student spends roughly 16 hours, or 10% of each week, formally “in class.” How students use the other 90% of their time is significant, because that’s when real learning can occur. Therefore, a general education program that seeks to produce magnificent gains in learning will seek to occupy some of that time by actively engaging students outside of class throughout their collegiate careers … and it will assess learning effectiveness accordingly.

Numerous attempts to assess general education have assumed the perspective that general education is a separate entity that is distinct from education in a discipline. The consequences of such unhitched teaching and learning have been summarized eloquently (AAC&U, 2002, p. 16). General education is not isolated from the disciplines (Weeks, 1996). It is the discipline-associated situation that governs the specifics of what are regarded, at a minimum, as good writing, speaking, and problem solving. For example, good, specific writing by students in biology, as in other disciplines,
requires a new vocabulary, new models, and a new attention to the needs of the specific reader. Students do not learn how to write only in English 101 and 102, introductory composition courses, no matter how high the grades they achieve there. Students learn to write by practicing and receiving feedback across the whole of their collegiate experiences. Therefore, assessment of general education for deep learning will focus on assessing the system, not individual courses. One way to assess general education as a system is, of course, within the academic disciplines.

The preceding reasoning suggests that assessing general education is probably not so much a challenge of assessing the ineffable as assessing the inchoate. In contrast to ineffable, which means “indescribable,” the word inchoate comes from the Latin cohum, meaning a strap used to hitch an item like a plow to the ox that pulls it. Therefore, inchoate means “not hitched up.” A tale told by Winthrop University’s Margaret Tebo-Messina (Tebo-Messina, D’Amato, Kitchings, Krech, & Underwood, 1995) may serve well here. Several years ago, Winthrop wished to assess aesthetic awareness as a learning outcome of its general education efforts. Not possessing the answer, Tebo-Messina carefully picked a team of six faculty members and sent them to an assessment conference with the charge, “Figure out a way to assess aesthetic awareness.” The team included an architect, a musician, a poet, a historian, a painter, and a thespian. Several days later the team returned with an answer. “Regardless of our disciplines, we agree that one indicator of aesthetic awareness would be the capacity to interpret metaphor.” Consequently, Winthrop realized that its general education students all took a course in literature, theater, or art wherein they would encounter, respectively, something on the order of Coleridge’s poem, *Rime of the Ancient Mariner*, a performance of Tchaikovsky’s ballet, *Swan Lake*, or a copy of Rodin’s sculpture, *The Thinker*. The assessment task then became simply one of tapping into already-existing assignment opportunities that ask students to interpret the metaphors inherent in these media. The lesson provided by Winthrop University is that assessment of general education attributes, such as aesthetic awareness, is not indescribable but, rather, merely not hitched up to appropriate assignments.
Authentic Assessment

Metaphors are useful for thinking about general education assessment as well as actually doing it. Consider again the airplane. Suppose it became known to you that pilots, including the one in the last commercial flight you took, become qualified solely on the basis of four written exams: airframe mechanics, flight dynamics, meteorology, and navigation. Qualifying standards are high; the minimum passing score is 95%. Nevertheless, how secure do you feel about flying with one of these pilots? It probably won’t escape your attention that there is no application of knowledge, no real take-offs and landings, no real problem solving in a real airplane, and so you are understandably nervous about flying. In actuality, real pilots get experience in real airplanes. They learn to solve problems, including some that are too dangerous to learn about in real airplanes. For this, they use cockpit flight simulators. Pilots can practice different applications of knowledge, analyze choices, and improve solutions to problems. When students err and crash in a simulator, nobody gets hurt. The instructor simply pushes the reset button and the simulation begins anew. Authentic assessment is assessment of a real task in a real or simulated environment under supervision. Authentic assessment has a long history in aviation, medicine, counseling, nuclear engineering, law, and other professional disciplines. From an educational viewpoint, a positive aspect of the archtypical flight simulator is that it combines teaching, learning, and assessing into a real task. Because the assessment is embedded in teaching and learning, it is largely transparent for both the student and the instructor.

