4.1.2 Distinctions Between Assessment and Evaluation

by Marie Baehr (Vice President for Academic Affairs, Coe College)

Educators use two distinct processes to help students build lifelong learning skills: assessment and evaluation. Assessment provides feedback on knowledge, skills, attitudes, and work products for the purpose of elevating future performances and learning outcomes. Evaluation determines the level of quality of a performance or outcome and enables decision-making based on the level of quality demonstrated. These two processes are complementary and necessary in education. This module draws important distinctions between assessment and evaluation, underscoring the need for both processes to occur at separate places and times, and ideally through different roles (4.1.4 Assessment Methodology and 1.4.7 Evaluation Methodology).

Inconsistent Use of the Terms

In the last fifteen years, much has been written about assessment and evaluation, but the terms have not always had distinct meanings. As accrediting agencies have become increasingly interested in improvement, it has become imperative to have a word that describes feedback for improvement that is distinct from one that describes the determination of quality. To add another layer of confusion from the literature, the word “formative” (used as an adjective with assessment or evaluation) has typically been used to describe an improvement process, while the word “summative” has been used to describe a decision-making process (Brown, Race, & Smith, 1996). However, the words “formative” and “summative” mean “as it is being created” and “addition of all things,” respectively. A process to determine quality can both be accomplished either as a performance is being created or after it is completed, so other words should be used to distinguish the two processes.

In the literature of the last several years, assessment has usually been used to indicate that at least some hint of improvement is expected in the assessment process (Bordon & Owens, 2001; Palomba & Banta, 1999). Similarly, evaluation is usually used to indicate that some sort of judgment of quality will be made. The Faculty Guidebook is consistent in its delineation of these two processes of improvement and judgment. Assessment is the term used to look at how the level of quality of a performance or outcome could be improved in the future; it includes strengths that should be sustained as well as high-priority areas for improvement. The assessment process is not concerned with the level of quality; only with how to improve the level of quality. Evaluation is the term used to describe the determination of the level of quality. The evaluation process focuses only on the actual level of quality with no interest in why that level was attained.

Assessment and evaluation both have their purposes, and, when used correctly, both can add significant value to teaching/learning. However, there can be detrimental effects when the people involved have not agreed whether the process is evaluation or assessment, or when the Assessment Methodology gets confused with the Evaluation Methodology.

Key Attributes

Although assessment and evaluation are used for different reasons, they do have some similar steps. Both involve specifying criteria to observe in a performance or outcome. Both require the collection of data and other evidence by observing the performance or by looking at the outcome or product. Both require a performer and a person who collects information about the performance. Both processes also conclude with a report of the findings which include all the similarities and at least as many differences. The relationship between the people involved is different in the assessment and evaluation processes. In both cases a person (either evaluator or assessor) observes or collects evidence about a performance or outcome; another person (either assee or evaluatee) performs or develops an outcome. In both cases a person (either the assee or client) requests the process (either evaluation or assessment). In assessment, the locus of control rests with the performer; in evaluation, it rests with the observer. The report to the performer (assee or evaluatee) is also vastly different. In the assessment process, the report includes information about why the performance was as strong as it was, and describes what could be done to improve future performances. In assessment, there is no mention of the actual quality of the performance; only how to make the next performance stronger. There is no language indicating the level of quality, such as “good,” “terrible,” “terrific,” or “horrible.” Conversely, in the evaluative report, only information regarding the actual quality of the performance is given. This might be in the form of a grade or a score or an evaluative comment, such as “good work.” The purpose of the evaluative report is to report the level of quality and possibly any consequences based on the determined level of quality. It is not used to suggest improvements in future performances.

