Kenyon College

Assessment Workbook:
A Guide for the Perplexed

2009

In collaboration with:
Reaccreditation 2010
Laurie Finke, Professor of Women's & Gender Studies; Self Study Coordinator
Paula Turner, Associate Provost

Introduction

What is Outcome Assessment?

For our purposes let's define assessment as a continuous process used by the College (a) to evaluate the degree to which all College programs and services contribute to the fulfillment of the College's primary mission; and (b) for documenting and improving the College's effectiveness. Faculty have been engaged with assessment of student learning in academic programs for at least the last decade. However for many divisions of the college this process of self-evaluation will be new. Most of us think of assessment as something that academic programs do, but increasingly calls for accountability from outside the college (i.e. the U. S. government, state legislators) are asking colleges and universities to do continuous assessment of all programs and services, not just curriculum.

Why an assessment handbook?

For many of us who teach at Kenyon and who have been dealing with it for the last two decades, assessment is still a dirty word. The mandate to "assess student outcomes" feels like an onerous task that has been foisted on us by bureaucrats from the outside. We fear that whatever assessment is (and we aren't always sure we know), it will create a lot of busy work for a faculty and staff already overburdened by work and that the data we collect will end up by being sucked into an informational black hole, never to see the light of day (or maybe that's just what we hope). However, most faculty, in fact most employees, at Kenyon, do have a genuine desire to foster in students a lifelong commitment to learning. And we have proven time and again that we are not averse to doing whatever hard work is required to achieve that end.

Assessment is Not Going Away

It is not a fad. I got involved in my first assessment project in 1988, twenty years ago, and if anything the call is becoming shriller and more insistent. The 2006 Spellings Report[1] put the measurement of student learning at the center of its recommendations. The Commissioners wrote:

---

[1] "A Test of Leadership: Charting the Future of U. S. Higher Education," issued by the Commission appointed by Margaret Spellings, U. S. Secretary of Education.

> "Student achievement, which is inextricably connected to institutional success, must be measured by institutions on a "value-added" basis that takes into account students' academic baseline when assessing their results.[2] This information should be made available to students, and reported publicly in aggregate form to provide consumers and policymakers an accessible, understandable way to measure the relative effectiveness of different colleges and universities." (Section 1:4)

And their call is for quantifiable results, preferably from external vendors who will charge a lot of money to purchase instruments which, while they may not be closely connected with our educational mission, will allow us to be compared with other colleges and universities, thus encouraging further the business of college ranking:

> Higher education institutions should measure student learning using quality-assessment data from instruments such as, for example, the Collegiate Learning Assessment, which measures the growth of student learning taking place in colleges, and the Measure of Academic Proficiency and Progress, which is designed to assess general education outcomes for undergraduates in order to improve the quality of instruction and learning. (p. 23)

And calls of this kind are generating a multi-million dollar industry in standardized testing and other measurement procedures, which we can now add to the SAT, GRE, MCAT, LSAT that already suck up our students' time and money. Before we begin to pour money into external forms of evaluation, let's use some common sense and see what we can do for ourselves. Let's try to cut through the bureaucratese. Most importantly, let's ask "what's in it for us?" How can we make assessment work for us by crafting plans that give us useful information that we want and can use.

What are the obstacles to developing workable assessment plans?

1) We don't really quite understand what assessment is or what its uses are.
2) We don't have the tools or experience necessary to formulate realistic, streamlined assessment plans that will work (by which I mean produce usable information);
3) We don't have the time to devote to it in the press of our daily work load; it falls through the cracks because of other more pressing duties (grading that set of tests, completing that search)
4) The information doesn't go anywhere. We collect it and it sits. It doesn't do anything productive for us.

---

[2] Ouch! Ugly dangling participle Commissioners!

It stands to reason that the way to overcome these obstacles is

1) to try to learn more about workable assessment; to explore best practices;
2) to learn some new tools, beginning with very simple measurements that take little time to complete and which yield useable results;
3) to start small and build by assessing one goal at a time;
4) make assessment part of the work we are already doing rather than something added on (i.e. grading);
5) collect information that is useful and then use it.

This workbook offers some advice on how to create a feasible assessment plan that will provide useful feedback with the minimum amount of extra work. Its intended audience is faculty, administration, and staff who are charged with the task of developing and implementing an assessment plan.