What would be nice would be to have an academic equivalent of the flight simulator so that assessment could be embedded transparently in collegiate teaching and learning. That is, it would be nice to have, in the words of Steve Ehrmann (1998), an academic simulator. The academic simulator exists. It can be defined as a reflective, scholarly engagement between student and dedicated professor that results in a visible product or behavior. Because the product or behavior is visible, it serves as an assessable performance indicator for the curriculum that produced it. The task for the student can range from a single class assignment
to a multifaceted senior project involving a team. When using authentic assessment with the academic simulator, the closer any simulation is to the real thing, the better is the appraisal. Key to good assessment, of course, is picking a good indicator for the academic simulator to produce.

Assessment Indicators and Principles of Good (Effective) Assessment Practice

The indicator concept is important in assessment. It is usually impossible actually to measure learning because what happens in the mind is fundamentally invisible. How does one know if a person is a “reflective critical thinker” or possesses “global awareness”? Truly, we cannot look at our students’ faces and “see it in their eyes.” Students have to construct something or behave overtly in order to make their learning visible. Therefore, one can monitor student learning through an effective indicator, which is a product of student learning—an artifact yielded up by an assignment—or a student behavior. The product or behavior serves as evidence that what is expected to be learned is actually being learned. To return to a previous example, how would a college know whether its students were actually becoming aesthetically aware? Well, one might design an assignment in which students actually had to interpret metaphor. Whether the activity is interpreting the choreography of Swan Lake (music appreciation), explicating the meanings suggested by The Thinker (art appreciation), or unraveling the implications of The Rime of the Ancient Mariner (literature appreciation), interpretation of metaphor is a performance indicator for student attainment of aesthetic awareness. Many other indicators exist, of course. Thus, a performance indicator is an item of evidence that indicates the extent to which a curriculum is functioning toward achieving its intended goals.

The concept of assessment indicators has proved to be both fruitful and elusive. “To be useful as an indicator, a particular piece of data must communicate something important about what is happening in a complex domain” (Ewell & Jones, 1996). The fruitful part of indicators is that they are the more accessible
proxies for something that is less accessible, but necessary to assess. So, if achievement of aesthetic awareness as a learning outcome is overly difficult to assess directly, one can use interpretation of a metaphor as an indicator of that outcome, as described above. The elusive part is that indicators are never perfect. They may not completely align with the desired outcome. Consequently, many faculty members expend great time and energy attempting to discover or create ideal indicators. Other professors abandon the project in the belief that outcomes in some fields (often their own) simply can’t be assessed. Even worse, students may discover superficial formulas that allow them to mimic the desired outcome, especially if the indicator itself is superficially related to the outcome. Hence, the maxim applied to social science research applies to assessment as well: Use multiple measures over time.

Principles of good, effective assessment do exist and they are the same whether they are applied to the disciplines, to general education, to web-assisted education, or to true distance education. These principles have been used successfully as the foundation for James Ratcliff’s (1995) “Ten Steps in Developing an Assessment Plan,” and James Nichols’s (1995) “Five Column Assessment.” The Six Principles outlined below and used as an outline in this paper are derived from those of Glassick, Huber, and Maeroff (1997), who suggested their use as a foundation for pursuing scholarly questions across all disciplines.

1. Establish clear goals—we identify the question.
2. Ensure adequate preparation—we see the task appropriately.
3. Employ appropriate methods—we can do the task.
4. Produce significant results—we do care about both products and processes.
5. Arrange for effective presentation—we can escape from the tyranny of external demands.
6. Practice reflective critique—we can improve.

