Final Report of the
Ad-hoc Committee for Review of 2009 Academic Program Learning Outcome Assessment Reports
September 2, 2010
Prepared by Laura E. Martin
WASC Coordinator

Introduction

This report summarizes the work and results of the ad-hoc committee to review the 2009 Academic Program Learning Outcome Assessment Reports (PLO Reports) formed under the auspices of the WASC Steering Committee in Spring 2010, and charged to 1) provide formative feedback to individual academic programs regarding their first PLO assessment efforts, 2) establish baseline data describing the degree to which academic assessment efforts are developed as evaluated using the *Rubric for the Report on PLO Assessment* (Appendix A), 3) identify common assessment or student learning-related strengths, weaknesses or potential issues as potential foci for further study or action, and 4) identify program assessment efforts to be highlighted in the EER Report.

The committee reviewed twenty-three PLO Reports, representing 15 undergraduate majors, six stand alone minors, one graduate program, and Core 1 (Appendix B). Reports for four undergraduate majors and 2 stand alone minors[1] were unavailable at the time of review. The committee's conclusions, results, and methods follow.

Conclusions

As represented in the *Rubric for the Report on PLO Assessment,* a foundational purpose of annual assessment and reporting is to cultivate an intentional, transparent, collaborative, and thereby programmatic approach to fostering student intellectual development within a degree granting program. Over time, these activities should result in a well-aligned and integrated set of courses (curriculum) taught to support student achievement of a comprehensive set of program learning outcomes (PLOs), for which the criteria and standards of student performance are detailed more specifically through instruments like programmatic rubrics.

The *Rubric for the Report on PLO Assessment* supports development of these attributes through criteria and standards that ask faculty to

- Collaboratively develop a shared set of explicit expectations for student intellectual development by graduation (*Assessable PLO* and *Valid Evidence* criteria).
- Collaboratively develop and apply tools to investigate student development and achievement of these expectations that should be a product of the program's curriculum rather than a single course (*Valid Evidence* and *Reliable Results* criteria).
- Take action as a program to improve student intellectual achievements and to refine the assessment practices (*Results Summary* and *Conclusions and Recommendations* criteria).

[1] Majors: Bioengineering, Economics, Management, and Materials Science and Engineering. Minors: American Studies and Service Science. These latter programs have fewer than 5 students each.

- Record these actions and decisions to allow current and future faculty to connect their own course curricula and instructional practices to these results as appropriate (all criteria).

Programs should be working to reach the *Developed* standard.

Based on the aggregated results of report reviews, the ad-hoc committee came to the following conclusions related to the degree of development of assessment practices across academic programs and associated needs.

1) Two notable strengths of the reported assessment practices were the widespread use of assessment results to inform curricular change and, for many programs, the degree of faculty involvement in programmatic assessment activities.

2) For most programs, continued development is needed to achieve a transparent, collaborative, programmatic approach to assessment. Development efforts should be focused primarily on the practices associated with three foundational criteria of the rubric - *Assessable PLOs*, *Valid Evidence*, and *Reliable Results*.

3) A majority of programs need to develop one or more lines of indirect evidence of student learning that complements the direct evidence gathered.

4) To support development, programs should continue to evaluate their annual assessment practices by using the *Rubric for the Report on PLO Assessment.*

5) To support development, efforts should be made to further establish a common assessment-related vocabulary and set of expectations and tools across academic programs.

Related to student learning, the committee recognized that

6) Communication skills, particularly writing, may benefit from development efforts focused at the disciplinary level. Further investigation is warranted.

Additionally, the committee noted that it will be important to

7) Update the annual PLO Assessment Report format to ensure that programs archive with their report the rubric used to evaluate student work together with representative examples of scored student work. This will support calibration and comparison when revisiting the PLO as well as the examination of student learning achievements during program review. WASC will also expect this material to be available in future accreditation reviews.

8) Remind programs that have revised PLOs to update catalog, websites, and syllabi and consider the implications of these modifications for the alignment of the program's curriculum across courses and for course-level student learning outcomes.

Finally, the committee identified the following programs as potential candidates to be highlighted in the Educational Effectiveness Review Report: Core 1, Environmental Engineering, History, Mechanical Engineering, Physics, and Spanish Minor. The committee recognized that the final selection would depend upon the larger context of the report.

