DIRECT ASSESSMENT OF STUDENT LEARNING

Jen Sweet
Office for Teaching, Learning, and Assessment
Workshop Outcomes

Participants will know how to:

- Differentiate between Indirect and Direct Assessment
- Identify and Apply Direct Assessment Methods
- Write Effective Prompts
- Develop Scoring Guides and Rubrics
Direct versus Indirect Assessment of Student Learning
Direct and Indirect Assessment of Learning

- *Direct* evidence of student learning is tangible, visible, and compelling evidence of exactly what students have and have not learned.

- *Indirect* evidence consists of proxy signs that students are probably learning. Indirect evidence is less clear and convincing.
Important Considerations

Who decides what was learned and/or how well it was learned?

Does the assessment measure the learning or is it a proxy for learning?
Who Decides what was Learned?

General Rule:

In Direct Assessment, a professional makes a decision regarding what was learned and how well it was learned.
- ex. faculty or staff evaluated reflection paper

In Indirect Assessment, the student decides what was learned and how well it was learned.
- ex. surveys, teaching evaluations
**Is Learning being Measured Directly, or is the Assessment Based on a Proxy for Learning?**

**Direct Evidence**: Students have completed some work or product that demonstrates they have achieved the learning outcome  
- exs. project, paper, performance

**Indirect Evidence**: A proxy measure was used, such as participation in a learning activity, students’ opinions about what was learned, student satisfaction, etc.  
- ex. number of students who visited an office or office hours, grades
Course grades are based on many iterations of direct measurement.

However, course grades are an indirect measurement of any one learning outcome because:

- They represent a combination of course learning outcomes; performance on these outcomes are averaged out in a final grade.
- They frequently include corrections not related to learning outcomes, such as extra credit or penalties for unexcused absences.
Examples of Evidence: Direct or Indirect?

For each of the following examples, work with your group to decide whether you think each of these examples is direct or indirect measurement:
Student Reflections Evaluated by your Assessment Committee
Survey that Asks Students to Report their Learning for Particular Concepts
Retention and Graduation Rates for Students in a Particular Program
Teaching Evaluations
Site Supervisor Ratings of Student Performance during an Internship
Survey that Asks Students to Respond to a Scenario using Content Learned in Class
Peer Ratings of Student Reflections
Course Grades
Challenges of Direct Assessment

Limited or Inconsistent Exposure to Students
- Especially in co-curricular units

Finding Time for Trained Professionals to Evaluate
- Especially in small units/departments

Motivation to Participate in the Assessment
Limited or Inconsistent Exposure

Captive Audience!

Embed Assessment into your Student Contact Activities
Limited Faculty/Staff Time

Bring Faculty/Staff into the Process
- All faculty/staff participate from the beginning (determining what questions you have about student learning) to the end (dissemination of report; discussion of results).

Captive Audience

Team Up
- Find other departments/units interested in the same questions to potentially share the workload
  - Richer sense of what students are learning
  - Application of classroom learning

Plan Ahead!
Motivation to Participate

Bring the Students into the Process

- Explain what you are doing and why you are doing it
- Inform students of what will be done with the results of the assessment
  - Potentially, offer to make results available to students
- Think about giving students an opportunity to provide feedback on the assessment

It helps if it “counts” for something, even if it is not much
Questions about Direct vs. Indirect Assessment?
Writing Effective Prompts
A prompt defines a task:

a statement or question that tells students what they should do (ex. in a survey item, essay question, or performance).

Prompts should “prompt” students to demonstrate the learning outcome that is being assessed.

Two Main Categories of Prompts

- **Restricted Response** prompts limit the way in which students may present information.
  - Advantage: Easier to evaluate
  - Disadvantage: Restricts students ability to provide diverse, individualized responses that provide richer information

- **Extended Response** prompts give students more latitude in deciding how to respond, including the format, length, and construction of the response.
  - Advantage: Provides richer information by allowing for diversity in responses.
  - Disadvantage: Requires more time to evaluate responses.