We’ll build a general education assessment process by following this outline one step at a time.
Establish Clear Goals

Early in his many presentations on assessing general education, Peter Ewell asks his audience, “What are the assumptions and primary goals of general education at your institution?” It is difficult to overstate the importance of well-constructed goals. In the history of assessment since 1985, numerous institutions have succumbed to the beguiling temptation of beginning with methods rather than with goals: “Here is a wonderful method that we’ve purchased (or constructed). Let’s use it to assess something.” After spending some combination of energy, time, and money, an institution discovers, first, that it hasn’t learned anything useful and, second, that the faculty thinks (correctly here) that assessment is a waste. There are several graphic ways for demonstrating the importance of goals. Here is one of them: When we were younger and played with mazes, what was the quickest, easiest, and most efficient way to solve the maze? Of course: Go to the goal in the middle and work backward. The reason this works is because running the maze backward eliminates the wrong turns. The same process solves any “maze,” including a roadmap. For example, if one wishes most efficiently to find the best driving route from Tallahassee, Florida, to Calgary, Alberta, one runs one’s finger backwards from Calgary back towards Tallahassee. Conversely, if one wishes to navigate from Calgary to Tallahassee, one plots the route most efficiently by working backward from Tallahassee toward Calgary. There is a complex mathematics called game theory behind this kind of thinking, part of an idea developed by Nobel Prize recipient John Nash and popularized by the movie, *A Beautiful Mind*.

Faculties at institutions of higher learning have not found it difficult to enumerate goals for general education. A Web search through Google.com on the string, “general education” + goals yielded over 250,000 listings. Sites uncovered this way, plus statements in a selection of college and university catalogs, are more impressive in their convergence than in their divergence. In most respects, they duplicate and amplify those listed at the beginning of this essay: Students should learn to be critical thinkers; students should possess curiosity sufficient to become lifelong learners; students are expected to develop into skilled writers and speakers; students should possess a sense of ethics, and so on.
The discovery that goals are central to choosing a correct pathway did not await a Nobel Prize in mathematics:

“Cheshire Puss …,” asked Alice, “would you tell me, please, which way I ought to go from here?”

“That depends a good deal on where you want to get to,” said the Cat.

“I don’t much care where—” said Alice.

“Then it doesn’t matter which way you go,” said the Cat. (Carroll, 1994)

Even when it seems counterintuitive, it is useful and efficient to approach general education assessment by starting from the goals.

**Ensure Adequate Preparation**

Good assessment is akin to good diagnosis. One doesn’t prescribe an intervention before one has accurately diagnosed a condition. More than anything else, therefore, good assessment asks good questions. This is important because bad questions take just as much energy to ask (and answer) as good questions do. When one asks a good assessment question, one asks about evidence, which must be in the form of an artifact or a behavior. To be useful as an indicator of learning, the evidence must be visible. This comes in the form of things either produced or modeled by students, and to generate these indicators, students must *do*, not just *be*. The general education goals so richly represented in catalogs and web pages describe what faculties want students to *be*. Unfortunately, goals themselves are largely invisible and, hence, unassessable. How can a professor tell, for example, the extent to which a student *is* a good critical thinker, *is* a creative problem solver, *is* a perceptive writer, or *is* an ethical decision maker? Reliable answers do not come from, as some pedagogues say only partly in jest, “I can see it in their eyes.” Answers do come from actual student products and actions, which are
independently visible. To make evident their achievement of the aforementioned goals, students have to perform an analysis, solve a problem, write something, or make a decision involving ethics. Sometimes the evidence is in a form that can be judged only by a specialist in the discipline. Other times it can be judged by a lay person. In all cases, however, the evidence is visible. The process of making goals visible means translating them faithfully from being into doing. If goals express what we want our students to be, then objectives describe what we want our students to do. Objectives are indicators of goals.

Table 1: Assessment Use of Bloom’s Taxonomy

Bloom’s Taxonomy is a well-known description of levels of learning. A taxonomy such as this one may be a useful guide when defining or examining learning objectives for assessment. It is also useful for content analysis of such things as the seemingly simple Minute Paper.

