# HAWAI'I PACIFIC UNIVERSITY

# GUIDE TO LEARNING ASSESSMENT



Updated 2011-2013

Office of Academic Administration, Planning and Assessment

# TABLE OF CONTENTS

### <u>Note</u>: All documents are posted on PIPELINE, under Academic Administration Tab: Program Review, Assessment and Educational Effectiveness Channel

	<u>P</u>	age
SECTION 1.	OVERVIEW OF LEARNING ASSESSMENT	. 3
SECTION 2.	Completing Curriculum Maps & Assessment Plans & Timelines	4-6
SECTION 3.	CREATING MEANINGFUL ASSESSMENTS FOR PROGRAM REVIEW	7-10
SECTION 4.	TASKS FOR COORDINATING/COMPLETING LEARNING ASSESSMENTS1	1-12
SECTION 5.	LEARNING ASSESSMENT REPORTS	13
SECTION 6.	EXAMPLES OF ASSESSMENT METHODS: RUBRICS: THE AACU VALUE PROJECT	14
SECTION 7.	EXAMPLES OF ASSESSMENT METHODS: ASSESSMENT OF QUALITATIVE ARTIFACTS	5-17

## SECTION 1. OVERVIEW OF LEARNING ASSESSMENT

**Assessment** is defined as evaluation or appraisal. In universities, assessment includes multiple levels (see GUIDE to Learning Assessment, separate document):

- Course-level learning outcomes (learning assessment)
- Program/university learning outcomes (learning assessment)
- Program quality assessment (program review)
- Institutional outcomes assessments (using institutional outcome indicators)
- Course evaluations (students' perception of course quality)
- · Teaching quality assessments and peer reviews

**Learning assessment** to document and analyze the quality of HPU academic programs has focused on program-level learning and has been integrated into the academic Program Review model. This model achieves ongoing learning assessment in programs by requiring four assessments of program-level learning per academic year. These learning assessments address the program-level learning as it is defined in the program's posted learning objectives.

High-quality high-value learning assessments require the faculty to differentiate between teachers' methods of assigning grades in courses and the program's methods of aggregating evidence of learning from multiple sources (not limited to courses) as the basis for completing systematic and objective reviews. These are important distinctions for faculty to make to achieve effective examination of trends in learning and student success at the degree program level. Learning and student success associated with individual courses and/or individual teachers is considered to be a more "internal" dimension of assessment for which faculty are also responsible. An important distinction that has required ongoing clarification and emphasis is that learning assessments conducted by reviewing student work from courses is not the same as assessment of <u>course</u> quality. Program review learning assessments merely involve the identification of courses that can provide learning artifacts that demonstrate learning relative to one or more degree objectives.

Assessment guidelines are provided in this document to assist faculty with creating curriculum maps and assessment plans and with designing and completing learning assessments (such as using rubrics and assessing qualitative artifacts):

Completing Curriculum Maps & Assessment Plans & Timelines

Creating Meaningful Assessments for Program Review

Tasks for Coordinating/Completing Learning Assessments

Learning Assessment Reports

Examples of Assessment Methods: Rubrics: The AACU VALUE Project

Examples of Assessment Methods: Assessment of Qualitative Artifacts

# SECTION 2. GUIDE TO COMPLETING CURRICULUM MAPS & LEARNING ASSESSMENT PLANS & TIMELINES

### CURRICULUM MAPS

A curriculum map shows the alignment of program learning outcomes with required courses in the major. Curriculum maps also increase the internal coherence of degree program curricula by organizing and aligning program learning outcomes with all the learning outcomes within all the required courses in the curriculum.

### **Example of a Program Curriculum Map**

	<b>Objective 1</b>	<b>Objective 2</b>	<b>Objective 3</b>	<b>Objective 4</b>	<b>Objective 5</b>	<b>Objective 6</b>
Course #						
Course #						
Course #						
Etc.						

### LEARNING ASSESSMENT PLANS & TIMELINES

Ongoing learning assessment in the majors is assured by completion of at least 4 learning assessments per year relating to degree learning outcomes (see **GUIDE to Learning Assessment**).

**Program Learning Assessment Plans** align types of student work available in specific courses in the major that can be collected as artifacts/evidence of learning relating to program learning outcomes.