For a couple of useful (short and readable; don't you love the titles?) discussions of assessment see

Susan R. Hatfield, "Black Holes and Gaseous Processes: Really Big Assessment Mistakes."

Anna Leahy, "Cookie-Cutter Monsters, One-Size Methodologies and the Humanities," *Inside Higher Education*, Jan. 29, 2009.

Wabash College: Center of Inquiry Blog
http://www.liberalarts.wabash.edu/center_of_inquiry_blog/

Kenyon's Response to Assessment

In the academic division, assessment can be traced back to the 1995 Assessment Plan Kenyon submitted to the North Central Association (our accrediting body) for approval. When many of the data collection methods we proposed proved unworkable, the 2000 Reaccreditation Steering Committee went back to the drawing board and drafted the assessment plan that we currently use for academic programs. All departments and programs at Kenyon that offer a major are required to file an assessment plan that shows how they will assess their major program. These plans are called DOAPs and the annual assessments they produce are called DOARs. All disciplinary departments are required to assess General Education objectives in an annual report called a GEAR (interdisciplinary programs have been exempted).

While RAAS reports good compliance from departments with the assessment program, the quality of the assessment plans and of the information they generate is uneven. Reports from RAAS over the last decade suggest that faculty creating these

reports do not always fully understand what is being asked of them and they are struggling to comply. They fill out the reports but the information they send is not always useful. A review of the GEARS for instance from 2007-08 suggest that faculty are quite proficient at formulating learning goals (they can articulate what they want students to learn in any given class or program), but they find it harder to figure out how to measure student's attainment of those goals. As a result one senses frustration on the part of departments with the whole process.

The 2000 Reaccreditation Self-Study captured this attitude and I quote at length from it because the attitudes they describe still persist, despite the fact that we have been using our current assessment plan for a decade. They wrote,

> ". . .several of the Department Outcome Assessment Plans are too ambitious to be practical, others focus on highly subjective judgments or on student opinion, and still others seek to describe and track program activities rather than outcomes. To date, department chairs have been given very little direction in how to create and implement effective Department Outcome Assessment Plans and have received no feedback on the plans they have submitted. The Steering Committee believes that some of the plans could be simplified and refocused on more rigorous assessments of student achievement." (p. 69)
>
> "Very few faculty, however, are enthusiastic about the annual outcome assessment process. The Steering Committee believes that faculty will only accept the value of the annual assessment reporting requirements when they see evidence that resource allocation decisions are explicitly tied to the results of assessment. . . .Until that happens, annual assessment reports will be seen by many as bureaucratic excess." (p. 70)

This handbook is an attempt to offer faculty some practical help in designing and implementing assessment plans that can succeed in providing the college and departments useful information to aid in curriculum development and allow us to describe to students, prospective students, their parents, our donors, and other interested parties, what we can say with confidence our students learn.

I should stress that this is the first in a series of steps that I hope we will take in making this process of understanding student learning more transparent, more useful.

Why Do Assessment at All?

We already do assessment, all the time. We call it grading.

Much of the literature on assessment stresses the difference between grading and assessment. I believe, however, that this is part of what alienates faculty and makes assessment feel like bureaucratic excess.

Assessment is Grading—Sort of:

Assessment is really not brain surgery. It isn't even particularly innovative. It has been around as long as teaching has. It's called grading.

There isn't a single soul on this campus who thinks that students should not be judged for the work that they produce, that they should not have to prove that they have learned something. Assessment feels to us like a task imposed from outside the institution, while grading is our definition of what we do. As a faculty, we can spend months, even years, arguing about grade inflation but don't want to spend an hour talking about what we expect our students to learn and how we might measure it? So why not work from our strengths? Why not try to see the connections between grading and assessment and use them to build an assessment program that has a chance of working. We might actually at the same time improve our grading practices, obviating the need to discuss grade inflation.