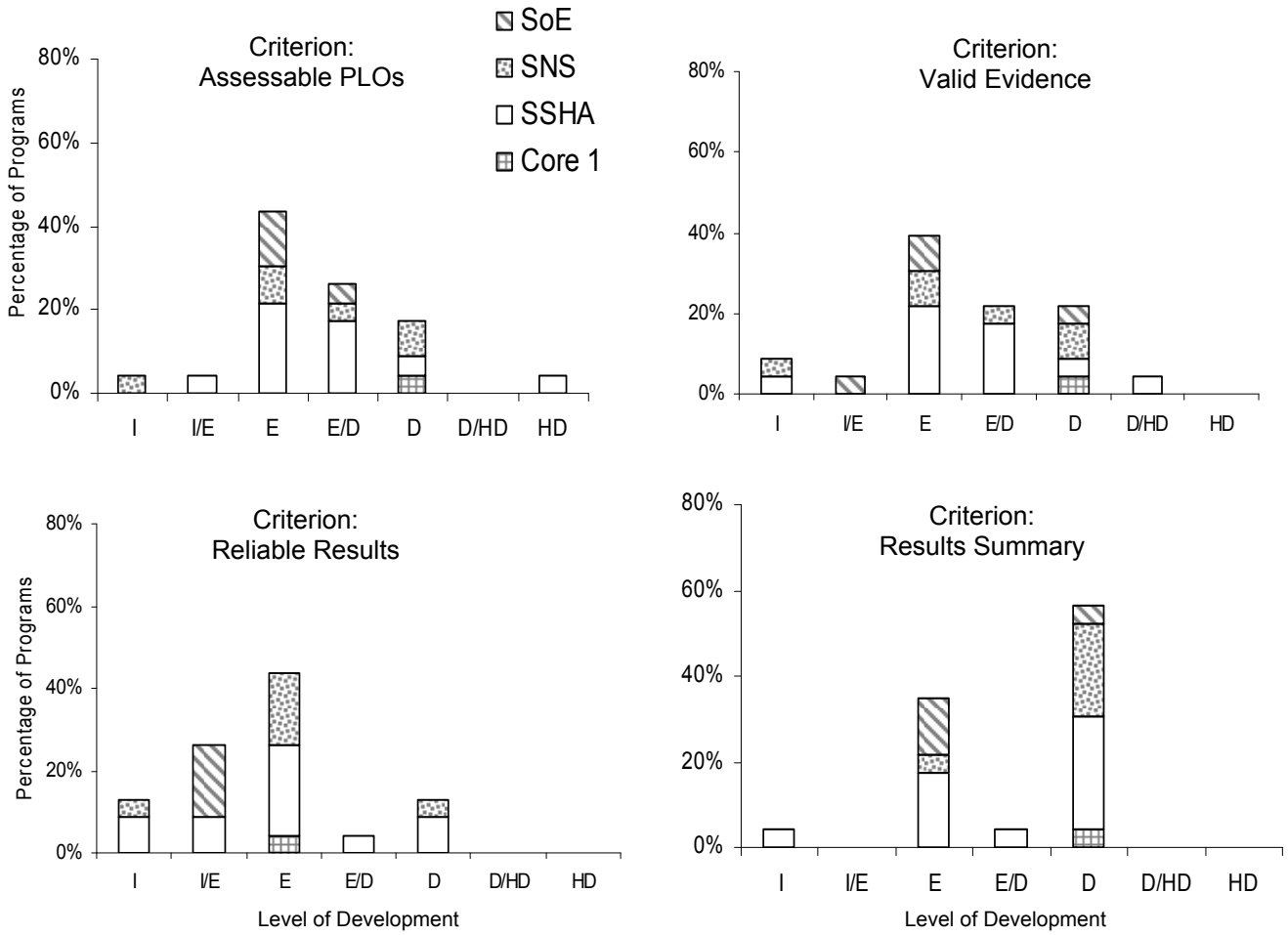
Specific details related to be the above conclusions are available the Results section of this report.

Results

Establishing Baseline Data on Program Assessment

PLO Reports communicated a wide diversity of assessment-related strengths, many of which were unique to individual programs. Broad faculty involvement was the most frequently recognized strength; it was noted in response to 39% of the reports.

Figure 1 summarizes the degree of program development for each of the rubric's five criteria, in the aggregate and by School, as assessed by the ad-hoc committee.