Importance of Good Prompts

- Good prompts are a critical part of teaching and learning process
  - Communicate expectations
  - Inspires students to do their best and achieve learning outcomes

Laying the Foundation

- Decide what you want students to learn from the experience.
- Determine how the learning aligns with your learning outcomes.
- Develop a meaningful task or problem related to identified learning outcome(s).
- Determine the methods you will use to measure (scoring guide, rubric, reflection, etc.) students’ learning.
“You are there” Scenarios

- You are on the subway and overhear a conversation about . . .
- You are a corporate trainer leading a diversity workshop . . .
- You are a consultant working with a community organization . . .
- You are a business executive leading a high stakes meeting . . .
Prompt Example

Program Learning Outcome:
Students will be able to identify options for post-graduate study/work and understand the implications for each.

Prompt:
Please consider the options you’ve learned about to continue your education after you receive your bachelor’s degree. Select one you think might be a good fit for you and briefly discuss:
- Master’s Degree
- Joint Master’s/PhD program
- Professional School
- Trade School

Work with your group – what do you think of this prompt? Do you see any problems with this prompt? Can you create a better prompt?
Discussion

Did you Identify any Problems?

Ideas for a Better Prompt?
A Few Problems:

- First part of the outcome is not directly addressed because options have been given.
- Data from this prompt indicated students did not understand the implications for post-graduate study options because, by in large, students did not identify any implications. Follow-up with a focus group, however, revealed the students actually did have good knowledge of the implications and were able to discuss them in meaningful ways.

Example of a Better Prompt

Throughout this program, we’ve discussed different options for continuing your education after you complete your bachelor’s degree. Select two or three you feel might be a good fit for you and discuss the implications of each option based on your future goals and aspirations.
Questions about writing prompts?
Scoring Guides and Rubrics
As the name implies, scoring guides generally provide a structure and definition of how student performance will be judged.

A Scoring guide should be tied directly to the student learning outcome it is measuring.

Examples of Common Scoring Guides:

- Rubrics
- Structured Observation Guides
Elements of a Rubric

- **Criteria for Evaluation of Students’ Performances**

- **Defined Levels of Performance**
  generally associated with a numeric value (ex. 1-4)

- **Description of each Level**
  detailed information regarding the qualities that should be or should not be present in a product/performance to receive a rating at each level

Identifying important aspects (elements, behaviors, components, qualities, features, characteristics) of a performance or product to be assessed.

Questions to consider:
- What do I want my students to learn from carrying out this process or producing this product?
- What would a high quality product look like? What are its essential, defining characteristics or features?
- What does the student need to do to complete the performance task or produce the product that I have in mind?
- What are the specific steps the student will need to follow to complete a task?

Deciding on a range of score points for each of the important aspects.

Questions to consider:
- What is the minimum number of performance levels I need to adequately describe the range of performance I am seeing in student work?
- The range of score points can be binary (e.g., Achieved/Did Not Achieve, Met/Not Met) or more than two categories (e.g., Exceeded/Met/Did Not Meet).
  - A common number of categories is between 3 and 7. However, the choice of number of score points should not be arbitrary.

Myford, C. (Spring 2006). Performance assessment, PowerPoint presentation
Two Main Types of Rubrics

1. Holistic Rubrics

2. Analytic Trait Rubrics
   Developmental Rubrics

Holistic Rubrics

A holistic rubric gives one score for an entire work or product. The rubric combines all important components of the student’s performance to arrive at a single judgment of the quality of the final product.

**Advantages**
- Emphasis on what the learner is able to demonstrate, rather than what s/he cannot do.
- Saves time by minimizing the number of decisions raters make.
- Can be applied consistently by trained raters increasing reliability.

**Disadvantages**
- Does not provide specific feedback for improvement.
- When student work is at varying levels spanning the criteria points it can be difficult to select the single best description.
- Criteria cannot be weighted.


Example of Holistic Rubric

Articulating thoughts through written communication – final paper/project.