Those in the assessment movement have done us a disservice by constantly telling us that assessment isn't grading. What repels most faculty about assessment is a sense that this is something new and different, that it will be complicated, that it will take a lot of time, and that no one will ever look at it or use it. Fair enough; there are a lot of days I feel the same way. But let's ask, what are the ways in which grading is assessment? When we grade we offer a direct assessment of student learning outcomes. We give a test and the test covers those things we want students to know (the outcomes). We have them write an essay and we evaluate that essay based on the extent to which it demonstrate that the students understand what we want them to understand (more outcomes). Even if we don't articulate a set of "learning objectives" for our assignments, we still do have them in mind. When we evaluate a test, a quiz, a paper, an art project, a musical composition, we assess student learning. If I assign a paper and most of the students do poorly on it, you can bet I am asking myself what went wrong? I am looking to see what I need to change (either the assignment or the teaching) so that the students have a better chance of succeeding. I am sure there isn't one of us who doesn't do the same. In the assessment literature, that's "closing the feedback loop." I have taken something I've learned from the assessment (the test, paper, or whatever) and used it to improve my teaching and hence the students' chances of learning. So assessment really isn't all that alien; what's new is the mandate to make these steps explicit rather than implicit in the grade.

Assessment is also not Like Grading

The assessment literature stresses how different assessment is from grading. And to some extent that's also true. But the basic principles are the same. The difference is the scale. First of all a grade really isn't evaluation of student learning; it's a

reification of the evaluation that is going on, which must remain implicit in the grade. What I mean by this is the following: an instructor takes all the things she want students to learn from an assignment (and there might be dozens of things ranging from content to skills) and crams them willy-nilly into one letter.

Let's take the case of three students, all receiving grades of C+ on a paper. One student receives a C+ because, although well-researched and with a strong thesis, the paper is poorly executed with ill-conceived paragraphs, numerous grammatical errors, and a significant number of typos. Student 2 receives a C+ because, although well researched, the paper fails to move from a summary to a thesis that exhibits a creative and analytical response to the research. Student three receives a C+ because, although well argued and demonstrating true creative flair, the student fails to do enough research (a stated component of the assignment). The grades reflect an assessment, but the grade alone cannot tell us which of the learning objectives embedded in it have been achieved and which have not. I count something like six different learning objectives in the example above, all of which are incorporated into that C+. I really think that is a weakness of grading and one that should give us pause. The students who receive that C+ may never figure out which of the learning objectives they failed to demonstrate (at least the grade can't tell them).

Figure 1: Breakout of Learning Objectives for Hypothetical Example

Objective	Student A	Student B	Student C
Research	Good	Good	Poor
Thesis	Good	Poor	Adequate
Organization	Poor	Adequate	Adequate
Creativity	Adequate	Poor	Good
Analysis	Adequate	Poor	Good
Grammar and Usage	Poor	Adequate	Adequate

Assessment asks us to be able to differentiate among those three C+. That is it focuses on teasing out and evaluating the objectives that underlie our assignments (the things we want students to learn). Figure 1 shows how an assessment of the assignment might break out the learning objectives implicit in the letter grade.

Another critical difference between grading and assessment that the above example suggests is that grading is primarily directed toward the individual student. We rarely consider the collective performance of all students in a class, department, or division. The one place where we do tend to look at student grades collectively is in examining grade inflation. So our grade inflation discussion is also connected to questions of assessment. But just as an individual grade cannot tell us about an individual student's learning, our grade inflation statistics cannot tell us collectively

what students are actually learning because the grade inflation statistics are as reified as the original grades upon which they were based (that is, the single number, say an average GPA of 3.3 again crams all sorts of learning outcomes for many, many students into a single number and so cannot really tell us much about the effectiveness of our programs.)

So what information might we take out of the example above (remember it is entirely made up and bears no relationship to any actual student or students or classes they might take)? It might enable us to make an argument that moves beyond the judgment that "our students write badly." A reading of the data enables us to tease out which writing skills are weakest. In this example, I'd say the biggest problem areas seem to be grammar and organization. Other areas seem at least adequate. Research skills seems to be the strongest. This information might be more useful to faculty in thinking about possible remedies than simply saying students write badly (of course in a real situation, information from more students will make the picture more nuanced).

What is Student Outcome Assessment?

The example above is outcome assessment. Outcomes Assessment attempts to measure what students have actually learned in relation to what we teach them. In fact, despite all of the jargon assessment can be boiled down to three simple questions:

1) What (specifically) do we want students to know?
2) How do we know that they have learned it?
3) How can we help them to learn it more effectively?

Simple as this sounds, we need to make some basic distinctions and demystify the language.