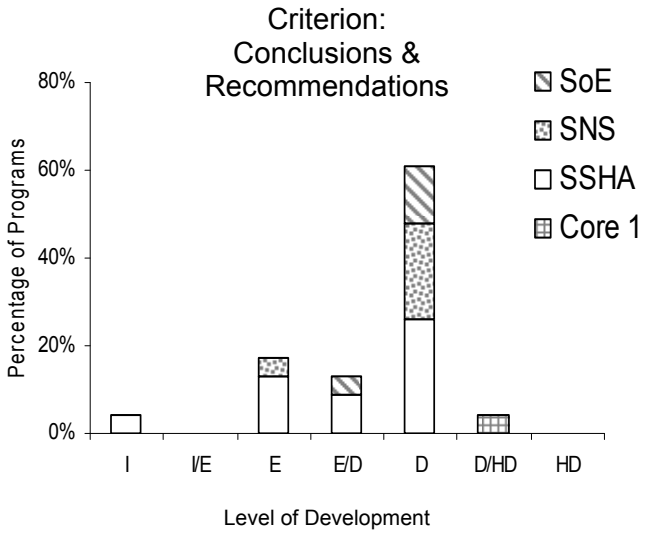


Figure 1. Results of the committee-based assessment of the level of development of the 2009-2010 Program PLO Reports as evaluated using UC Merced's *Rubric for the Report on PLO Assessment*. For each of the rubric's five criteria, the percentage of reports judged to be at each level of development - Initial (I), Emerging (E), Developed (D), Highly Developed (HD) or intermediate to two levels (for example, I/E) - is summarized below by School (n=23).

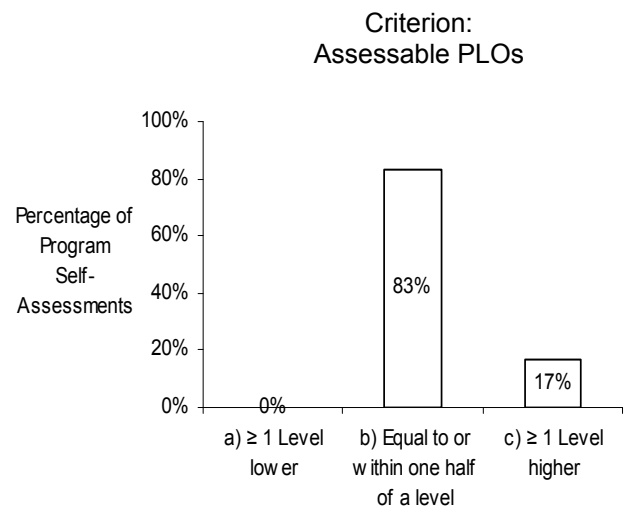
These results indicate that

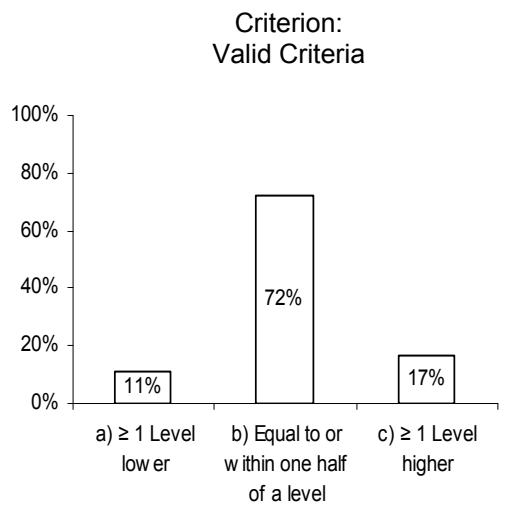
1) In the aggregate, UC Merced's academic programs are moving toward *Developed* assessment structures and practices, with most reporting programs assessed to be *Emerging* or better for all five of the rubric's criteria.

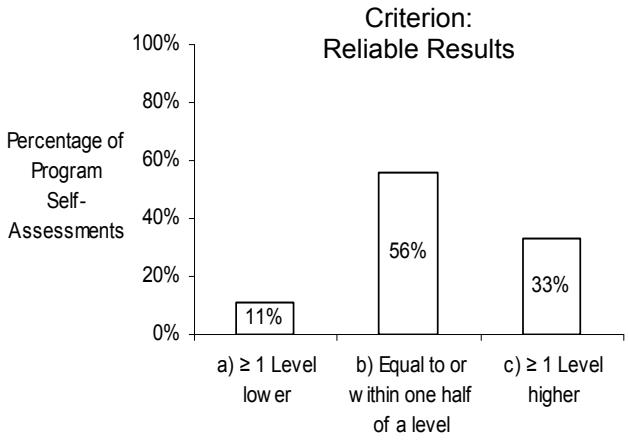
2) The criteria associated with the key foundational practices of transparent, collaborative programmatic practices - *Assessable PLOs*, *Valid Evidence*, and *Reliable Results* - are most in need of development, with approximately 50% of programs assessed as *Emerging* or lower in these categories.

