



# Assessment Methodology

by Daniel K. Apple, Pacific Crest and Marie Baehr, Elmhurst College

The Assessment Methodology is a tool to help one better understand the steps needed to do a quality assessment. By following this process you can learn what you need to know and change what you need to change in order to improve a performance or a product. The discussion and examples of the use of this methodology are geared toward assessment of student learning. Much of the terminology used in this methodology is taken from the Assessment Overview Module.

## Definitions

Key terminology associated with the Assessment Methodology are defined below.

**Assessment** — A process of measuring and analyzing a performance or product to provide quality, timely feedback for improvement.

**Attribute** — A qualitative or quantitative measurable characteristic of a criterion.

**Criterion** — A focus area of quality.

**Evidence** — A collection of specific qualitative and quantitative data on attributes.

**Instrument** — A specific tool used to obtain evidence.

**Method of collection** — What you will use to collect information on the attributes.

**Range** — A span that includes all the quality levels.

**Scale** — A means for determining the quality level of the evidence.

## Assessment Methodology

The Assessment Methodology consists of four main steps along with a set of sub-steps. The methodology is as follows:

### 1. Develop guidelines for the assessor to follow when assessing a performance or a product.

*Both the assessee and the assessor should, before the performance:*

- Define the purpose of the performance or the product.
- Define the purpose of the assessment.
- Determine what is appropriate to be assessed.
- Agree on what should be reported and how it should be reported in the assessment report.

### 2. Design the approach to be used for the assessment.

*Both the assessee and assessor should:*

- Inventory a list of possible criteria to be used as part of the assessment.
- Choose the criteria from the list in step 2a that best meet the previously established guidelines (Step 1). If appropriate, determine the attributes that indicate quality for each criterion.
- For each attribute (or simple criterion), determine the evidence needed to perform the assessment.
- Agree on the scale and range to be used in looking at each piece of evidence.
- Agree on the method of collection that will be used to collect evidence.
- Determine the specific instruments that will be used to collect evidence.
- Set up a plan to collect the evidence in a timely manner.

### 3. Collect and analyze the evidence.

*The assessor should:*

- Collect the evidence agreed upon in Step 2.
- Use the collected evidence to determine and document the strengths and areas to improve.
- Offer feedback during the performance, if appropriate and agreed upon beforehand, with the assessee.

### 4. Report the findings to the assessee.

*The assessor should:*

- Create the assessment report for the assessee, using the results from Step 1d as a guide.
- Analyze various contributions when evidence of attributes suggests a poor performance or product. Determine what part is due to the evidence collected, the criteria chosen, the use of the product (if appropriate), and/or the performance or product itself.

## **Discussion of the Assessment Methodology**

### **Develop guidelines for the assessor to follow when assessing a performance or a product.**

The first step in setting up an assessment is to define the purpose for the performance or product and the purpose for the assessment. With this information, the person whose performance is being assessed or the group who have developed a product (assessee) can better determine what is important to assess, and the person who is observing the performance or collecting information about the product (assessor) is equipped to give accurate and appropriate feedback.

After determining these two purposes, the two parties should collaborate to determine what is appropriate to assess. This depends on the nature of the activity being performed or the purpose of the product, the skill of the person performing that particular activity or the background of the people who developed the product, the level of assessment skill on the part of the person assessing, and the assessor's knowledge of what he or she is assessing. Finally, the assessee and assessor must decide on the form and content of the assessment report, what the report should include and how it should be reported.

### **Design the approach to be used for the assessment.**

In designing an approach for assessment, both parties should collaborate to generate a list of possible criteria that could be used by the assessor to give feedback to the assessee. From this list, both should agree and select the most important criteria that best fit within the guidelines from the first step in the methodology. In most cases, this list should contain no more than four criteria.

For each chosen criterion, determine appropriate attributes to look for during the performance or in the product. Note that in some cases where the assessment is more narrowly focused, the criterion may be manageable enough without defining attributes. It is important for both the assessor and the assessee to discuss how evidence will be collected, the expected range the assessor might find when collecting evidence, and the scale the assessor will use when collecting the evidence. Typically the range will be set by the assessee's abilities while the scale will be set by the assessor's abilities. Finally, the assessor and the assessee should agree on a plan to collect the needed information as well as the specific instruments that will be used to collect the desired evidence.