Program Learning Objectives	Key Courses Relevant to Program Objectives (Sources of Student Work for Learning Assessments)	Student Work to be used for Learning Assessments & Method (Papers, exams, portfolios, etc.)
Objective 1.	JADMxxx	Term Paper
	JADM xyz	Case study
Objective 2.	JADM abc	Final exam
	JADM frc	Term Paper
Objective 3.	JADM xxx	etc
	etc	
Objective 4.	etc	etc
Objective 5.	etc	etc
Objective 6	etc	etc

Example of a Program Learning Assessment Plan

**Learning Assessment Timelines** show the schedule of assessments the faculty are conducting each year (showing courses & artifacts of student work to be sampled from the courses). The course learning artifacts align with degree outcomes to assess learning relevant to these degree outcomes.

Courses in which student work is being collected:								
1) for degree	Fall 10	Spring 11	Fall 11	Spring 12	Fall 12	Spring 13	Fall 13	Spring 14
objectives								
(list which								
objectives)								
2) additional	Fall 10	Spring 11	Fall 11	Spring 12	Fall 12	Spring 13	Fall 13	Spring 14
courses								
supporting								
the major or								
other majors								

Example of a Learning Assessment Timeline

### Designing the Program's Approach to Learning Assessments

Multi-section and online courses are given priority when selecting courses as sources of learning artifacts for assessments. Whenever relevant to the program, assessments are structured so as to document comparability of learning with respect to (a) learning in online and classroom modalities of courses; and (b) learning in courses offered Downtown and in Military Campus Program locations. Additional assessments are conducted for the general education program.

A teacher's grades are not sufficient for the learning assessments for program review. If a teacher has used a rubric or rating scales to grade student work, these ratings may be collected. However, "external" or peer review of a substantial sample of the student work must be completed to fulfill the principle of an *objective* view of the student work.

Learning assessments on student work from courses is not the same as assessing the quality of <u>courses</u>. It is merely the identification of courses that can provide learning artifacts that demonstrate learning relative to one or more degree objectives.

Multi-section courses are often the source of student work and sections will be offered during various terms. So collection of samples of student work needs to be planned ahead in order to have the student work ready to complete a learning assessment. Or it will need to be completed in stages in order to include the respective sections. Multi-section courses may yield varying forms of student work. To complete the learning assessment across multiple sections, a generic scale or rubric may be needed to estimate the extent to which each artifact meets the degree objective.

Note that the goal is to achieve equivalent forms of student work from multi-sections courses, to assure quality of <u>learning *opportunities*</u> as well as assessment of the <u>quality of learning</u>.

### Suggested Work Plan for Developing the Assessment Plan and Timeline

1. <u>Start</u> by reviewing all the upper-division and 2000-level courses in the discipline that serve the degree and/or other degrees to review when/where scheduled.

#### SECTION 2. - CONTINUED

2. One approach is to create two tables each for the Assessment Plan & the Timeline.

• For the <u>first</u> Assessment Plan table, <u>each</u> degree objective is matched to a course that produces student work that enables assessment of that learning objective. Courses often fit more than one objective; make the best possible matches on student work that will enable the assessment using a "best-fit" approach.

• For the <u>second</u> Assessment Plan table, define 6-8 more (different) courses that are sources of student learning relevant to the degree and/or to other degrees. These courses should also be matched to degree objectives. If the program provides courses that serve another degree, the program is responsible for assessing learning in these "service" courses.

- For the <u>Timeline Tables</u>, first consult the total list of course sections and when/where all the sectons are offered. Select the relevant courses. Then arrange the courses across the timeline to show 2 assessments scheduled in each term. Note that the plan for collection of student work will begin 1-2 terms before the assessment is scheduled to be completed.
- 4. For each course/assessment, designate a <u>faculty lead person</u> to organize selection of course sections that provide samples of the necessary comparisons (as possible) on location and modality. This faculty member is responsible for contacting the various instructors and following through to collect the student work.
- 5. The sampling process depends on the total number of sections and the number of possible papers. There should be at least 10 papers from each section. It may be useful to collect all the possible samples and then select 10 at random from each section. If the faculty propose and document a different sampling method, this may be equally valid for the assessment activity. The sampling process should be included in the Assessment report.
- 6. Completion of assessments may need to be "rolled" across 2-3 terms, if multiple terms are required to access the work. That is, faculty may complete an assessment on the sets of student work available from some sections and complete the assessment in a subsequent term when remaining work can be added to the analysis.
- 7. The <u>program review chair</u> is responsible for (a) assuring that 2-3 or more faculty are convened to complete the learning assessment, (b) reporting the assessments to faculty, and (c) assuring that faculty review the results and discuss the meaning and potential opportunities for improvements based on the assessment data. *See the posted guidelines for completing learning assessments (Creating Meaningful Assessments for Program Review).*