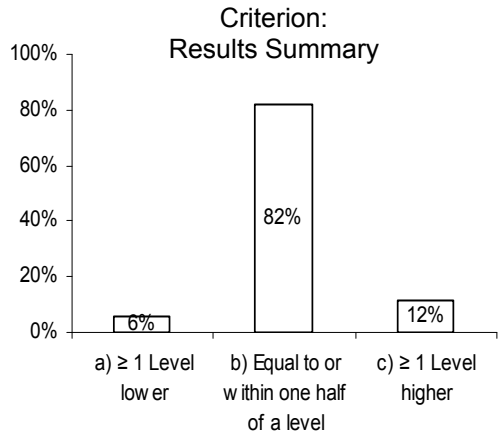
3) Assessment results are being connected to curriculum and instructional practices even in these earliest stages, as a majority of programs (~60%) reached the *Developed* or *Highly Developed* standard with respect to the *Results Summary* and *Conclusions and Recommendations* criteria. Indeed, approximately 90% of reporting programs identified changes to be made to the curriculum as a result of the assessment process.

Figure 2 summarizes the degree of agreement between the committee's assessment of program development and each program's self-assessment with respect to the rubric's five criteria.









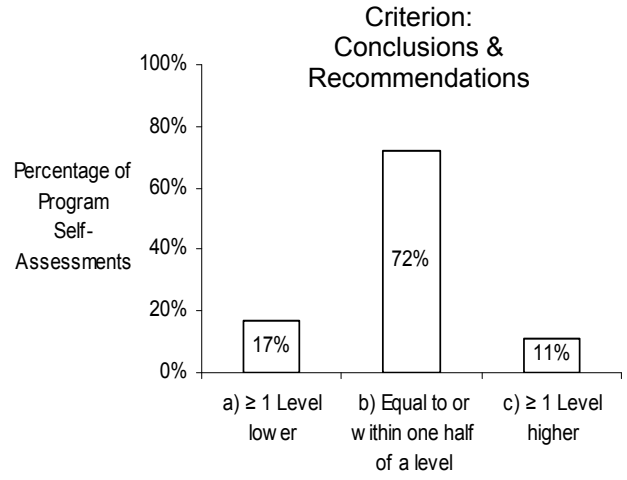


Figure 2. Comparison of program and committee assessments of PLO Reports using UC Merced's *Rubric for the Report on PLO Assessments*. For each of the rubric's five criteria, a figure below describes the percentage of program self-assessments were a) at least one level of development lower than the committee's assessment, b) the same as or with-in one-half level of committee's assessment, and c) at least one level of development higher than the committee's (n=18 for all criteria, except Results Summary with an n=17).

The strong agreement between committee and program assessments for nearly all criteria

1) Lends support to the conclusions regarding the current level of assessment-related development among reporting academic programs.

2) Indicates that the rubric is supporting the development of a shared, foundational, understanding of the expectations of assessment practices and reporting across academic programs.

The agreement between committee and program assessments was weakest for the *Reliable Results* criterion, with a third of program self-ratings exceeding those of the committee by at least one level of development. Allowing for the possibility that low committee scores and relatively high program self-ratings may reflect abbreviated reporting rather than actual practice, the results suggest that many programs need to formalize the calibration (norming) process. This will increase confidence in the conclusions of assessment work and promote shared understandings that feedback into course-based instruction to better support achievement of programmatic curricular goals.

To more precisely identify the assessment practices in need of development, Table 1 provides the most common suggestions made to programs and their relationship to the three criteria specifically identified for attention - *Assessable PLOs*, *Valid Evidence*, and *Reliable Results*. Suggestions 1, 2, 5, and 6 describe practices essential to establishing a collaborative, programmatic approach to improving student learning through assessment.

Table 1: The most common suggestions for improving assessment practices in relation to the rubric criteria each supports.

Committee Suggestion	**Rubric Criteria Addressed by Committee Suggestion**		
	Assessable PLOs	**Valid Evidence**	**Reliable Results**
1) Further elaborate rubrics to describe more precisely the criteria and standards used to evaluate student work so that a) faculty engaged in the assessment process are using shared rather than personal criteria to evaluate student work, and b) faculty, including future faculty, who are not directly involved in the assessment process can see their colleagues' expectations for student learning in relation to a PLO.	x	x	
2) In the absence of a capstone course or assignment, strategically select examples of student work from a number of courses so that the assessment results enable conclusions about the program's efficacy in bringing about student learning as opposed to that of a single course. Alternatively, explain why student work from a single course is expected to represent learning facilitated by the program's curriculum, rather than just that of the course.		x	

Committee Suggestion	Rubric Criteria Addressed by Committee Suggestion		
	Assessable PLOs	Valid Evidence	Reliable Results
3) Consider more explicitly the alignment between the assessed student work and the PLO. Do the assignments elicit student responses that allow the program to evaluate achievement of the PLO?	x	x	
4) Develop lines of indirect evidence to complement and support results of direct assessment. [2]		x	
5) Formalize calibration processes in order to increase confidence in conclusions and develop shared understandings to take back to individual courses.			x
6) Determine inter-rater reliability in order to, as necessary, take steps to improve agreement among faculty and confidence in conclusions.			x

Emerging themes

The committee identified several recurrent issues or themes related to both student learning and assessment.