<table>
<thead>
<tr>
<th>Level</th>
<th>Some Cognitive Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation</td>
<td>Appraisal of own or someone else’s Analysis or Synthesis</td>
</tr>
<tr>
<td></td>
<td>Exam question at this level: Evaluate another physical therapist’s program to strengthen the rotator cuff.</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Assembly of Application</td>
</tr>
<tr>
<td></td>
<td>Exam question at this level: Design a physical therapy program to strengthen each component of the rotator</td>
</tr>
<tr>
<td>Analysis</td>
<td>Disassembly of Application</td>
</tr>
<tr>
<td></td>
<td>Exam question at this level: How does the throwing motion stress each component, in turn, of the rotator cuff?</td>
</tr>
<tr>
<td>Application</td>
<td>Use of Understanding</td>
</tr>
<tr>
<td></td>
<td>Exam question at this level: Why does throwing a curve ball cause rotator cuff injury?</td>
</tr>
<tr>
<td>Understanding</td>
<td>Management of Knowledge</td>
</tr>
<tr>
<td></td>
<td>Exam question at this level: How does the rotator cuff help you to raise your arm?</td>
</tr>
<tr>
<td>Knowledge</td>
<td>Memorization of facts, language, concepts, principles, theories</td>
</tr>
<tr>
<td></td>
<td>Exam question at this level: Name the muscles of the rotator cuff.</td>
</tr>
</tbody>
</table>
Table 1A: Related Verbs for Bloom’s Taxonomy

Some relevant and useful verbs for assessments at each level (after Lynda Harding, California State University, Fresno).

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>Understanding</th>
<th>Application</th>
<th>Analysis</th>
<th>Synthesis</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>define</td>
<td>arrange</td>
<td>apply</td>
<td>analyze</td>
<td>arrange</td>
<td>appraise</td>
</tr>
<tr>
<td>identify</td>
<td>classify</td>
<td>compute</td>
<td>calculate</td>
<td>assemble</td>
<td>assess</td>
</tr>
<tr>
<td>indicate</td>
<td>comprehend</td>
<td>construct</td>
<td>categorize</td>
<td>collect</td>
<td>choose</td>
</tr>
<tr>
<td>know</td>
<td>describe</td>
<td>demonstrate</td>
<td>compare</td>
<td>compose</td>
<td>compare</td>
</tr>
<tr>
<td>label</td>
<td>discuss</td>
<td>dramatize</td>
<td>contrast</td>
<td>construct</td>
<td>contrast</td>
</tr>
<tr>
<td>list</td>
<td>explain</td>
<td>employ</td>
<td>criticize</td>
<td>create</td>
<td>decide</td>
</tr>
<tr>
<td>memorize</td>
<td>express</td>
<td>give</td>
<td>debate</td>
<td>design</td>
<td>estimate</td>
</tr>
<tr>
<td>name</td>
<td>identify</td>
<td>examples</td>
<td>determine</td>
<td>formulate</td>
<td>evaluate</td>
</tr>
<tr>
<td>recall</td>
<td>locate</td>
<td>illustrate</td>
<td>diagram</td>
<td>integrate</td>
<td>grade</td>
</tr>
<tr>
<td>record</td>
<td>manage</td>
<td>interpret</td>
<td>differentiate</td>
<td>organize</td>
<td>judge</td>
</tr>
<tr>
<td>relate</td>
<td>paraphrase</td>
<td>investigate</td>
<td>disassemble</td>
<td>perform</td>
<td>measure</td>
</tr>
<tr>
<td>repeat</td>
<td>recognize</td>
<td>operate</td>
<td>distinguish</td>
<td>plan</td>
<td>rate</td>
</tr>
<tr>
<td>select</td>
<td>report</td>
<td>practice</td>
<td>examine</td>
<td>prepare</td>
<td>revise</td>
</tr>
<tr>
<td>underline</td>
<td>restate</td>
<td>predict</td>
<td>experiment</td>
<td>produce</td>
<td>score</td>
</tr>
<tr>
<td></td>
<td>review</td>
<td>schedule</td>
<td>inspect</td>
<td>propose</td>
<td>select</td>
</tr>
<tr>
<td></td>
<td>suggest</td>
<td>shop</td>
<td>inventory</td>
<td>set up</td>
<td>value</td>
</tr>
<tr>
<td></td>
<td>summarize</td>
<td>sketch</td>
<td>question</td>
<td>synthesize</td>
<td>weigh</td>
</tr>
<tr>
<td></td>
<td>tell</td>
<td>translate</td>
<td>relate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>translate</td>
<td>use</td>
<td>solve</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Good indicator statements usually contain action verbs. Table 1A lists action verbs in an array that is associated with the six cognitive levels of Bloom’s taxonomy. As such, objectives containing these verbs, or their synonyms, tend to indicate the cognitive levels at which performance is expected.