Table 1 clarifies the similarities and differences between the two processes. The modules 4.1.1 Overview of Assessment, 1.4.6 Overview of Evaluation, 4.1.4 Assessment Methodology, and 1.4.7 Evaluation Methodology give supporting explanations.
**Table 1 Differences Between Processes of Assessment and Evaluation**

<table>
<thead>
<tr>
<th></th>
<th>Assessment</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What is the purpose?</strong></td>
<td>to improve the quality of future performances</td>
<td>to determine the quality of the present performance</td>
</tr>
<tr>
<td><strong>Who requests it?</strong></td>
<td>assessee</td>
<td>client</td>
</tr>
<tr>
<td><strong>Who performs?</strong></td>
<td>assessee</td>
<td>evaluatee</td>
</tr>
<tr>
<td>**Who observes the perfor-</td>
<td>assessor</td>
<td>evaluator</td>
</tr>
<tr>
<td>mance?</td>
<td>assessor and assessor</td>
<td>client (with possible consultation with the evaluator)</td>
</tr>
<tr>
<td><strong>Who sets criteria?</strong></td>
<td>assessee and assessor</td>
<td>client (with possible consultation with the evaluator)</td>
</tr>
<tr>
<td><strong>Who uses the information?</strong></td>
<td>assessee (in future performances)</td>
<td>client (to make decisions)</td>
</tr>
<tr>
<td><strong>When can feedback occur?</strong></td>
<td>during or after a performance</td>
<td>during or after a performance</td>
</tr>
<tr>
<td><strong>On what is feedback based?</strong></td>
<td>observations; and strongest and weakest points</td>
<td>level of quality based on a set standard</td>
</tr>
<tr>
<td><strong>What is included in the report?</strong></td>
<td>what made the quality of the performance strong; and how might one improve future performances</td>
<td>the quality of the performance, often compared to set standards</td>
</tr>
<tr>
<td><strong>Who receives the report?</strong></td>
<td>assessee</td>
<td>client</td>
</tr>
<tr>
<td><strong>How is the report used?</strong></td>
<td>to improve performance</td>
<td>to make judgments</td>
</tr>
</tbody>
</table>

**Case Studies**

Examples of the use of the assessment process or evaluation process can be found in 4.1.1 Overview of Assessment or 1.4.6 Overview of Evaluation respectively. This section addresses ways that evaluation and assessment can become confused.

**Case 1:** The person observing a performance believes he or she is assessing, but the performer perceives the feedback as evaluative because the performer has not worked with the observer to set up criteria and valuable feedback.

**Dysfunctional Partners**

One of the first steps in the Assessment Methodology is for the assessor and assessee to determine the performance or outcome criteria for which the assessee would like to gain feedback. If this step is skipped, no matter how well-meaning the person giving feedback may be, the feedback is likely to be perceived by the assessee as judgmental. Since the control in assessment rests with the assessee, feedback will be used for improvement only if the person receiving the feedback wants to use feedback from the assessor.

**Parent-Child Relations**

All parents want their children to improve. However, parents also want their children to perform at acceptable levels of quality. When a parent gives feedback for improvement using evaluative language to a child in an area in which the child has no desire to improve, the child will perceive this feedback as judgmental. For instance, there is a big difference in the message sent between saying, “Your room is a mess. Clean it up now or you will be punished,” and “If you put your books away and make your bed, your room would look much nicer.”

**In-Class Assessment Exercises**

Students are more used to feeling that they are evaluated by instructors, rather than assessed. Part of the reason for this perception is that instructors do evaluate students by giving grades. Part of the reason is that students are not often included in determining what should be fed back to them. In order for assessment of student learning to work effectively, students must participate in determining the criteria that will be used for their feedback. For example, after giving an assignment that requires a draft, you could ask students to tell you in what areas they would like feedback for improvement. In this way they would have to determine the areas where they feel improvement would make a difference, and it would help clarify that the purpose of the draft is not for a “free” grading cycle.
Case 2: A person, observing a performance and using the same criteria, gives assessment feedback as well as evaluative judgments.

Interim Feedback on Work Products

Students are often dismayed when they make all the suggested improvements on a paper that was turned in for comment as a rough draft and they do not receive an “A” on the final product. In this case, the instructor has given feedback for improvement without determining the quality of the paper. The student perceives that if he or she improves in the areas noted, he or she will have an excellent paper. One way to avoid this problem while strengthening the assessment process is to ask the students to request feedback on the draft based on set criteria.