With respect to student learning, several programs noted that

- Student written communication skills and mathematical expression confounded the assessment of student knowledge and intellectual skills.

With respect to assessment, the committee noted the need to

- Develop a programmatic approach to assessment that involves identifying student work that reflects learning due to the program's curriculum rather than that of a single course.
- Develop explicit rubrics that are appropriate/well-aligned to the assessment task.
- Continue to address the challenge of assessing forms of evidence that are not easily susceptible to criteria, for example, complex narrative.
- Develop indirect evidence of student learning that complements direct evidence.
- Improve faculty calibration.

[2] Only 43%, of the 95% of programs that worked with direct evidence, used some form of indirect evidence as well. The absence of indirect evidence contributed in part to approximately 60% of programs being identified as emerging or emerging/developed for the *Valid Evidence* criterion. Programs that gathered indirect evidence, particularly through student interviews, typically described these results as being very informative and as practices that will be continued. In several cases, student perspectives on the program independently supported faculty observations or conclusions based on direct evidence of student performance. WASC also expect multiple lines of evidence.

Programs to Potentially Highlight in the Educational Effectiveness Review Report

Based on the quality of the assessment work and results, the following programs were identified as potential candidates to be highlighted in the Educational Effectiveness Review Report: Core 1, Environmental Engineering, History, Mechanical Engineering, Physics, and Spanish Minor. The committee recognized that the final selection would depend upon the larger context of the report.

Methods

The ad-hoc committee included six individuals, consisting of faculty and staff from each School and Core 1/General Education (Appendix C). Each PLO Report was reviewed by a primary and secondary reviewer guided by a template (Appendix D) constructed around the *Rubric for the Report on PLO Assessment* (Appendix A). Following an in-person norming session focused on an example report, reviews were conducted asynchronously with primary reviewers forwarding completed reviews to secondary reviewers. After reading the PLO Report, secondary reviewers then considered the primary reviewer's summary comments and, as necessary, noted any supplemental or discrepant points. Primary and secondary review responsibilities were split evenly among committee members. Workload was distributed in this way to reduce the workload associated with reading and evaluating a large number of reports near the end of the semester.

Through this process, reviewers

1) Rated the program's level of development for each of the rubric's five criteria;
2) Identified two or three assessment practices to be strengthened, based on the rubric criteria identified in step 1 to be most in need of development; and
3) Identified one to two strengths of each program's assessment practices.

For the few cases in which the two reviewers' rubric-based ratings disagreed by one level of development (ex. emerging versus developed), the average rating was calculated (ex. emerging/developed). For half-step differences in rater scores (ex. intermediate/developed versus developed), the shared level of development was calculated (ex. developed). Reviewers never disagreed by more than one level of development.

To identify frequently observed assessment or student learning-related strengths, weaknesses or potential issues, reviewers' narrative comments were coded and the frequency of each code was calculated. Using these results, the committee identified common assessment or student learning-related themes or issues to be addressed during an in-person meeting.

Finally, to gain some sense of how useful the rubric is for communicating assessment-related expectations and to gauge each program's impression of the quality of its assessment work, the committee's rubric scores were compared to the rubric-based self-evaluations each program reported in its PLO Report.

Appendix A: UC Merced's *Rubric for the Report on PLO Assessment*.