The use of indicators infers standards, that is, criteria for what constitutes good performances or acceptable products. For instance, an examination of general education skills in the sciences might precipitate the question, “How would we know if undergraduate biology students are proficient critical thinkers in their discipline?” To assess critical thinking in biology, one might design a task whereby students encounter factual biological observations and, consequently, must construct explanatory hypotheses. The adequate performer might produce one hypothesis and skillfully marshal the evidence in support of it. The excellent performer would go beyond
one hypothesis to offer two or more alternatives that are consistent with the available evidence, and then suggest what additional evidence would distinguish between them. The inadequate performer would offer only a claim, with the evidence either misapplied or absent. This example reveals the nature of critical thinking as a general education concept applied within a discipline. The qualities of a testable scientific hypothesis and its relation to the evidence are part of the domain of science. A scientific hypothesis gains strength when alternative explanations are able to be discarded, that is, when one attempts to disprove or falsify the hypothesis rather than try to prove it true (Karl Popper summarized in Godfrey, 1985). Stated another way, the disciplines approach their examination of ideas differently. Scientific disciplines establish provisional and falsifiable hypotheses as windows through which to view the empirical world. Science depends on objective evidence, new additions to which can overthrow a hypothesis at any time. The discovery of chromosomes roughly 100 years ago overthrew the hypothesis that blood contains the heritable principle, even though the notion of truly heritable blood lines had governed the successful breeding of horses, dogs, and cats for over 2,500 years. Therefore, assessment in the sciences can rightfully include examining the readiness of students to discard even cherished hypotheses when confronted with unavoidable, contrary evidence.

In contrast to the sciences, disciplines in the humanities deal with the subjective nature of experience. Falsification of hypotheses through examination of objective evidence is not what the humanities do. Rather, the humanities appraise the human condition—sometimes by observing people and sometimes by contemplating texts—and examine ways that we humans either discover or create meaning in our circumstances. For instance, how do people regard life and what causes some of them to trade it away in order to uphold an ideal? Is there a universally right behavior, or does ethical rightness vary from one culture to another? What did Benjamin Franklin mean when he said, “There never was a good war or a bad peace”? Accordingly, assessment in the humanities might investigate students’ capacities to describe the history behind a text and to show how different religious, cultural, economic, or political settings have prompted different uses of its wisdom. Distinctions between the
Disciplines reinforce the view that general education gains power when applied to, and filtered by, a discipline’s way of knowing. Assessment of general education also gains power when used in a disciplinary context, and this happens when virtual goals are translated into visible, discipline-related tasks. Indicators are always imperfect, but they are indicators nevertheless.

**Employ Appropriate Methods**

The list below presents some of the many assessment methods that can be employed within the disciplines to assess general education learning. Assessment is many things on different campuses, but whatever those things may be, they are always *multiple measures over time.* One of the first things many effective assessment programs have done is to take stock of what is already going on. Surprisingly, a lot is usually happening, but it just isn’t being called assessment. Regarding already ongoing efforts, some of the best assessment advice ever established was given out early in the assessment movement by Pat Hutchings, first director of the American Association for Higher Education (AAHE) Assessment Forum: “Cultivate assessment methods that match the culture of your institution.”