Input and Output Assessment

We must first distinguish between inputs and outputs. Our teaching are the "inputs." The curriculum, requirements, particular classes, senior exercise assignments, lectures, discussions, seminars, are all inputs. As instructors, we have the most control over inputs and we are the most comfortable with assessing them. For instance, in our faculty evaluation procedures, part of our evaluation of teaching is based on faculty letters whose evidentiary basis comes from observing classes and reviewing syllabi and assignments. When we have external reviews, those reviewers are generally looking at inputs (increasingly, however, external reviewers are interested in knowing more about outcomes). We are pretty comfortable with this kind of evaluation. But what the evaluation of inputs, however high quality, cannot tell us is whether our students are actually learning the things we want them

to learn. For this we need to look more closely at what students are actually doing; these are "outcomes." We are often anxious about outcome assessment because we believe, as the adage goes, that we can lead the horses to water but we can't make them drink. We are anxious about having our work evaluated based on what students have learned because we cannot control student behavior. We can only control what we do in class. Assessment doesn't necessarily change that, but at the very least it might help us to see which teaching strategies succeed and which do not, thus offering some measure of the effectiveness of our teaching. It might help us to eliminate procedures that take up a lot of time, but which yield insignificant results.

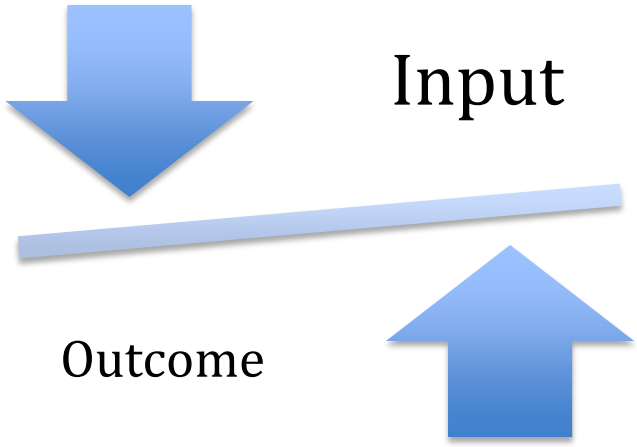

Figure 2: Input and Outcome

As Figure 2 suggests, what assessment strives to do (and what we already do implicitly) is to establish a feedback loop between inputs and outcomes. As we evaluate our students' work, we go back and examine our teaching (curriculum, syllabi, assignments) and refine them. Again the assessment of student outcomes simply makes explicit a process that goes on informally and implicitly all the time.

Is there a benefit to making what is implicit in our work more explicit? In this case I think yes. There is a psychology to grading that tends to make us remember and dwell on the negative outcomes of our grading. We tend to forget the things that went right. Since I believe that in fact our students do have transformative experiences at Kenyon, I also believe that making our assessment of student learning more explicit might show us some surprising things that our students are learning and that we are doing very well and it might help us more effectively target those things that need attention.

Direct and Indirect Assessment

We can evaluate our students learning either directly or indirectly. As part of the evaluation process for faculty, TPC and the provost look at student evaluations for all courses and at letters written by students about the faculty member's teaching. These are indirect assessments of student learning. In these assessments, students are telling us what they think they have learned. Other examples of indirect assessment include exit interviews and surveys. But of course indirect assessments, however useful, do not necessarily demonstrate that students have learned what they say they've learned. A test is a direct assessment of learning (see above on the relationship between grading and assessment). If I want students to demonstrate that they know how to find the hypotenuse of a right triangle or they know the causes of the Spanish American war or what a metaphor is, I can ask those question on a test and if they answer correctly, I have direct proof that they learned it.