RUBRIC FOR REPORT ON PLO ASSESSMENT

Center for Research on Teaching Excellence

	Criterion	Initial	Emerging	Developed	Highly Developed
ASSESSMENT METHODS	**Assessable Program Learning Outcome (PLO)**	PLO does not identify what students can do to demonstrate learning (vague, immeasurable verb statements like "students understand major theories"). No rubric developed.	PLO indicates how students can demonstrate learning. Action verb may be general and the PLO may not be observable or measurable. Assessment criteria[1] have not been identified or are incomplete. Rubric in early stages of development.	PLO describes how students can demonstrate learning, identifying observable and measurable results. Criteria are articulated in the form of a rubric, criteria and standards[1] may need further development to be more meaningful and consistently applied.	PLO specifically describes how students can demonstrate learning. Rubric clearly articulates explicit criteria and standards[1] for assessing the PLO, identifies the most important aspects of student learning, and includes descriptions of student performance at varying levels.
	Valid Evidence	It is not clear that potentially valid evidence is collected for the PLO <u>and/or</u> individual faculty use personalized rather than programmatic criteria and standards[1] to assess student work or performance.	Faculty have reached general agreement on the types of evidence to be collected for the PLO but may not include both direct and indirect forms. Evidence needs to be further focused or aligned with PLO or emerging criteria to produce truly meaningful and useful results.	Faculty collect relevant & sufficient evidence for each outcome, including both indirect and direct evidence. Assessment instruments (ex. rubric) assess the level of student attainment. Evidence is aligned with the PLO and assessment criteria to enable meaningful results and conclusions.	Assessment criteria have been pilot-tested and refined over time, usually shared with students. Direct and indirect evidence are designed to mutually inform conclusions. Feedback has led to refinements in the assessment process.
	Reliable Results	Reviewers of student work are not calibrated to apply assessment criteria in a uniform way; there are no checks for inter-rater reliability	Reviewers are calibrated to apply assessment criteria in a uniform way <u>or</u> faculty routinely check for inter-rater reliability.	Reviewers are calibrated to apply assessment criteria in a uniform way <u>and</u> faculty routinely check for inter-rater reliability.	Reviewers are calibrated, and faculty routinely find assessment data to have high inter-rater reliability.
RESULTS & CONCLUSIONS	**Results Summary**	Results (data table or other means) are not included in report.	Results (data table or other means) are included but unclear or missing key data.	Results clearly delineated for each line of evidence in tabular or other summary formats. May reference benchmarks or other expectations.	Results clearly delineate each line of evidence, indicating various levels of achievement. Includes benchmarks.
	Conclusions & Recommendations	Report identifies implications but no recommendations for improvement in student learning or assessment practices and no explanation of how these claims are derived. No reasoning offered in support of claims.	Report identifies some conclusions, implications, and recommendations for improvements regarding student learning or assessment, but the claims are vague or questionably related to results. Support for claims is occasionally insufficient. Questions of validity or reliability are not discussed. Results may be discussed by limited number of faculty, recommendations may be difficult to implement due to lack of convincing results and/or limited faculty involvement or support.	Report clearly articulates conclusions, implications and recommendations for improvement regarding both student learning and assessment and which could be drawn from results. Includes some consideration of the reliability and validity of results. May offer vague support for some claims. Results have been discussed by many faculty and recommendations likely to be implemented due to faculty involvement and support and quality of assessment work.	Report articulates a well-reasoned critique of conclusions, implications, and recommendations that could be drawn from the results for both student learning and assessment. Includes a well-reasoned discussion of validity and reliability of results. Faculty discuss results, plan needed changes, secure necessary resources, and implement changes. Efforts to collaborate with others, such as librarians or student affairs professionals, to improve results.

[1] Criteria are the specific skills or abilities to be measured. Standards describe the levels of performance for a given criterion (ex. proficient to exemplary).

Based on C. Jenefsky & JFKU Program Review Council (2008) and WASC rubrics (2007)

Laura E. Martin & Anne Zanzucchi, Center for Research on Teaching Excellence, University of California, Merced. 2009.

RUBRIC FOR REPORT ON PLO ASSESSMENT

Center for Research on Teaching Excellence

This rubric has five major criteria:

(1) **Assessable Program Learning Outcomes**: Program learning outcome should be reasonable and appropriate for the degree level. If national disciplinary standards are available, the PLO may reflect those priorities. To be assessable, it should involve specific, active verbs with supporting details describing how students will demonstrate their learning. For instance, avoid verbs of general cognition such as "know" or "understand" and instead use verbs like "demonstrate by" or "solve" that show how learning is applied. Through discussion of examples of student work and perhaps course-specific rubrics used by faculty, faculty groups have agreed on explicit criteria[2] and elaborated a program-level rubric. For more information, see <http://crte.ucmerced.edu/program-learning-outcomes-resources>.

(2) **Valid Evidence**: To be valid, evidence must be discussed among faculty and aligned with both the expectation(s) described by the PLO and the criteria[2] faculty use to evaluate student learning. Valid evidence is also linked to sample size or sampling approach, so as to be representative of a norm. For more information, see the appended selection on sample sizing from Linda Suskie's *Assessing Student Learning: A Common Sense Guide* (2004).