# SECTION 3. CREATING MEANINGFUL ASSESSMENTS FOR PROGRAM REVIEW

Meaningful assessment of student learning for program reviews involves faculty in addressing student learning evidence, as well as evidence about the program's teaching-learning academic quality that is associated with student learning. This work additionally involves realistic consideration of the workload implications and faculty decisions about managing the workload associated with Program Review. Ultimately, the work of each Program Review feeds into the collective work that is to be done by the University to demonstrate compliance with higher education quality standards. This Section addresses the following:

- I. Differentiating Multiple Levels of Learning Assessment
- II. Defining Sources of Evidence of Academic Quality Relating to Student Learning
- III. Workload & What it Takes to Complete Learning Assessments for Program Review
- IV. The Collective Work to be Done by the University

### I. Multiple Levels of Learning Assessment

- Differentiate between levels of learning to be assessed at course, program & institutional levels:
  - 1. Student performance in a course by an instructor to assign a grade. And/or: Student performance in multi-section courses to evaluate how the course is going (re course objectives & course quality).
  - 2. Student performance to document learning evidence relating to a degree objective (for program review).
  - 3. Student outcome data to document student success; can relate to course, program, college, or institutional level. Student success includes course completions, course drops, grades, graduation time and rates, etc.
- Also differentiate <u>student learning</u> from <u>course quality</u> from <u>teacher competence</u>:
  - > Student learning is assessed by looking at student work & performance.
  - Courses provide basis for the student work, but assessing quality of a <u>course</u> requires different methods (examples: syllabus reviews; QM rubric for online courses).
  - Assessment of teacher competence is established by the faculty and may include in-class observations, student evaluations and other methods to be defined.

Level of Learning Assessment	Methods
1. Assessment of student performance in a course by an instructor, to assign a grade. A teacher assesses student performance & assigns a grade. Based on student performance, student feedback & teacher insightfulness, there may be course improvements based on this assessment.	Based on course objectives, teacher-made criteria, rating scales and/or rubrics are used to evaluate student work.
1 cont. Assessment of student performance to document learning in multi-section courses to evaluate how the course is going (re course objectives & course quality) Teachers assess student performance to look at learning within a multi-section course.	<ol> <li>Assemble samples of student work (from previous terms) from various course sections representing modalities/localities as is relevant.</li> <li>Rating scales or rubrics are used to assess the student work relative to a <u>course</u> objective</li> <li>Results are tabulated &amp; summarized, &amp; reviewed by faculty to determine whether improvements are needed</li> </ol>
2. Assessment of student performance to document learning evidence relating to a degree objective (for program review).	<ol> <li>Assemble samples of student work (from previous terms) from various course sections representing modalities/localities as is relevant.</li> <li>Rating scales or rubrics are used to assess the student work relative to a <u>program</u> objective</li> <li>Results are tabulated &amp; summarized, &amp; reviewed by faculty to determine whether improvements are needed</li> </ol>
3. Assessment of student outcome data to document student success; can relate to course, program, college, or institutional level.	<ul> <li>Ranges across:</li> <li>1. Defined learning expectations</li> <li>2. Learning opportunities</li> <li>3. Student engagement</li> <li>4. Student performance/student work (#3 above)</li> <li>5. Feedback given to students for improvement</li> <li>6. Institutional indicators of student success</li> </ul>

### II. Sources of Evidence of Academic Quality Relating to Student Learning

- 1. Clearly defined expectations for learning & performance & related criteria
- 2. Relevant learning opportunities
- 3. Student engagement
- 4. Student performance / student work
- 5. Feedback to students for improvement
- 6. Institutional indicators of student success

