



Section How To: I. Objectives

[According to elements IA and IB of the Assessment Progress Template Evaluation Rubric (below), good objectives are clear, specific, and oriented relative to students. See “Exemplar” for how objectives are incorporated into the complete hypothetical report.]

**I. Student-centered learning objectives**

1 – Beginning	2 – Developing	3 – Good	4 – Exemplary
<b>1. Student-centered learning objectives</b>			
<b>A. Clarity and Specificity</b>			
No objectives stated.	Objectives present, but with imprecise verbs (e.g., know, understand), vague description of content/skill/or attitudinal domain, and non-specificity of whom should be assessed (e.g., “students”)	Objectives generally contain precise verbs, rich description of the content/skill/or attitudinal domain, and specification of whom should be assessed (e.g., “graduating seniors in the Biology B.A. program”)	All objectives stated with clarity and specificity including precise verbs, rich description of the content/skill/or attitudinal domain, and specification of whom should be assessed (e.g., “graduating seniors in the Biology B.A. program”)
<b>B. Orientation</b>			
No objectives stated in student-centered terms.	Some objectives stated in student-centered terms.	Most objectives stated in student-centered terms.	All objectives stated in student-centered terms (i.e., what a student should know, think, or do).

[For illustrative purposes, the Center for Assessment and Research Studies has created a hypothetical degree program (a B.A. in 80s pop culture). The faculty of this program have written four objectives, that drive the other steps of their assessment process. Provided below are poor and better examples of two of these objectives.]

**Description of process for developing objectives:** In early spring of 2003 all program faculty participated in an objective writing process for the BA program in 80s pop culture. We began with the following question: What knowledge, skills, or attitudes should our students possess by graduation? The initial list consisted of 20 objectives. Over the course of the semester the faculty combined some of the objectives and dropped others. Finally, the faculty endorsed four universal objectives for program graduates.

**Commented [K1]:** A brief history of how the objectives were written helps provide context for the reader and aids with institutional memory.

Examples

Example 1	
Poor	The students will understand basic components of 80s pop culture.
Better	Students graduating from the BA program in 80s pop culture will identify (a) relevant musicians, (b) TV shows and movies, (c) fads, and (d) technology of the period.

Example 2	
Poor	Faculty will teach writing skills.
Better	Students graduating from the BA program in 80s pop culture will write a cogent argument about how a political event in the 80s shaped pop culture. These papers should (a) contain a coherent argument, (b) use references appropriately, (c) be well organized, (d) and consist of sentence-level mechanics that enhance the readability of the paper.

**Commented [K2]:** Not specific.

**Commented [K3]:** Verbs like “know” or “understand” are too broad.

**Commented [K4]:** Fails to elaborate on the major components of 80s pop culture.

**Commented [K5]:** Note that this objective provides much more guidance to the rest of the assessment process: specificity about whom is assessed, relevant verb (identify) associated with the desired behavior, and detail regarding what the students should identify.

**Commented [K6]:** Objectives should be stated in terms of what students should know, think, or do. This objective is stated in terms of what faculty will do.

**Commented [K7]:** Again, a good objective like this one gives important clues about what an assessment will entail.

**Two other examples of good objectives**

Students will interpret quantitative results based on an analysis of (a) methodology, (b) graphs, and (c) tables.

Students graduating from the BA program in 80s pop culture will deliver effectively a presentation with an (a) engaging introduction, (b) a logical and fluid body, and (c) a conclusion that reinforces the main ideas of the presentation and closes smoothly.

Section How To: II. Course/Learning Experiences

According to Element II of the Assessment Progress Template Evaluation Rubric (below), all objectives should have classes and/or activities linked to them. See “Total Example” for how this component is incorporated into the complete hypothetical report.

**II. Course/learning experiences that are mapped to objectives**

<b>1 – Beginning</b>	<b>2 – Developing</b>	<b>3 – Good</b>	<b>4 – Exemplary</b>
No activities/ courses listed.	Activities/courses listed but link to objectives is absent.	Most objectives have classes and/or activities linked to them.	All objectives have classes and/or activities linked to them.

For illustrative purposes, the Center for Assessment and Research Studies has created a hypothetical degree program (a B.A. in 80s pop culture). The faculty of this program have written four objectives that drive the other steps of the assessment process. Provided on the following pages are two acceptable ways of constructing a “curriculum map” relative to these objectives.