 Supervisor as Mentor

Often chairs of departments are expected to mentor their non-tenured faculty in their department at the same time that they are expected to make decisions on continuing employment. Although the individuals might agree on criteria to use, it becomes difficult for the assessees to feel in control of using or not using the feedback as he or she sees fit, since, at some point, the assessor will become the evaluator. Although this is sometimes unavoidable, the problem can be reduced by choosing the criteria differently in the two cases. In the mentoring situations, the non-tenured faculty member should choose the criteria for focus, while in the evaluative situations, the chair should. In both cases the criteria need to be known by both parties.
**Case 3:** A person who is more comfortable with the evaluator role is put in the role of assessor.

**Expert Assessing a Novice**

Sometimes, someone who is so ingrained in an area of expertise is unable to stop judging the quality of a novice performance. Though all criteria and scale are agreed upon, the expert as assessor can sometimes give the feedback in evaluative terms without realizing it. This sometimes happens when faculty start teaching right after they have earned their graduate degree. They are not prepared for the limited understanding and skills of the students who are taking their class. Rather than mentoring the students to help them build their knowledge and skills, the faculty members are sometimes apt to evaluate students as unmotivated and poorly prepared.

**First-Time Assessor**

Often, when one is used to giving feedback on the level of quality only, someone can feel uncomfortable giving “critical” feedback to an assessee, feeling that pointing out areas to improve is the same as criticizing the performance. This can cause even more problems when the assessee also perceives the feedback as evaluative (Case 1). Practice and building trust help this situation the most but it can also help if the assessor imagines what feedback he or she would have wanted if he or she had been the performer. It is important for the assessee to send the message that he or she would like to have the feedback from someone he or she trusts.

**Concluding Thoughts**

Discussion in this module is intended to strengthen outcomes from assessment and evaluation in teaching/learning situations. Assessment is a process used to improve a performance or outcome. Evaluation is a process used to determine the quality of a performance or outcome and to make decisions based on the quality. Both processes can be formative (undertaken while an educational process is ongoing) or summative (taken at the conclusion of an educational process). Before starting either assessment or evaluation it is essential for instructors to clarify the purpose of the process. It is then critical to communicate this purpose to everyone involved and to establish whether this will be conducted as assessment or evaluation. Finally, one should be cautious whenever an assessor will ultimately be an evaluator or when assessment is initiated without buy-in of the assessee.

**References**


4.1.9 SII Method for Assessment Reporting

by Jack Wasserman (Mechanical Engineering, University of Tennessee at Knoxville) and Steven W. Beyerlein (Mechanical Engineering, University of Idaho)

Assessment results are most likely to be put into action by an assessee when they are concisely stated, supported by evidence, and delivered in a positive manner. This module outlines a format for informal assessment reports that meets these needs. Known as the SII method, it includes a thoughtful description of assessee strengths, areas for improvement, and insights that can be transferred to other contexts. The SII method is assessee-centered in its language, specific in its use of data from a specific learning context, and enlightening in its recommendations for future action.

The Role of Self-Assessment

Psychological studies of highly successful people across all domains of intelligence—linguistic, musical, mathematical, scientific, interpersonal, kinesthetic, intrapersonal, and spiritual—reveal that these extraordinary individuals share three behaviors that are the source of sustained personal growth (Gardner, 1998).

- These individuals stand out in the extent to which they reflect, often explicitly, on the events of their lives
- These individuals stand out less by their impressive “raw powers” than by their ability to identify and then exploit their strengths
- These individuals fail often and sometimes dramatically, but they stand out in the extent to which they learn from their setbacks and convert defeats into opportunities

Extraordinary individuals, therefore, possess a strong internal process of thinking about their circumstances, their performance capabilities, and their opportunities for effecting change. The SII method strives to make these attributes explicit in the dialogue between assessor and assessee. It embodies several characteristics known to improve critical thinking, including positiveness, process-orientation, a recognition of contextual details, and the role of emotion as well as reason in human behavior (Brookfield, 1987).

Organization of the SII Report

While the assessee is performing, the assessor must collect information consistent with the chosen criteria (4.1.4 Assessment Methodology). It is important for the assessor to note the strong points of the assessee’s performance (things done well) and why they were considered strong; the areas in which the assessee’s performance could be improved, along with suggestions for how the improvement could be made; and any insights that might help the assessee in other contexts. The SII format provides a succinct way to communicate these findings in a cooperative learning environment.