Sources of Evidence of Academic Quality	Assessment Methods to Generate Relevant Evidence (Accumulate Over Time)
<ol> <li>Clearly defined expectations for learning &amp; criteria</li> <li>Learning outcomes defined, published &amp; widely shared (WASC)</li> <li>Equivalent across locations &amp; modalities as is possible</li> </ol>	<ol> <li>Curriculum Map – Show alignment of course objectives/content with program objectives</li> <li>Syllabus reviews – Examples include: quality of learning experiences, level of writing expectations, other factors defined by faculty.</li> <li>Standard syllabus created for multi-sect courses</li> </ol>
2. Relevant learning opportunities Learning opportunities are relevant to course objectives & require/support student engagement Equivalent across locations & modalities	<ol> <li>Syllabus analyses by faculty teams</li> <li>Student self reports on extent &amp; quality of engagement</li> </ol>
<ul> <li>Cont - Expectations for student work are:</li> <li>1. Relevant to course objectives</li> <li>2. Observable &amp; attributable to individual students (can be shared for review by a peer/external reviewer in form that can be communicated</li> <li>3. Appropriate for course level based on faculty expectations in the program</li> <li>4. Comparable across multi-sections of courses</li> <li>5. Equivalent across locations/modalities</li> </ul>	<ol> <li>Available (saved): Tests, papers, in-class work, presentations, field trip and project reports</li> <li>Peer reviews</li> <li>Syllabus reviews</li> </ol>
<b>3</b> . <b>Student engagement</b> Office hours / other forms of contact with instructors such as field trips, projects. Equivalent across locations & modalities as possible	<ol> <li>Syllabus review</li> <li>Walk-around monitoring</li> </ol>
<b>4</b> . <b>Student performance: Student work</b> Learning as evidenced in coursework meets faculty expectations & is validated by peer review Equivalent across locations & modalities as possible	Learning assessments completed on student work
<b>Student performance: Grade distribution</b> Grade distribution. Equivalent across locations & modalities as is possible	Reviewed by program faculty who decide on standard or criteria for concern
5. Feedback to students for improvement	? Review of feedback offered on assignments
6. Indicators of Student Success Equivalent across locations & modalities as is possible. Retention to next term/year. Graduation rates. Gainful employment & happiness.	Can look at many aspects of retention – how did your students do last year compared with HPU average? Student surveys. Or School-wide or program specific

### III. Workload & What it takes to Complete Learning Assessments for Program Review

### 1. Create Learning Assessment Plans and Timelines – Tasks Include:

- Lead/ organize the work & engage other faculty (including MCP)
- Review list of all sections of all courses in the discipline past year, this year, and next term
- · Assemble syllabi from key courses to consult as needed
- Design a plan that aligns learning & student work in the various <u>courses</u> with <u>degree objectives</u>.
- Or construct a "curriculum map" that aligns all course objectives with degree objectives.
- Create the timeline after the plan is completed (schedule of assessments & collection student work)

### 2. Complete 2 Learning Assessments Per Semester, Four Per Year

- Collection the student work has to be completed in term before assessment conducted
- Faculty decide student work to request & if samples (at least 7-10/section, multi-section courses); student work is what faculty decide is relevant evidence of 1 or more degree objectives.
- Requests to relevant faculty (include online, MCP) to save student work (& assignment instructions)
- Develop rating scale (or rubric or criteria) with faculty for reviewing the student work
- Make copies of samples of student work so inter-rater agreement can be established (number the artifacts; make copies of rating scales for reviewers to make ratings)
- Faculty meet to review the student work; take time to review task, scale (or rubric), work to be reviewed. Review some samples of work to establish inter-rater agreement.
- Collate & summarize results
- · Analyze results & evaluating needs for improvements through faculty meeting discussion
- Document discussion & results including strengths & weaknesses, & whether concerns relating to comparability of online/classroom or Downtown/MCP, if these were part of the review.
- Document improvements made, if needed
- Enter information into Learning Assessment Report Template & TracDat; forwards report to appropriate entities.

