

SPEAKING THE SAME TONGUE



Coming to Consensus on the Language of Assessment and Evaluation

Ice Breaker

- Assign terms to tables
– ask them to define
- Compare definitions ...
if everyone is the same
we can go home!

Language is important

- It is our symbolic route to shared understanding
- Shared understanding allows us to use the results of others' work to:
 - Serve as the foundation for generating new knowledge;
 - Solve problems in new contexts.



Language is problematic

- Natural language is fluid, metaphorical
- Science requires language that is *precise*
- Imprecision in the language of assessment and evaluation impedes progress in the development and application of new knowledge



Our Historical Context

- Language evolves and diverges in different contexts and cultures
 - "Assessment of Learning" / "Student Evaluation"
 - "Program Evaluation" / "Program Assessment"
- In Pharmacy:
 - Accountability and accreditation
 - Continuous quality improvement
 - AACP Academic Affairs committee report*

Hollenbeck RG. American Journal of Pharmaceutical Education Vol. 63, 1999S

We propose definitions for:

- | | |
|---|---|
| <ul style="list-style-type: none"> ■ Defining Excellence: <ul style="list-style-type: none"> ▪ Quality ▪ Criteria ▪ Standards ■ Quality Measurement <ul style="list-style-type: none"> ▪ Measurement ▪ Validity ▪ Reliability | <ul style="list-style-type: none"> ■ Assessment Versus Evaluation <ul style="list-style-type: none"> ▪ Assessment ▪ Evaluation <ul style="list-style-type: none"> ▪ Constructive / Formative ▪ Judgmental / Summative ■ Quality Culture |
|---|---|

Terms used regularly in the literature that we have NOT addressed: formative assessment, summative assessment, test, grade, goal, objective, outcome, competency, and ???

Part 1 - Defining Excellence

Defining Excellence

- **Quality** – attribution of superior value
- **Criteria** - Characteristics of a performance, object, group, or institution that together determine its value or quality to defined stakeholders
- **Standard** - A measure or cut-off point that determines whether a particular characteristic is deemed acceptable to a defined set of stakeholders or not

Quality of Critical Thinking*

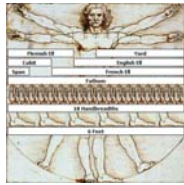
	Exceeding Expectations	Satisfactory	Developing	Novice
Explanation of issues	Issue/problem is stated clearly and described comprehensively, with all relevant information necessary for full understanding.	Issue/problem is stated, described, and clarified so that understanding is not seriously impeded by omissions.	Issue/problem is stated but leaves some terms undefined, ambiguities, boundaries unexplored, backgrounds unknown.	Issue/problem to be considered critically is stated without clarification or description.
Evidence	<p>CRITERIA (indicated by a red arrow pointing to the 'Exceeding Expectations' column)</p> <p>STANDARD (indicated by a purple oval around the 'Satisfactory' column)</p>			
Context and Assumptions				

*Part of the CT VALUE Rubric, AAC&U 2009

Part 2 - Quality Measurement

What is "quality" Measurement?

- The process of **systematically** assigning **numbers** or **qualitative descriptors** associated with **increasing value** to characteristics of a performance, object, group, or institution for the purpose of determining the **degree to which they possess** a certain **characteristic** being studied.



Measurement Quality

- **Reliability** – The results of measurement reflect the performance in question and not measurement errors.
- **Validity** – an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the **adequacy** and **appropriateness** of interpretations and actions based on test scores and other modes of evaluation.

Validity

- **ADEQUACY** determined by 3 types of evidence:
 - **Construct-related evidence** –tested construct embedded in a conceptual framework so that the relationship between scores and other variables clearly indicates the meaning of the score (e.g., IQ test measures “intelligence”)
 - **Content-related evidence** – items, tasks, or questions are representative of the domain (e.g., IQ test items cover all aspects of intelligence discussed in the scientific literature)
 - **Criterion-related evidence** – test scores are systematically related to one or more outcome criteria (e.g., IQ related to academic performance)
- **APPROPRIATENESS** of test interpretation with respect to individual and social consequences

Messick, 1989

Validity Testing

- Inspection of items
- Assessing internal correlations
- Evaluating stability of scores
- Administering the test to individuals who think aloud
- Varying test procedures systematically
- Trying to improve scores
- Studying group differences
- Seeking correlation with practical criteria and other tests

Applications in Academia

High stakes evaluations must be reliable and valid.

- | | |
|--|--|
| <ul style="list-style-type: none"> ■ Questionable validity and reliability: <ul style="list-style-type: none"> ■ Traditional orals ■ Essays ■ Global rating scales ■ Self, peer assessment of knowledge | <ul style="list-style-type: none"> ■ Good validity and reliability: <ul style="list-style-type: none"> ■ Multiple choice questionnaires (with good questions) ■ Key features ■ Clinical work sample ■ OSCE ■ Progress test |
|--|--|

Part 3 - Assessment and Evaluation

Assessment versus Evaluation

- **Assessment** - The process of collecting information about a performance, object, group, or institution to *better understand it*



assess means to "sit beside"

Applications

- SWOT analysis
- Tests of preference, style, aptitude
- Grounded theory analysis of focus group data
- Self-study
- Reflection