One of the keys to learning how to assess is to start simple. Often the evidence you collect for analyzing quality can be measured on a basic scale. For example, if you are asked to assess an oral presentation, one of the attributes could be "eye contact." A veteran assessor

might collect evidence by determining the eye contact on a scale of 1 to 10. However, a novice assessor could use a scale of none, some, lots. Both scales elicit information to create constructive feedback.

It is important for the assessor and assessee to agree on the method of collection. This method could be in the form of observing, looking at test results, filling out check lists, or collecting questionnaires. Note that the method of collection is more general than the particular instrument chosen. The assessee needs to be a part of the discussion of choosing the method, but the assessor alone chooses the particular instrument to use, i.e. the particular method, once the general method is mutually satisfactory to both parties.

Both the assessor and assessee should also determine HOW the evidence can be collected in a timely manner.

### **Collect and analyze the evidence.**

Once the design is in place, it is up to the assessor to collect the agreed upon evidence. Once the evidence is collected, the assessor must make sense of it by looking at what it says about the attributes and criteria. This information can then be used to document strengths and areas to improve. The process itself can help the assessor with insights.

In the case where the assessment plan includes the collection of evidence during the performance, rather than waiting for the final assessment report, the assessee may ask the assessor for feedback during the performance, called "real-time" feedback or authentic assessment. If appropriate to the situation and agreed upon prior to the start of the performance, the assessor may offer feedback (to the assessee) during the performance. For example, a basketball coach may give feedback to a player during a game but it is more difficult for an orchestra conductor to give feedback to musicians during a concert.

### **Report findings to the assessee.**

The final step of the methodology is for the assessor to provide the report to the assessee. The assessment report documents the evidence collected and provides a discussion on how it relates to each attribute and/or criterion. An assessment report also includes feedback about how the assessee can improve future performance.

When a performance has not gone well from the perspective of the assessor, it can typically be attributed to the poor quality of one or a combination of the following:

- the performance or product itself,
- the evidence collected,
- the choice of criteria or attributes, or
- a person's use of a product being assessed.

Finally, the assessee may offer feedback about the assessor's performance so that the assessor can improve his or her assessment techniques in the future.

## Examples of Assessing

### Scenario #1

The members of the physics department at the local university have worked during the past year to develop student learning goals and alter the required curriculum to support the newly formulated goals. They would like to have someone outside the department check to see how the designed curriculum supports the stated goals. They ask a member of the physics department at a neighboring college, Professor Baker, to take a look at the stated goals and course descriptions.

Product: curriculum  
Assessee: Members of the physics department responsible for curriculum design  
Assessor: Professor Baker

### Develop guidelines

#### 1a) Purpose of the product:

The product is to be used so that students completing the curriculum will fulfill the goals.

#### 1b) Purpose of the assessment:

The assessment will help to keep the links strong and show how to improve the links between the curriculum and the stated learning goals.

#### 1c) Determine what is appropriate to be assessed:

The department has five goals and developed 13 major courses. Because Professor Baker has a background in physics and has been active in curricular design, those involved feel that all links should be investigated.

#### 1d) Agree on what should be reported:

The links seen between the goals and the courses will be reported as well as the areas of strengths and the areas to improve.

### Design the approach

#### 2a) Inventory possible criteria:

1. the number of links each course has to goals,
2. parallel building of goals, and
3. the length of time a student spends in class fulfilling each goal.

#### 2b) Choose criteria:

The department determines that collecting data for criterion #1 and criterion #2 makes the most sense.

#### 2c) Determine attributes:

For criterion #1, the number of links is sufficient, and no further attributes are necessary. For criterion #2, the goals touched upon within a term, and the sequence of terms each goal is touched upon are attributes.

#### 2d) Agree on the scale and range:

For criterion #1, the number of links ranging from 0 to 5 are the scale and range.

For criterion #2, attribute #1, the number of goals touched upon each term from 0 to 5 are the scale and the range; for attribute #2, the scale will be the number of years a goal is touched upon, and the range will be from 0 to 4.

#### 2e) Method of collection:

The method of collection will be tabulation of information by Professor Baker.

#### 2f) Instruments used:

The instrument that will be used is a grid of courses versus goals that Professor Baker will fill in using abbreviations for none, introduce goal, develop goal, test goal fulfillment.

#### 2g) Set up a plan:

Professor Baker plans to take the stated goals and the course descriptions to fill in the grid.