### IV. The Collective Work to be Done by the University

- 1. Deans & program faculty groups meet periodically to establish the commitment to completing learning assessment plans, timelines, and assessment reports. This is a leadership/management challenge and opportunity.
- 2. Deans & program faculty groups meet at least once each term to review and <u>document</u> that enrollment and retention data have been reviewed and consideration given to actions that may be needed to increase student success.
- Deans/departments document the collective review of institutional-level data each semester, including giving consideration to actions that may be needed to increase student success: enrollment and retention graduation survey end-of-course evaluation results
   Course completion rates

# SECTION 4. TASKS FOR COORDINATING/COMPLETING THE LEARNING ASSESSMENTS

 <u>Collect the Learning Artifacts (Student Work) for Assessments</u>. Consult schedule of classes past and present to decide when student work can be collected from the instructors in the various class sections available. If multiple sections are available, probably best to sample 10 or so artifacts from each section. Create timeline to schedule assessments and collection of student work.

Contact each instructor and request the student work (and instructions to students for the assignment) be saved, indicating which term and what student work. Note that online courses can often be sampled from previous terms because the student work is still in the system and instructors can retrieve it. May requires multiple requests to instructors in course sections, to assure that student work is collected on time. Follow up as needed to obtain the student work samples. The MCP Curriculum Area Liaison (CAL) can assist with this process. The student work must be collected in an earlier term so that it is <u>ready</u> for the learning assessments sessions in the current term.

### 2. Develop the Rating Scale/Rubric or Assessment Criteria.

- a) For each assessment, meet with faculty who will review the student work. Orient the group to the task, development of the ratings, and the work to be reviewed.
- b) Make copies of samples of student work so each reviewer has copies of the work to be reviewed to address inter-rater agreement.
- c) The artifacts are discussed in light of the assignment instructions and the degree objective to be assessed.
- d) The group designs the scale/criteria to be applied in the assessment. (See other resources posted on PIPELINE for guidelines for developing rubrics and rating scales.) Make copies of the rating scales on sheets of paper so each rater can make a rating.
- e) Some samples of work are "double reviewed" and then discussed to clarify the way the scales are to be applied (inter-rater agreement).

### 3. Complete the Assessment Ratings With Faculty Reviewers.

- a) If good agreement is demonstrated, individual reviewers can proceed to review student work so long as it is not from their own course sections.
- b) Collate and summarize the ratings into the Learning Assessment Report format. Review results in a faculty meeting, preferably by the full faculty group. Data charts and other additional data can be attached to the report.
- c) Discuss with program faculty to acknowledge the strengths & weaknesses, including whether there are any concerns relating to online/classroom or Downtown/MCP comparisons, if these were part of the review. Discussion also defines improvements as needed,
- d) Complete the report using the Learning Assessment Report Template and forward the report to appropriate entities.

### 4. Relating Workload to Compensation – Resources to Support Program Review.

1. One estimate of the workload for teaching a 3-credit course is a total of 145 hours for the semester: <u>each week</u> requires; 3 hours in class, 3-4 hours of preparation and grading papers and 2 office hours = 9 x 15 weeks = 135; in addition, add 10-15 hours for grading exams/papers at middle and end of the term to a total of 145 hours.

2. One estimate for <u>leading the learning assessment work</u> is as follows: development of the Learning Assessment Plan/Timeline = 5 hours; coordination and completion of one learning assessment = 18-20 hours. In one year's time, development of the plan and coordination/leadership of the required 4 assessments would be about 85 hours.

- 3. For faculty members who are fast-tracking the assessments to complete the program review report:
  - a) Learning Assessments It should be possible to complete a total of 7 learning assessments in one term along with the Learning Assessment Plan/Timeline ( $20 \times 7 = 140$  hours + 5 = 145 hours);
  - b) Program Review Report It is likely that completion and coordination of the data and information analyses, writing the report, and coordinating the external review will be equivalent to one course release or stipend.

4. It is <u>assumed</u> that other program faculty will help by contributing some hours to the review of student work.

5. It is <u>budgeted</u> for adjunct faculty to be paid at \$25 per hour for helping with assessments. This must be approved in advance. See procedure for using adjuncts with assessments.