(3) **Reliable Results**: Reliable results stem from agreement among faculty about the standards[2] used to evaluate student work, usually as articulated in a faculty-developed, program-wide rubric. Agreement about how to apply these standards in the evaluation of student work (i.e. calibration) is rooted in discussion and practice. Some questions to consider are: How do faculty promote calibration? How do faculty check for calibration? I.e. when faculty apply a rubric to student work, how consistently do they reach the same conclusions (i.e. exhibit inter-rater reliability)? If results are inconsistent, how can inter-rater reliability be improved?

(4) **Summarizing Results**: When drafting a results chart (data table or other means), it is important to consider multiple audiences. How would faculty within your department understand the results? If viewed by outside stake-holders like students, faculty from other programs, administrators, parents, etc., would they reach similar conclusions? Comparing the results to previous results in your program, expectations your program has set for student learning, or to results of similar programs within or outside of the UC (i.e. benchmarking) can provide context for interpreting the results.

(5) **Conclusions and Recommendations**: An effective conclusion closes the loop by analyzing results and implementing change. The narrative should address some probable conclusions based on the results. For example, if students were not given a clear incentive to participate in a particular assessment, the results may not be completely reliable as students may not have been motivated to perform at their best. Specific actions and a timeline for implementation should also be provided since the goal is to gather data to improve both student learning and the ability to engage in effective assessment. Changes might include improving the assessment process or curriculum, examining curriculum content in support of skill development, changing pedagogical practices, stimulating faculty discussion, simply re-examining program learning outcomes, or identifying ways student support services (tutoring, advising, the library) might contribute to increased student success.

---

[2] Criteria are the specific skills or abilities to be measured. Standards describe the levels of performance for a given criterion (ex. proficient to exemplary) and in doing so enable their measurement.

Based on C. Jenefsky & JFKU Program Review Council (2008) and WASC rubrics (2007)

Appendix B: 2009 PLO Report Submission Record as of Spring 2010 when the review was conducted.

School	Program Name	Program Type	2009 PLO Report Submitted?
SoE	Bioengineering	Major	No
	Computer Science & Engineering	Major	Yes
	Environmental Engineering	Major	Yes
	Materials Science & Engineering	Major	Yes
	Mechanical Engineering	Major	No
	Environmental Systems	Graduate	Yes
SNS	Applied Mathematics	Major	Yes
	Biology	Major	Yes
	Chemistry	Major	Yes
	Earth Systems Science	Major	Yes
	Physics	Major	Yes
	Natural Sciences Education Minor	Minor	Yes
SSHA	Anthropology	Major	Yes
	Cognitive Science	Major	Yes
	Economics	Major	No
	History	Major	Yes
	Literatures & Cultures	Major	Yes
	Management	Major	No
	Political Science	Major	Yes
	Psychology	Major	Yes
	Sociology	Major	Yes
	American Studies	Minor	Yes
	Arts	Minor	Yes
	Philosophy	Minor	Yes
	Spanish	Minor	Yes
	Service Science	Minor	Yes
	Writing Minor	Minor	Yes
College One	Core 1	GE	Yes

Appendix C: Roster of the Ad-hoc Committee for PLO Report Review, Spring 2010.

Name	Title & Relevant Committee Memberships	School/ Unit Affiliation
Gregg Camfield	Professor, Chair of the WASC Steering Committee; member Senate-Administrative Council on Assessment	SSHA
Tom Hothem	Lecturing Faculty; member of the WASC Steering Committee	SSHA & Core 1 (GE)
Valerie Leppert	Associate Professor	SoE
Laura Martin	WASC Coordinator, Assessment Coordinator, member of the WASC Steering Committee, member Senate-Administrative Council on Assessment	Center for Research on Teaching Excellence
Masa Watanabe	Director of Student Success, School of Natural Sciences	SNS
Jack Vevea	Associate Professor, member of the WASC Steering Committee	SSHA

Appendix D: Review Template

PLO Report Review: Instructions and Form

Background:
The goals of this PLO Assessment Report Review are to (1) provide feedback to programs on their assessment efforts, (2) identify and report back to each School's faculty any assessment or student learning issues common to the School's programs, and (3) identify programs whose results might serve as case studies in our EER Report. To support this work, we will also (4) rate each program's assessment efforts against the *Rubric for the Report on PLO Assessment*.

Instructions:
Primary Reviewers:
For each PLO Report you review please complete the primary reviewer sections of the Review Form, then forward the completed forms to the secondary reviewer.

Secondary reviewers:
Please review the PLO Reports and the primary reviewer's responses to the Review Form. In the secondary reviewer sections of the form, please note any differences with the primary reviewer's conclusions, or any additional thoughts, you might have. Finally please submit completed Review Forms to Laura, lmartin@ucmerced.edu.