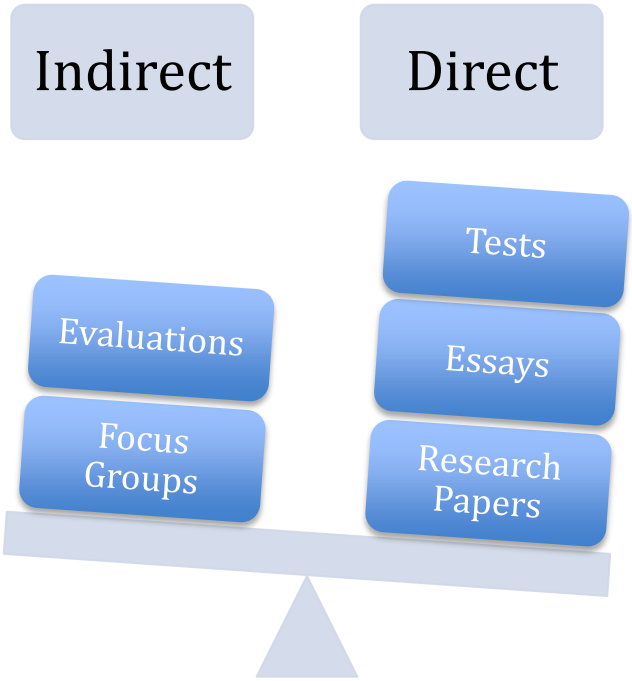

Figure 3: Direct and Indirect Assessment

Increasingly external accrediting bodies are asking for direct assessments of student learning outcomes. For instance, the Higher Learning Commission's Criteria for Reaccreditation, Core Component 3a requires us to demonstrate that "Assessment of student learning includes multiple direct and indirect measures of student learning."

For some examples of direct and indirect assessment tools from Kenyon departments, see Best Practices.

Developing an Assessment Plan

The components of a good assessment plan are identified in the literature as follows:

1. Create a mission statement.
2. Develop a set of measurable learning objectives or goals (This is by far and away the single most important step.)
3. Identify feasible methods of measuring those goals (I repeat feasible).
4. Gather the data using those measures
5. Analyze the data and draw some conclusions from them
6. Make appropriate changes based on those conclusions.

Return to step 2 and repeat

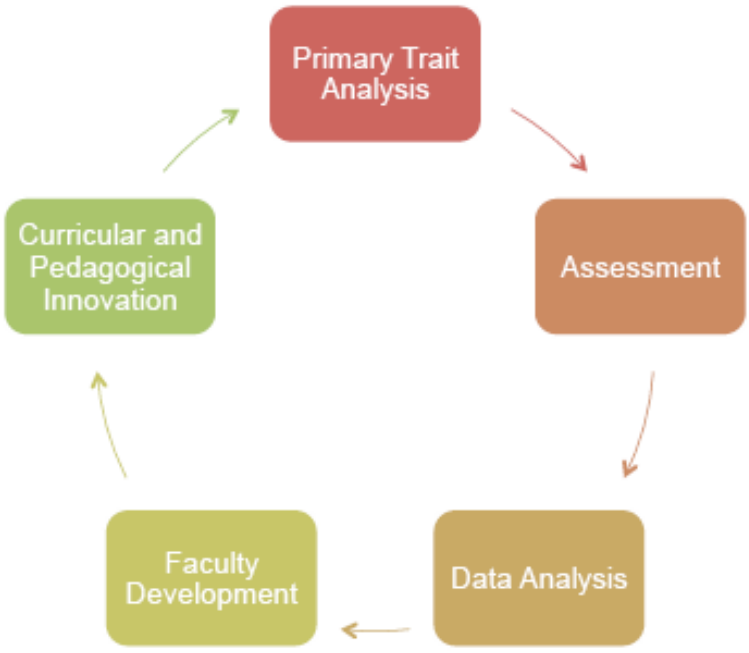

Figure 4 Assessment Implementation Cycle

Let's go through these steps.

1.Mission statement. Keep it brief. Focus on what you most value. Don't be afraid to highlight conflicts within the field if you do not have consensus on a mission. After all what makes research fun is the variety of ways in which we can approach the same thing.

2. It is counterproductive to try to measure outcomes before you know what outcomes you want to measure. So it is important to have a clear sense of what you believe students should be learning. How you state a goal is as important as the goal itself. Aim for something you can measure.

Bear in mind that you do not have to measure every goal every year. You could, for instance, identify one or two goals to measure each year and cycle them through on say a four-year cycle. The advantage of this approach is that it allows some time for you to implement changes and give those changes a chance to work before you measure them again. You probably aren't going to get usable results from a curriculum revision in only one year anyway.

3. Beg, borrow, and steal outright here. As Hemmingway once said, "good writers borrow; great writers steal outright." Assessment has been going on for a very long time and there are already out their many good approaches to it. There is probably one for every type of assessment imaginable. The most important thing to remember is to choose measurements that seem to you appropriate to your goals. Use measurements that fit well with the types of evaluations you already do. For instance, in a curriculum that relies heavily on testing to measure outcomes, a standardized test will probably be the best assessment tool. For a curriculum that relies heavily on writing, there are plenty of holistic writing evaluations that are simple and can serve to tell you something about what students can and can't do. Even ETS has figured out how to do this. You probably don't need to reinvent the wheel.