## SECTION 5. FORMAT FOR REPORTING LEARNING ASSESSMENTS

### **REPORT OF LEARNING ASSESSMENT**

<u>Instructions</u>: Complete in Word document and maintain record in Program Review Portfolio in colleges. Expect updates as Blackboard system comes up.

- 1. Program:
- 2. Program Review Chair:
- 3. Degree objective addressed by assessment:
- 4. Brief description of student work/artifact:
- 5. Course # and title: Number of students included: Number of sections sampled: Comparisons of location (MCP vs DT/HL)? (concern? yes / no / explain) Comparisons of modality (Classroom or Entirely Online)? (concern? yes / no / explain)
- 6. How was the student work reviewed 1 faculty member, group, etc? What criterion or rubric was used (can be attached or pasted on p.2):
- 7. Summary of results include % meeting criterion/standard
- 8. Conclusion(s) what concern found (or no concern)
- 9. Actions to be taken and/or improvements to be made (if any):
- 10. Implications for program capacity and/or resources:

# SECTION 6. EXAMPLES OF ASSESSMENT METHODS: <u>RUBRICS</u>

Rubrics are multidimensional scales used to rate student learning artifacts to assess the quality of learning demonstrated in that artifact. The general properties of rubrics include the following:

- 1. Rubrics must use multiple rating scales to assess multiple dimensions of learning, such as the quality of writing, comparisons/contrasts, adequacy of references, use of examples/applications, etc. The dimensions are defined to address the various types of learning that are being evaluated.
- 2. The results of faculty ratings of student work are aggregated separately for each dimension of learning so that each dimension can be considered separately as to whether the student learning is adequate or improvements may be needed.
- 3. It is not acceptable practice to rely on one rating scale for a learning objective. It is also not good practice to create a "total score" for a rubric by adding up the dimension ratings to one total. In either case, the holistic rating eliminates the opportunity to evaluate student learning on the various dimensions.
- 4. The rating scales are essentially the same as Likert scales used in social science and typically the scales have 4 points. As such, the results of rubric ratings are shown as numbers and percentages at each level as well as for each dimension. It is not mathematically acceptable to compute mean scores for these scales.

There is an extensive literature available on the construction and use of rubrics. One excellent set of rubrics and supporting information about the use of rubrics has been published by the Association of American Colleges and Universities (AACU). These

http://www.aacu.org/value/rubrics/index.cfm

Intellectual and Practical Skills

- Inquiry and analysis
- <u>Critical thinking</u>
- <u>Creative thinking</u>
- Written communication
- Oral communication
- <u>Reading</u>
- Quantitative literacy
- Information literacy
- <u>Teamwork</u>
- <u>Problem solving</u>

### Personal and Social Responsibility

- <u>Civic knowledge and engagement—local and global</u>
- Intercultural knowledge and competence
- Ethical reasoning
- Foundations and skills for lifelong learning
- <u>Global learning</u>

Integrative and Applied Learning

• <u>Integrative and applied learning</u>

## **SECTION 7. EXAMPLES OF ASSESSMENT METHODS:**

# GENERIC ASSESSMENT SCALES/RUBRIC FOR QUALITATIVE STUDENT WORK

This document presents one approach for the challenge of assessing qualitative student work collected from multi-section courses. Qualitative student work may be in the form of term papers, essays or essay items on examinations, or presentations in oral and/or formats.

<u>The Process</u> – Here is an example of the flow of work for faculty conducting program review learning assessments with qualitative student work:

- Faculty collect student work from multi-section courses when implementing the learning assessment plans for degree objectives.
- A major assessment challenge is encountered if the teachers/sections of that course use various written or presentation assignments instead of a *standard assignment*.
- The result is that student work samples collected from the various sections include different kinds of qualitative assignments.
- This variance adds to the challenge of addressing comparability of learning in <u>Downtown and MCP</u> sections as well as <u>classroom and online</u> sections.

<u>The Generic Model</u> - One idea for constructing the assessment instrument is to work from a generic model that faculty can use as the foundation for the assessment method. The generic model can be made more specific to the teacher's instructions that guided the student work on the assignment that is being assessed. But its value is that it can be used generically and applied to different assignments across different sections of the same course.