Within the disciplines, the general education goals-translated-into-objectives are assessable through embedded assignments and activities, and many examples come to mind. An individual might start by following Pat Hutchings’s advice and ascertaining what is already happening. One way to do this is to employ Amy Driscoll’s (California State University, Monterey Bay) Curriculum Alignment Grid (Driscoll, 1998). A purpose of the grid is to make visible the student learning objectives that are implicit in a disciplinary syllabus. One begins by asking a professor or professors to describe what they want students in their courses to “look like” at semester’s end. In other words, what are the learning goals for students in the course? The goals are translated into major learning objectives and these are written in abbreviated form across the top of the grid. Typically, 5–12 course goals work well. Down the left side of the grid appears the daily schedule (e.g., Day 1, Day 2, Day 3)
plus readings (Reading 1, Reading 2, Reading 3), exams, projects, and so on. The professor(s) put Xs in the grid where various learning objectives are covered during class meetings, readings, exams, and projects.

So far, this technique has the same properties as numerous similar matrices claimed to be used for assessment in the disciplines and in general education. The usual practice has been to examine a battery of such matrices created by a departmental faculty and, through a combination of luck, cajoling, and pressure from above, assure that all the departmental learning objectives show up on someone’s syllabus. The operating assumption has been, “We covered it, so the students have been exposed to it, and therefore they must have learned it.” Used only in this way, a matrix may serve as syllabus assessment, but it is not an assessment of student learning. Driscoll’s technique takes an extra step. She asks students to indicate for each class session, reading, exam, and project which of the course objectives they have perceived. By the end of the course, all students have contributed information sufficient to fill in a student version of the grid. Now one looks for alignment of professors’ intents with students’ perceptions. Ideally, for effective learning to occur,
students should generally focus on what the professor intends for them to focus on. What factors in the learning environment contribute to good alignment between a professor’s intents and students’ perceptions? This assessment technique acquires added magnitude when two or more professors in a department collaborate. They need not agree on all course objectives, merely some of them. The more closely that student attention aligns with the faculty’s major objectives, especially general education objectives, the more effective is the learning likely to be. Using the grid as an assessment of student focus within a discipline helps to reveal what students perceive as important. Ideally, we want students to focus on what professors have identified as most important in a course or curriculum. This course/curriculum alignment grid makes embedded general education goals visible, first to the professor and then to the students.

Other powerful opportunities for assessing general education goals within the disciplines involve performances, portfolios, and projects. Such artifacts of student engagement are routinely available in some disciplines and can be incorporated into general education assessment, usually with little cost. Laboratory, field, and stage performances, especially if a team is involved, can enable assessment of problem solving, respect for diversity, oral communication, and leadership. Portfolios make visible such things as creative thinking, critical thinking, ethics, organizational skills, and written communication. Projects open to inspection such areas as aesthetics, computer skills, creativity, ethics, persistence, and problem solving. Additionally, a benefit of assessing general education through the disciplines is the capacity to overlay Bloom’s taxonomy on the assessment in order to examine the cognitive level of student achievement and its integration into disciplinary epistemology.

Finally, although transcript analysis is not a discipline-specific assessment tool, it is mentioned here because of its power to integrate general education assessment into disciplinary thinking. Peter Ewell (for example, 2002) often mentions transcript analysis in his presentations on general education assessment. By combining transcript analysis with other forms of general education assessment, one can address four important ideas:
1. What are the goals of general education? Are they explicitly stated? Where do students encounter them? Are they aligned with other educational goals for students?
2. What is the structure of general education … really?
3. What is the evidence that the process in place is producing the desired results?
4. If change is desired, where (or who) is the lever for changing curricular performance?

Ewell’s advocacy for transcript analysis emanates from question 2 above. The general education curriculum lauded in the catalog may not actually be the one that students take. For example, do students actually select courses in the pattern that is outlined by the catalog, or do they choose courses based on what is available at 10 o’clock in the morning? Do they take courses in the order specified by prerequisites, or do they do quite well by meeting requirements out of sequence? Which upper division disciplinary courses intentionally invoke and build upon lower division general education skills? Transcript analysis, in combination with other assessments, has great potential to reveal how the disciplines integrate general education with disciplinary knowledge.