**Above Average (4):** The audience is able to easily identify the focus of the work and is engaged by its clear focus and relevant details. Information is presented logically and naturally. There are no more than two mechanical errors or misspelled words to distract the reader.

**Sufficient (3):** The audience is easily able to identify the focus of the student work which is supported by relevant ideas and supporting details. Information is presented in a logical manner that is easily followed. There is minimal interruption to the work due to misspellings and/or mechanical errors.

**Developing (2):** The audience can identify the central purpose of the student work with little difficulty and supporting ideas are present and clear. The information is presented in an orderly fashion that can be followed with little difficulty. There are some misspellings and/or mechanical errors, but they do not seriously distract from the work.

**Needs Improvement (1):** The audience cannot clearly or easily identify the central ideas or purpose of the student work. Information is presented in a disorganized fashion causing the audience to have difficulty following the author’s ideas. There are many misspellings and/or mechanical errors that negatively affect the audience’s ability to read the work.

An Analytic Trait Rubric divides the product or performance into important components or traits, then evaluates each one separately.

**Advantages**
- Can provide useful feedback on areas of strength and weakness.
- Criteria can be weighted to reflect the relative importance of each dimension.
- Can give specific feedback on each important dimension.

**Disadvantages**
- Takes more time to create and use than a holistic rubric.
- Unless each point for each criterion is well-defined raters may not arrive at the same score.


# Example of Analytic Trait Rubric

Articulating thoughts through written communication— final paper/project.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Needs Improvement (1)</th>
<th>Developing (2)</th>
<th>Sufficient (3)</th>
<th>Above Average (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clarity</strong> (Thesis supported by relevant information and ideas.)</td>
<td>The purpose of the student work is not well-defined. Central ideas are not focused to support the thesis. Thoughts appear disconnected.</td>
<td>The central purpose of the student work is identified. Ideas are generally focused in a way that supports the thesis.</td>
<td>The central purpose of the student work is clear and ideas are almost always focused in a way that supports the thesis. Relevant details illustrate the author’s ideas.</td>
<td>The central purpose of the student work is clear and supporting ideas are always well-focused. Details are relevant, enrich the work</td>
</tr>
<tr>
<td><strong>Organization (Sequencing of Elements/Ideas)</strong></td>
<td>Information and ideas are poorly sequenced (the author jumps around). The audience has difficulty following the thread of thought.</td>
<td>Information and ideas are presented in an order that the audience can follow with minimum difficulty.</td>
<td>Information and ideas are presented in a logical sequence which is followed by the reader with little or no difficulty.</td>
<td>Information and ideas are presented in a logical sequence which flows naturally and is engaging to the audience.</td>
</tr>
<tr>
<td><strong>Mechanics (Correctness of grammar and spelling)</strong></td>
<td>There are five or more misspellings and/or systematic grammatical errors per page or 8 or more in the entire document. The readability of the work is seriously hampered by errors.</td>
<td>There are no more than four misspellings and/or systematic grammatical errors per page or six or more in the entire document. Errors distract from the work.</td>
<td>There are no more than three misspellings and/or grammatical errors per page and no more than five in the entire document. The readability of the work is minimally interrupted by errors.</td>
<td>There are no more than two misspelled words or grammatical errors in the document.</td>
</tr>
</tbody>
</table>

Developmental Rubrics

A type of analytic trait rubric

Not evaluating an end product or performance, but interested in answering the question, “to what extent are students who engage in our programs/services developing this skill/ability/value/etc.?”

Generally, this type of rubric would be based on a theory of development.