<u>The example on the next page</u> shows use of 4-6 evaluation criteria that align with the assignment instructions and can be used to rate student performance on the <u>assignment</u>. The levels of quality for each of the criteria may be left as general terms or modified with specific words to create a <u>rubric</u>.

<u>Which is better? a rubric or a rating scale</u>? The correct answer is that <u>either</u> can be used, so long as the assessment activity includes some preliminary faculty work to address <u>interrater agreement</u>.

What's the difference? <u>Rubrics</u> = rating scales + words that describe each level on the scale <u>Rating scales</u> = the words are optional; the numbers range from low to high Typical assessments use 4-point scales (for example: 0-1-2-3 or 1-2-3-4)

Is a rubric better? Not necessarily! Social scientist have long debated whether numerical rating scales are more *valid* if words are attached to the numbers to explain what the number means. Social scientists also debate whether a scale should have 4 categories, or 5 or 7 and whether an even number is better than an odd number, *But this debate has not reached a definite conclusion*.

### SECTION 7. - CONTINUED

<u>Interrater Agreement</u> – How do you increase the validity of the assessment? By spending some time at the beginning of an assessment meeting to discuss inter-rater agreement.

- Learning assessments for program review require that samples of student work be reviewed/rated by an instructor other than the teacher in the course section from which the student work was sampled.
- Before starting to rate the work of another instructor's students, each of the faculty review the same 5-10 artifacts and record their ratings. Then they discuss each paper and their ratings and agree on how the ratings will be applied. This step is entered into the written report of the assignment.

### Example of Generic Assessment Criteria & Scales [Can be made more specific for assignment]

- 1. Sufficiency of content/ substance / depth, such as length or number of words
- 2. Quality of content areas, topics, use of theory (or description, analysis)
- 3. Critical thinking (quality of synthesis, critique, argument, comparative analysis, depending on assignment)
- 4. Application rating, or values or ethics
- 5. Use of sources, references; information literacy rating (finding, citing, evaluating, as required)
- 6. Quality of writing (organization of work, sentences and paragraphs, format, and/or appearance/presentation)

### Sample Format for Faculty Ratings Using a 4-point Scale

Artifacts are numbered (names can be removed)

Instructor: :or Section \_\_\_\_\_ Student # \_\_\_\_

### Ratings are from 0 to 3 (4-point scale):

- 4 = Exceptional or excellent work (might be equivalent of "A" grade)
- Exemplary; performance characteristics reflecting the highest level of performance. 3 = Good work (might be equivalent of "B")
- Accomplished; performance characteristics reflecting mastery of performance.
- 2 = Acceptable work (might be equivalent of "C")
  - Developing; performance characteristics reflecting development and movement toward mastery of performance.
- 1 = Poor or missing (might be equivalent of a D or F grade).Beginning; performance characteristics reflecting a beginning level of performance.

PROCEDURE: Read instructions from instructor & think re what student thought was expected in paper.

### **CRITERIA FOR 1-4 RATINGS** (DIMENSIONS OF LEARNING):

Rating

- **#1 Length/depth** (relative to expected length; this is an indicator of "depth")
- #2 Use of Theory or Topic Development How a topic is developed or use of theory or frameworks to analyze a problem or topic
- **#3 Argument or Comparison/Contrast** Use of critical thinking to develop an argument or present a comparison/contrast re topic development in #2
- **#4 Application or Rating Relating to Quality of Work on Values or Ethics** Write a brief explanation here...
- #5 Use of Sources, References Information literacy rating
- #6 Writing Clarity, grammar, flow, syntax

#### Section 7. - Continued

<u>Collecting the Ratings</u> - Save the faculty assessments (ratings) in a grid/table format: Faculty record their assessment ratings of each piece of work on a separate slip of paper that identifies the work evaluated (you can have number coded the paper and the course number and section and term. Create a cumulative record of the ratings in worksheet.

<u>Collating and Analyzing the Results</u> – Do not calculate a mean when using 4-point rating scales! Present results as frequencies with percentages at each level of the criterion. The summary of results should show a total, a result for the Downtown/MCP comparability and a result for the Online/Classroom comparability. The faculty may set a standard for the percentage they want to see at a certain level of the scale for each criterion. For example: "80% of students will be at least at the 3rd level using a 4-point scale; 70% will be at least at the second level.