Produce Significant Results

One of the most powerful devices available to assess general education in the disciplines is Primary Trait Analysis (PTA) (Walvoord & Anderson, 1998) and it works marvelously well with the academic simulator described earlier. PTA has the advantage of being flexible. It works on everything from single classroom assignments to entire curricula. PTA has the advantage of being embedded in what professors do, namely, grading student work. Accordingly, extra expenditure of time and energy is small compared to other assessment methods. PTA has the advantage of adjusting performance standards to local expectations. Thus, one can apply PTA to first-year students and to seniors in a way that controls for cognitive maturity. Moreover, one can do studies either with or without external benchmarks. Other advantages include using information
already available, bringing to consciousness the mostly subconscious processes that go into recording grades, and looking at performance strengths and weaknesses in individual pieces of an assignment, course, or curriculum. Walvoord and Anderson have described Primary Trait Analysis clearly and extensively. Web-based tutorials exist (e.g., SIUE). At its core, PTA takes what professors already do—record grades—and converts that process into assessment.

Each teaching professor has a view of what s/he wants students to accomplish. This view, even if it is an unconscious one, pictures ideal student achievements at the end of a particular class, a unit of instruction, or an entire curriculum. At the end of an assignment or course, students who achieve the goals and “look like” the ideal tend to get As; those who look a bit less like the ideal get Bs, and so on. Because students (and professors) aren’t perfect, achievement of goals is usually uneven. Students may excel in one area and be merely adequate in another. Nevertheless, professors record a single, holistic grade that tends to sum up the student’s performance and provide an overall judgment of merit.

Primary Trait Analysis does not yield a single, holistic grade. Instead, it reveals parts. To put an earlier example to work, suppose one wishes to assess general education goals and has identified aesthetic awareness and problem-solving skills as major educational goals in the discipline of civil engineering. These items become primary traits and, when suitably translated, are roughly the same as learning objectives. The “good assessment question” is, to what extent does the curriculum produce students who possess these primary traits? One might design a task—in a single course or for a culminating senior assignment, involving one person or a team—around designing a bridge across a river for a city wishing to promote its own history. The bridge environment could have certain physical, economic, and cultural constraints. In completing the task, students may perform very well in negotiating the structural problems posed by the river, its flow, and its embankments, but do less well in devising a structure that harmonizes with the cultural and economic intents of the sponsors. The resulting projects would receive grades in the usual way, and they would also undergo Primary Trait Analysis. Regardless of grades in the civil engineering
course, PTA allows assessment practitioners to isolate the indicators of aesthetic awareness and problem solving, thereby helping to characterize the curricular fortification of these general education goals. Thus informed, the faculty could intervene to improve student attainment of a desired level of aesthetic awareness … or honor its high attainment! The bridge design project just described exists as the aesthetically celebrated, new Lewis and Clark Bridge across the Mississippi River at Alton, Illinois.

Primary Trait Analysis derives its name from the notion that some aspects of student learning are more important than others. That is, some traits are primary. Some assessment thinkers have described primary traits as those things that characterize students five years after graduation. They are the meaningful things that students should readily recall that they learned in a course, a curriculum, or a collegiate career after five years or more have passed. Therefore, PTA has the salutary effect of making professors ask, “What do we really want our students to learn as a result of their time with us?” By making these choices, professors register what is important to assess. The opening paragraphs of this paper suggest that general education goals head the list.

Arrange for Effective Presentation

To recite again from Peter Ewell’s many general assessment presentations, “Institutional effectiveness is a comparison of results achieved to goals intended.” In preparation for this paper, a content analysis of over 100 requests for assessment reports, encompassing regional and national accrediting agencies, state boards, boards of trustees, and deans, revealed three fundamental questions:

1. What have you learned about students’ learning?
2. What are you going to do about what you have learned?
3. When, where, and how are you going to do it?