4. Try to find fairly simple and efficient ways of collecting and organizing the data. You don't need a lot of quantitative analysis. Simply counting the number of students who are outstanding, proficient, adequate, or inadequate on a particular outcome is helpful. The idea is not to dazzle everyone with a lot of charts and graphs and bells and whistles (always what I try to do!); it's to provide information that you can use in your departments and classes. If it takes a long time to do, it probably is too complicated.

5. If you have completed the first four steps well (by which I mean simply and with your own needs in mind) you should have some information that you can use, that can feed back into your thinking about inputs.

6. Make corrections, changes, or requests for resources as necessary.

Establishing Goals

Perhaps the most difficult step in assessment is the process of formulating learning objectives that are meaningful and measurable. Here is a fascinating and disturbing discovery. A study that looked at the best practices in assessment at three institutions considered to have exemplary assessment programs discovered that 80% of the learning goals being measured in those programs fit within the first three and lowest cognitive domains of Bloom's learning taxonomy: remember, understand, and apply.[3] Higher-level critical skills, which are harder to assess, make up only 20% of learning goals being measured. These include important skills like analyzing, evaluating, and creating (Presentation by Christopher A. McCullough and Elizabeth A. Jones at 2008 Higher Learning Commission Annual Meeting).

Here are some tips to think about as you attempt to establish goals you can work with.

1. As I said above, develop a reasonable number of objectives. If you have too many, your assessment will become onerous and it will be harder to target specific improvements you might make. On the other hand too few will limit your ability to portray the complexity of student learning and you might miss areas where further attention is warranted. 5-8 seems a reasonable number of goals if they are well chosen and crafted. It makes sense to assess one or two goals per year and cycle them on a regular schedule, allowing you to make adjustments and some time to pass before you reassess.
2. Link your goals to college-wide objectives. Ample consideration should be given to the institutional mission, goals, and objectives to demonstrate clear alignment. We generally do a good job on this one.
3. Perhaps the most important part of a learning goal or objective is the verb. Use verbs that suggest qualities that can be measured. See below for lists of action verbs that generally work well in articulating measurable learning outcomes. Aim for a mix of Bloom's domains.

[3] In 1956, Benjamin Bloom headed a group of educational psychologists who developed a classification of levels of intellectual behavior important in learning. During the 1990's a new group of cognitive psychologist, lead by Lorin Anderson (a former student of Bloom's), updated the taxonomy reflecting relevance to 21st century work. I reproduce the latter in my discussion.

Action Verbs for Formulating Measurable Student Learning Objectives

Knowledge	Comprehension	Application
Count	Associate	Add
Define	Compute	Apply
Describe	Convert	Calculate
Draw	Defend	Change
Identify	Discuss	Classify
Labels	Distinguish	Complete
List	Estimate	Compute
Match	Explain	Demonstrate
Name	Extend	Discover
Outlines	Extrapolate	Divide
Point	Generalize	Examine
Quote	Give examples	Graph
Read	Infer	Interpolate
Recall	Paraphrase	Manipulate
Recite	Predict	Modify
Recognize	Rewrite	Operate
Record	Summarize	Prepare
Repeat		Produce
Reproduces		Show
Selects		Solve
State		Subtract
Write		Translate
		Use
Analysis	**Synthesis**	**Evaluation**
Analyze	Categorize	Appraise
Arrange	Combine	Assess
Breakdown	Compile	Compare
Combine	Compose	Conclude
Design	Create	Contrast
Detect	Drive	Criticize
Develop	Design	Critique
Diagram	Devise	Determine
Differentiate	Explain	Grade
Discriminate	Generate	Interpret
Illustrate	Group	Judge
Infer	Integrate	Justify
Outline	Modify	Measure
Point out	Order	Rank
Relate	Organize	Rate
Select	Plan	Support
Separate	Prescribe	Test
Subdivide	Propose	
Utilize	Rearrange	
	Reconstruct	
	Related	
	Reorganize	
	Revise	
	Rewrite	
	Summarize	
	Transform	
	Specify	

Assessment Matrix and Curriculum Mapping

See attached spread sheet, where you can actually work out a visual diagram of these processes for each of major learning goals you have for your students.

Kenyon Best Practices

With the above in mind, see the Reaccreditation 2010 website for some examples of assessment instruments already in use at Kenyon in various departments and programs.