Assessment versus Evaluation

- **Evaluation** - The systematic collection and interpretation of evidence, leading, as a part of the process, to a *judgment of value with a view to action*



Two Roles for Evaluation

- **Formative (Constructive) Evaluation** - A process of measuring the quality of a performance, work product or use of a process using criteria **to generate feedback aimed at improving future performances or products.**
- **Summative (Judgmental) Evaluation** - A process of measuring the quality of a performance, work product or use of a process against a set of standards to make a judgment or determination if, or to what level, the standards have been met. **Judgmental evaluation results in a consequential decision or score.**

"Assessment" and "Evaluation"

- *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs 2nd Edition, (1994)*
- Prepared by the Joint Committee on Standards for Education Evaluation composed of 16 groups representing American and Canadian educational, evaluation and psychological associations.
- **Assessment:** The act of determining the standing of an object on some variable of interest, for example, testing students and reporting grades.
- **Evaluation:** Systematic investigation of the worth or merit of an object (e.g., a program, project, or instructional material).

"Assessment" and "Evaluation"

- *The Student Evaluation Standards: How to Improve Evaluations of Students (2003).*
- Again, prepared by the Joint Committee on Standards for Education Evaluation composed of 16 groups representing American and Canadian educational, evaluation and psychological associations.
- **Assessment:** The process of collecting information about a student to aid in the decision making about the progress and development of the student.
- **Evaluation:** Systematic investigation of the worth or merit of a student's performance in relation to a set of learner expectations or standards of performance.

"Assessment" and "Evaluation"

- Accreditation Council for Pharmacy Education (ACPE)
- Royal Pharmaceutical Society of Great Britain (RPSGB)
- Canadian Council for Accreditation of Pharmacy Programs (CCAPP)

"Assessment" and "Evaluation"

- Canadian Council for Accreditation of Pharmacy Programs (CCAPP)
- **Assessment:** Processes for grading (formal and non formal), examining and other activities related to measuring learning and skill development in the individual student.
- **Evaluation:** Systematic collection and interpretation of evidence, leading, as part of the process to a judgement of the effectiveness or value of the educational program or any component of the operational structure of a Faculty.

"Assessment" and "Evaluation"

Number of occasions when respective terms "Assessment" or "Evaluations" appear in Standards Documents

Standards Prepared By	Assessment	Evaluation
ACPE ¹	71	63
RPSGB ²	22	5
CCAPP ³	20	36

- 1 Accreditation Council for Pharmacy Education
- 2 Royal Pharmaceutical Society of Great Britain
- 3 Canadian Council for Accreditation of Pharmacy Programs

"Assessment" or "Evaluation"

FROM ACPE STANDARDS:

The college or school must produce and make available to students and prospective students transfer credit and course-waiver policies, based on rational procedures and defensible assessments

Requisites may only be waived based upon an educationally sound assessment of the professional competencies (as set forth in Standard 12) that may have been achieved through continuing pharmacy education, other postgraduate education and training, and previous pharmacy practice experience.

The college or school should have progression policies that take into consideration assessments of professional behavior and academic integrity.

The college or school should obtain assessment of qualities and performance of preceptors from students in a manner that would not adversely affect the grading process.

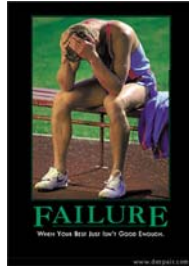
The methods of assessment and reporting employed should promote the development within the student of the ability to offer constructive criticism in a manner appropriate to interprofessional relationships. The assessment should include each preceptor's:

- ability to facilitate learning
- communication skills
- quality as a professional role model
- effectiveness related to pharmacy education

Part 4 – Constructive Evaluation and Quality Culture

We live in a "judgmental culture"

- The consequences of which include:
 - Fear, hypervigilance
 - Risk avoidance
 - Cheating
 - Assessment avoidance
 - Not learning and growing from mistakes
 - Defensiveness
 - Wasted time, wasted lives



We seek a "quality culture"

	Judgmental Culture	Quality Culture
Who initiates	Those in power	Everyone
Focus	Outcomes, deadlines	Progress, timelines
When	Provides closure	Is ongoing
Critical Stakeholder	Other	Self
Schedule	Rigid	Flexible
Output	Good / Bad	Strategies to improve
Outcome	Rewards and penalties	Increasing value
Effect on self esteem	Potentially disconfirming	Affirmation of strengths support for areas for improvement
Consequence of failure	Loss	Learning opportunity

How do we create a quality culture?

- First and foremost, those in power should model seeking feedback for their OWN performances and products:
 - Setting aside time with important stakeholders;
 - Working with them to define quality through the generation of criteria and standards;
 - Proactively seeking feedback from them;
 - Receiving their feedback with grace;
 - Using their feedback transparently.

How do we create a quality culture?

- In constructively evaluating others' performance and products:
 - Involving those who will be evaluated in identifying what will be evaluated and the criteria that will be used
 - Being respectful
 - Focusing on what's important*
 - Most important determinants of success
 - Most critical areas for improvement
 - Giving strategies for improvement, not just weaknesses

* Evaluating one thing well is much more useful than evaluating many things poorly

Group Discussion/Summary

Group Discussion

- Major insights about definitions proposed?
- How do we frame a quality culture?