Strengths—identify the ways in which a performance was of high quality and commendable. Each strength statement should address what was valuable in the performance, why this attribute is important, and how to reproduce this aspect of the performance.

Areas for Improvement—identify the changes that can be made in the future, between this assessment and the next assessment, that are likely to improve performance. Improvements should recognize the issues that caused any problems and mention how changes could be implemented to resolve these difficulties.

Insights—identify new and significant discoveries/understandings that were gained concerning the performance area; i.e., What did the assessor learn that others might benefit from hearing or knowing? Insights include why a discovery/new understanding is important or significant and how it can be applied to other situations.

These statements should be delivered in the order given above first to affirm the assessee and then to apprise him or her of opportunities for additional growth. An assessor should take care to cast these statements in a succinct manner and avoid using judgmental language. As a matter of convenience in written SII reports, each statement can be identified with the appropriate letter (S or I).

Rubric for Elevating SII Reports

The following rubric has been developed to help students visualize different levels of assessment quality and to rate the sophistication of their SII reports. As assessments move up the scale, there is a discernible shift from assessing effort to meaningfully assessing performance.

Level 1—Observation

Strengths and areas for improvement are presented as simple statements. The following statements are typical of this level:

(S) The presenter was energetic
(I) The introduction was too long
(I) The score was not the only goal
Level II—Comprehension of Key Issues

Strengths and improvements are clearly stated, and reasons are given for the strengths and suggestions for improvement. Insights tend to be related to the specific context of the assessment. The following statements are typical of this level:

(S) The enthusiasm of the presenter inspired the audience to ask many questions
(I) Much of the material in the introduction was secondary to the purpose of the talk
(I) The team kept the problem statement in mind, not just the score

Level III—Application in a Related Context

This feedback builds on comprehension of key issues and gives specific ideas for improving performance in a related context. The following statements are typical of this level:

(S) Taking time to practice your presentation can help you deliver your message in a confident and convincing tone
(I) The introduction should highlight a single hypothesis and explain why it is justified
(I) By focusing on the goal of good technical communication, rather than focusing simply on the score, the team reminded everyone about the educational objective of the project

Level IV—Transfer to a New Context

This feedback illustrates generalized understanding and is instructive in applying this understanding across a broad range of contexts. The following statements are typical of this level:

(S) Researching the background of your audience can help you stimulate interest in and attention to your message
(I) Section divisions appear to be seamless in a carefully planned and practiced presentation
(I) By communicating your interpretation of the underlying purpose of an activity, you help everyone assess whether they could have learned more from the activity

Implementing SII Reports

SII reports represent a powerful formative assessment tool that can be used with a great deal of flexibility in the classroom. The following techniques have proven successful in elevating and adding variety to SII reports.

Prioritize findings—Students share only the greatest strength, the greatest area for improvement, and the best insight. This encourages participants to rank the significance of their observations and to defend their thinking.

Limit response time—This is especially valuable for sharing oral assessment reports from multiple teams. Challenge participants to limit SII reports (all three parts) to less than 30 seconds.

Build common understanding—Participants are asked to rephrase what they hear in others’ SII reports. This process can help clarify muddy ideas as well as emphasize important discoveries.

Focus attention—The instruction directs attention to a narrow set of learning skills or performance criteria. Focusing the assessment helps to minimize motherhood-and-apple-pie statements; and instead connects the commentary with specific behaviors.

Rate performance on a scale—As a reference for writing SII statements, the instructor provides several scales or rubrics for ranking performance in key areas. Assigning numerical scores can trigger recollection of supporting evidence that adds more specificity to a written SII report.

Collective feedback—At the end of a reporting session (oral or written), the instructor may use the SII format to comment on the entire spectrum of reports. This serves to reiterate key findings and to establish performance expectations for future reporting sessions.

Concluding Thoughts

One of the driving forces for change in higher education is the need to develop students who are lifelong learners who can adapt to the ever-and-rapidly-changing world around us (Brookfield, 1987). Quality self-assessment provides a solid foundation for such self-growth (Gardner, 1998). By giving and receiving SII reports, learners at any level in the curriculum gain the practice and experience they need to become quality self-assessors and self-growers. SII reports support an assessment culture in which students are motivated to perform better and proactively seek to improve their own performance.

References