At their core, good assessment reports respond to these three fundamental questions, whether for an examination of general education
outcomes within the disciplines or for anything else. In amplified form, these three questions ask: (1) What has your assessment effort presently revealed to you about your students’ learning? What evidence supports these revelations? What aspects of your [general education] curriculum are working well? What aspects do you wish to improve? (2) What are you going to do with what you’ve found? What plans do you have? What actions will you take to deal with your discoveries? (3) When, where, and how will you do them? What are your next steps? How are you going to use the feedback information you’ve obtained to improve your curriculum and what is the next focus of your attention? In a remarkable confluence of principles for assessment of general education and for assessment of professional education, the Association for Advancement of Colleges and Schools of Business states, “In a self-evaluation process, a school assesses its accomplishments in relation to its mission and objectives, as well as on the criteria in the accreditation standards. [Validation by external peers confirms the internal] analysis of a school’s processes for achieving its mission, assessing educational outcomes, and planning for continuous improvement” (Association for Advancement of Colleges and Schools of Business [AACSB], 2001). The AACSB statement aligns well with the Carnegie Foundation/AAHE Principles for the Scholarship of Teaching (Shulman, 1999). In terms of documenting institutional effectiveness, one can satisfy many expectations of an assessment report for almost any questioner by responding to the three questions above.

Practice Reflective Critique

Of the several enduring themes that pervade assessment discussions in print and speech, one of the most prominent is discussion of continuing disconnects between the output of assessment reports and the input needed for curricular improvement. In short, the feedback loop is thin or nonexistent.

By the early 1990s, the assessment movement had made major inroads, with many institutions engaged in assessment activities and attendance booming at events such as AAHE’s assessment forums. On most campuses, however, faculty still
viewed assessment as externally imposed reporting requirements having little to do with their business of research and teaching, and they objected to the public’s oversimplified view that measuring learning was simple. If there was a “learning problem,” they tended to argue, it resided in student motivation and inadequate schooling prior to college. (Lazerson, Wagener, & Shumanis, 2000)

Within general education,

The NCPI [National Center for Postsecondary Improvement] survey reported that only a third or so of the institutions assessed student’s higher order learning skills, affective development, or professional skills. Alternative assessments—like the much-talked-about portfolios, capstone projects, and observations of student performance—were infrequent. In their “most disappointing finding,” the NCPI investigators also discovered that institutions rarely used their student assessment data in academic decision making. (Lazerson et al., 2000)

Numerous voices have made compelling cases for closing the feedback loop, and failure to do so is not limited to assessment of general education, either within the disciplines or as a separate entity. Those exhortations will not be repeated here. What will be recited are Tom Angelo’s “10 Guidelines for Assessing As If Learning Matters Most” (Angelo, 1999, p. 6). They apply well to assessments of all kinds, including assessment of general education within the disciplines. Angelo wrote, “If learning really matters most, then our assessment practices should help students develop the skills, dispositions, and knowledge” that are needed to:

- Engage actively—intellectually and emotionally—in their academic work;
- Set and maintain realistically high, personally meaningful expectations and goals;
- Provide, receive, and make use of regular, timely, specific feedback;
• Become explicitly aware of their values, beliefs, preconceptions, and prior learning, and be willing to unlearn when necessary;
• Work in ways that recognize (and stretch) their present learning styles or preferences and levels of development;
• Seek and find connections to and real-world applications of what they’re learning;
• Understand and value the criteria, standards, and methods by which they are assessed and evaluated;
• Work regularly and productively with academic staff;
• Work regularly and productively with other students;
• Invest as much engaged time and high-quality effort as possible in academic work.

These ten assessment guidelines affirm how assessment should feed back to reinforce, among other things, the significance of self-motivation, high performance standards, application of knowledge toward solving problems, good communication, and time on task. They resemble the general education goals that opened this essay, don’t they?

References


Hersh, R. (1994). What our publics want, but think they don’t get, from a liberal arts education. AAHE Bulletin (November), 8–10.