PLO Report Review Form

1) Name of Program:_____

2) Please assess the program's level of development with respect to each of the five criteria in the *Rubric for the Report on PLO Assessment* (Appendix B). Provide your conclusions, along with any supporting comments, in the table below as I (Initial), E (Emerging), D (Developed) or HD (Highly Developed). A program can be assessed to fall between two levels of development, for example, I/E or E/D.

Reviewer	Assessable PLO	Valid Evidence	Reliable Results	Results Summary	Conclusions & Recommendations
Primary					
Secondary					

3) Please provide the program with constructive feedback regarding its assessment practices. (These comments will be excerpted and shared with the program on behalf of this committee, so please craft these with your colleagues in mind.)

   a) In one sentence, describe a clear strength of the program's assessment efforts.

      Primary Reviewer:

      Secondary Reviewer:

   b) Based on the criterion (or criteria) identified in question 2 as needing the most development, and the corresponding supplemental questions provided in Appendix A, please identify two or three assessment practices to be strengthened.

      Primary Reviewer:

      Secondary Reviewer:

4) Please note what you might imagine to be emerging, shared themes related to the assessment process and/or student learning results.

Primary Reviewer:

Secondary Reviewer:

5) Briefly evaluate the potential of the program's assessment work as a case study for our EER report. Relevant criteria include:
   - Quality of assessment work, including most importantly evidence of assessment-based actions to improve student learning.
   - Illustrative of a commonly observed approach to assessment.
   - Illustrative of trends or common conclusions emerging from PLO Reports.

Primary Reviewer:

Secondary Reviewer:

6) Any outstanding thoughts or questions?

Primary Reviewer:

Secondary Reviewer:

APPENDIX A: A set of questions is provided below to help guide the identification of assessment practices to be strengthened in response to Question 3 above. To support this process, the questions are organized by the criteria that appear on *Rubric for the Report on PLO Assessment*. *Please pay particular attention to italicized questions as they were specifically identified in the most recent WASC Commission Action Letter as important areas of development and so should have priority in feedback.*

Assessable PLOs:

- *As written, is the PLO measurable? Does it involve specific, active verbs that such as "demonstrate by" or "solve" as opposed to verbs of general cognition like "know" or "understand"?*
- Is the PLO likely to be understood by students? Of use to students?
- To help faculty (and students as relevant) develop a shared understanding of what student mastery of the PLO looks like in practice, has a rubric been developed that articulates criteria[3] and standards[4] of performance (for each criterion)?

Valid Evidence:

- Is a rationale for the assessment strategy provided? Does the program explain why a particular piece of work, or a particular course, is an appropriate focus for examining student achievement with respect to the PLO?
- Related to the bullet above, does the assessment work have a program/PLO focus rather than course-level focus?
- *Does the assessment method include at least one form of direct evidence (i.e. actual student work)?*
- Is the assessment measure going to produce results that bear on the PLO? (I.e. Is it aligned with the PLO?)
- Will the sample size and sampling strategy produce results that represent the student norm?
- Are multiple, complementary forms of evidence used to more precisely identify areas in need of attention and to strengthen confidence in the conclusions? (For example, direct and indirect evidence?)

Reliable Results:

- Did the program use a rubric with explicit standards and criteria to review student work and, thereby, promote agreement among reviewers about student proficiency?
- Did at least two faculty members review each piece of student work?
- Were faculty reviewers calibrated or normed with respect to explicit standards and criteria used to asses student work in order to promote agreement among reviewers about observed student proficiencies?
- Did the program determine how consistently faculty reached the same conclusion with respect to a piece of student work (i.e. determine inter-rater reliability)?

Summarizing Results:

[3] "The qualities we look for in student evidence." (Driscoll and Wood, 2007) The specific skills or abilities to be measured.
[4] Standards describe the levels of performance for a given criterion (ex. proficient to exemplary).

- To gain a sense of the distribution of student performance relative to performance standards or levels of proficiency, does the program describe the percentage of students meeting specific levels of performance, for example, as described in a rubric?
- Does the program identify a goal for the percentage of students meeting minimum or higher levels of proficiency? Are the assessment results evaluated in relation to this goal?

Conclusions and Recommendations:

- Are the program's conclusions supported by the results?
- Are issues related to the validity and reliability of the results considered in drawing conclusions and identifying actions to be taken on the basis of those conclusions?
- As warranted, does the program propose some actions to be taken in response to their conclusions? Are the actions well-aligned with the conclusions?
- In order to promote improvements in student learning have the results, conclusions and proposed actions been shared with the faculty and approved by the faculty?