**Advantages**

- Useful when the goal of evaluation is to determine level of development rather than the quality of a final product.
- Rubric can be based on relevant developmental theory

**Disadvantages**

- Conceptually, more difficult to design.
- Requires close tie between assessment criteria and theory of development.
## Intercultural Maturity

<table>
<thead>
<tr>
<th>Domain</th>
<th>Initial Level of Development (1)</th>
<th>Intermediate Level of Development (2)</th>
<th>Mature Level of Development (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive</td>
<td>Assumes knowledge is certain and categorizes knowledge claims as right or wrong; is naïve about different cultural practices and values; resists challenges to one’s own beliefs and views differing cultural perspectives as wrong</td>
<td>Evolving awareness and acceptance of uncertainty and multiple perspectives; ability to shift from accepting authority’s knowledge claims to personal processes for adopting knowledge claims</td>
<td>Ability to consciously shift perspectives and behaviors into an alternative cultural worldview and to use multiple cultural frames</td>
</tr>
<tr>
<td>Intrapersonal</td>
<td>Lack of awareness of one’s own values and intersection of social (racial, class, ethnicity, sexual orientation) identity; lack of understanding of other cultures; externally defined identity yields externally defined beliefs that regulate interpretation of experiences and guide choices; difference is viewed as a threat to identity</td>
<td>Evolving sense of identity as distinct from external others’ perceptions; tension between external and internal definitions prompts self-exploration of values, racial identity, beliefs; immersion in own culture; recognizes legitimacy of other cultures</td>
<td>Capacity to create an internal self that openly engages challenges to one’s views and beliefs and that considers social identities (race, class, gender, etc.) in a global and national context; integrates aspects of self into one’s identity</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>Dependent relations with similar others is a primary source of identity and social affirmation; perspectives of different others are viewed as wrong; awareness of how social systems affect group norms and intergroup differences is lacking; view social problems egocentrically, no recognition of society as an organized entity</td>
<td>Willingness to interact with diverse others and refrain from judgment; relies on independent relations in which multiple perspectives exist (but are not coordinated); self is often overshadowed by need for others’ approval. Begins to explore how social systems affect group norms and intergroup relations</td>
<td>Capacity to engage in meaningful, interdependent relationships with diverse others that are grounded in an understanding and appreciation for human differences; understanding of ways individual and community practices affect social systems; willing to work for the rights of others</td>
</tr>
</tbody>
</table>

Scenario presented for a peer training program related to diversity:

Marieka walks into a study lounge to get some work done between classes and sits at an empty table. She overhears the following conversation between two students she knows from class.

Ed: “I can’t believe my brother didn’t get into Northwestern University – he had a 3.4 GPA in high school! You know that if he was black, he would have gotten in!”

Tiffany: “Yeah, they have all these programs to help ‘minorities,’ but if you’re white, no one cares. A lot of those kids who get special treatment because their great, great, great, great, greeeeaaatt grandpas were slaves are richer and went to better schools than we did. Slavery happened a long time ago – why are they still benefitting from it?”

Ed: “Worse than being white these days is being a white man. It’s nothing more than reverse discrimination!”

1. To what extent do you agree or disagree with Ed and Tiffany’s conversation?
2. If you were Marieka, how would you react?
Structured Observation Guides

A more qualitative type of scoring guide.

**Advantages**

- May allow for a richer description of student performance or work.
- May be useful for assessment of qualities that are difficult to operationally define, like attitudes or values

**Disadvantages**

- This is a more subjective approach to scoring.
- More difficult to analyze results.

## Example of a Structured Observation Guide

### Structured Observation Guide for a Presentation

<table>
<thead>
<tr>
<th>Effectiveness of Presenter in:</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicating the Purpose of the Presentation</td>
<td></td>
</tr>
<tr>
<td>Organizing the Presentation</td>
<td></td>
</tr>
<tr>
<td>Demonstrating Good Knowledge of the Topic(s)</td>
<td></td>
</tr>
<tr>
<td>Speaking with Clarity</td>
<td></td>
</tr>
<tr>
<td>Responding Appropriately to Participants’ Questions</td>
<td></td>
</tr>
<tr>
<td>Adhering to Time Constraints</td>
<td></td>
</tr>
<tr>
<td>Accomplishing the Stated Objective</td>
<td></td>
</tr>
</tbody>
</table>

Questions about rubrics and scoring guides?
References