### Collect and analyze evidence

#### 3a) Collect the agreed upon evidence.

#### 3b) Use the collected evidence to determine and document strengths and areas to improve.

#### 3c) Offer feedback during the performance.

### Report the findings

#### 4a) Create the assessment report:

Professor Baker's report consisted of the grid he created along with a summary of his findings, including a list of goals that seem to be developed most strongly and the goals that seem to be developed less strongly, along with suggestions of places some of the goals could be developed more strongly.

#### 4b) Analyze contributions:

The department and Professor Baker sat down after the department had a chance to read the report. Many of the points Professor Baker made were quite helpful to the department in modifying its program. However, in a few cases, the goals were being fulfilled in certain spots, but the course descriptions did not indicate that this was so. So, rather than modifying the curriculum, the department modified the description of the courses.

## Scenario #2

The program developed for Economics majors at City College includes a first-year course in economics. Professor Kramer, a second year faculty member, is teaching this course for the first time. He has asked Professor Chandler, chair of the Economics Department and one of the developers of the new program, to assess the learning of the students in his class this year.

|                     |                    |
|---------------------|--------------------|
| Performance:        | teaching           |
| Type of assessment: | summative          |
| Assessor:           | Professor Chandler |
| Assessee:           | Professor Kramer   |

### Develop guidelines

- 1a) The purpose of the performance is for students in the class to learn first-year economics.
- 1b) The purpose of the assessment is to find out what the students are learning well and what they should be learning better.
- 1c) Because Professor Kramer's teaching style is similar for most of what he teaches, both Professors Kramer and Chandler agree that the learning of one or two key concepts should be assessed. In looking through the material in the course, they both decide to focus on the concept of supply and demand.
- 1d) Dr. Kramer does not want Dr. Chandler to spend much time on reporting back. Dr. Kramer requests that Dr. Chandler report what the students seemed to learn well with a rationale for this determination. In addition he would like to know what the students seem to be missing and what might be changed in order for it to be learned better.

### Design the approach

- 2a) Criteria include 1) understanding the principle of supply and demand, and 2) applying the principle of supply and demand.
- 2b) Both professors determine that collecting data for #1 makes the most sense.
- 2c & 2d) Dr. Chandler will look at two attributes: 1) the ability to state the principle of supply and demand, with a scale of expectations that range from below expectations to above expectations; and 2) the ability to draw a supply and demand curve, with a grading scale from A to F.
- 2e) The method of collection will be answers to test questions and graded lab reports.
- 2f) The instruments will be: 1) a lab report where the students graph real data and answer questions about supply and demand and 2) a final exam question that requires the student to state the principle and give an example of its use.

- 2g) The assessor will collect the data before the assignments are handed back.

### Collect and analyze evidence

- 3a) The assignments are analyzed. It was found that almost all students could state the principle of supply and demand, and 2/3 of the students had examples that supported the understanding of the principle. For the lab assignment, several students misread the graphs and came up with conclusions that were inaccurate, but correct based on their wrong readings.
- 3b) Dr. Chandler determined that a strength was the ability to apply the principle, and an area to improve was the ability of the students to read graphs accurately.
- 3c) Not appropriate for this scenario.

### Report the findings

- 4a) Dr. Chandler's report included an evaluation of the students' work, some examples of work that he felt was above and below expectation for each attribute, and a summary of the strengths, areas to improve, and insights regarding Dr. Kramer's teaching of the curriculum.
- 4b) In discussing the report with Dr. Kramer, it was discovered that Dr. Kramer assumed that the students had learned how to read supply and demand graphs in the math course that was a prerequisite for the economics course. After receiving the assessment report, Dr. Kramer found out from the mathematics professor that this concept was not taught in the required mathematics course, but was taught in a different mathematics course. Based on this information, the prerequisite for the course was changed and Dr. Kramer vowed to find out early in the term whether or not the students could analyze graphs sufficiently.

### References

- Angelo, T. K., & Cross, P. (1994). *Classroom assessment techniques: A handbook for college teachers*. San Francisco: Jossey-Bass.
- Marzano, R. J., Pickering, D., & McTighe, J. (1993). *Assessing student outcomes: Performance assessment using the dimensions of learning model*. Alexandria, VA: Association For Supervision and Curriculum Development.
- Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.
- Wiggins, G. & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: Association For Supervision and Curriculum Development.