Example 1 (More detailed)

Coverage objective: 0 =No Coverage; 1 = Slight Coverage; 2 = Moderate Coverage; 3 = Major Coverage

	Obj 1 (Identification of 80s Components)	Obj 2 (Research Methodology)	Obj 3 (Writing Critically)	Obj 4 (Oral Comm)
PCUL201 (Introduction to the 80s)	3	0	1	0
PCUL301 (80s Music)	3	0	1	2
PCUL302 (80s Fads)	3	0	1	2
PCUL303 (80s TV and Movies)	3	0	0	2
PCUL304 (80s Technology)	3	1	1	0
PCUL361(Methods and Analysis)	0	3	1	0
PCUL401 (80s Politics and Culture)	1	1	3	0
PCUL402 (Profiles of 80s Icons)	1	0	1	3
PCUL403(The Music Video)	2	0	0	0
PCUL404(The 80s and Today)	0	2	3	1
PCUL480 (Capstone)	0	2	2	2

**Commented [K8]:** This type of curriculum map not only shows what classes link to which objectives, but also reveals the degree to which the classes are intended to link to the objectives. Here are a few observations one could make from this map: (1) The courses are aligned most closely with objective 1 (2) and objective 4 has the least coverage. (3) Relatively speaking PCUL403 (The Music Video) is less aligned with the program objectives than other courses.

Example 2 (Less detailed)

**Commented [K9]:** This type of curricular map is acceptable but provides less information than the first example.

	Obj X (Identification of 8s Components)	Obj X (Writing Critically)	Obj X (Oral Comm)	Obj 4 (Team Work)
PCUL201 (Introduction to the 80s)	x		x	
PCUL301 (80s Music)	x		x	x
PCUL302 (80s Fads)	x		x	x
PCUL303 (80s TV and Movies)	x			x
PCUL304 (80s Technology)	x	x	x	
PCUL361(Methods and Analysis)		x	x	
PCUL401 (80s Politics and Culture)	x	x	x	
PCUL402 (Profiles of 80s Icons)	x		x	x
PCUL403(The Music Video)	x			
PCUL404(The 80s and Today)		x	x	x
PCUL480 (Capstone)		x	x	x

### Section How To: 3. Methodology

The Methods section of the APT is the most involved. Essentially, programs make an argument that the methodologies used for their assessments are of sufficient quality to yield trustworthy results. This How-To piece gives programs a sense of what type of evidence constitutes a strong argument and, concomitantly, high ratings on the evaluation rubric. As with other sections, if you need assistance please contact PASS. *See “Total Example” for how the methodology section is incorporated into the complete hypothetical report.*

Although optional, CARS recommends that you provide a table at the beginning of this section summarizing the methods. This provides a quick overview for the reader. For example:

*The BA program in 80s Pop Culture uses four instruments for its assessment. This table summarizes the process involving these instruments. More detail about the methodology follows the table.*

	Corresponds to which objective(s)	Type of Measure	Data Collection	Expected Results
Graduation Test	1 (identification) & 2 (methods)	Direct	Census of students on assessment day.	Avg of 80% correct on identification items and 65% on methods items.
Writing Rubric	3 (writing)	Direct	Representative Sampling/Course Embedded	Avg of 3 (competent) for each writing trait.
Oral Communication Rubric	4 (oral communication)	Direct	Representative Sampling/Course embedded	Avg of 3 (competent) for each oral comm. Trait.
Graduation Survey	1, 2, 3, 4	Indirect	Census of students on assessment day.	Avg of 3 (moderate gain) for each item set representing objectives.

The Methods section of the APT is evaluated via five elements of the evaluation rubric. I will present those elements, followed by an annotated write-up for a multiple-choice test and a write-up for a performance assessment (writing), and then additional tips.

**III. Systematic method for evaluating progress on objectives**

1 – Beginning	2 – Developing	3 – Good	4 – Exemplary
<b>A. Relationship between measures and objectives</b>			
Seemingly no relationship between objectives and measures.	At a superficial level, it appears the content assessed by the measures matches the objectives, but no explanation is provided.	General detail about how objectives relate to measures is provided. For example, the faculty wrote items to match the objectives, or the instrument was selected “because its general description appeared to match our objectives.”	Detail is provided regarding objective-to-measure match. Specific items on the test are linked to objectives. The match is affirmed by faculty subject experts (e.g., through a backwards translation).
<b>B. Types of Measures</b>			
No measures indicated	Objectives are not assessed via direct measures (only with indirect measures).	Most objectives assessed with direct measures.	All objectives assessed using at least one direct measure (e.g., tests, essays).
<b>C. Specification of desired results for objectives</b>			
No a priori desired results for objectives	Statement of desired result (e.g., student growth, comparison to previous year’s data, comparison to faculty standards, performance vs. a criterion), but no specificity (e.g., students will grow; students will perform better than last year)	Desired result specified. (e.g., our students will gain ½ standard deviation from junior to senior year; our students will score above a faculty-determined standard). “Gathering baseline data” is acceptable for this rating.	Desired result specified AND justified (e.g., Last year the typical student scored 20 points on measure x. The current cohort underwent more extensive coursework in the area, so we hope that the average student scores 22 points or better.)
<b>D. Data collection &amp; Research design</b>			
No information is provided about data collection process or data not collected.	No information is provided about data collection process or data not collected.	No information is provided about data collection process or data not collected.	No information is provided about data collection process or data not collected.
<b>E. Additional validity evidence</b>			
No additional psychometric properties provided.	No additional psychometric properties provided.	No additional psychometric properties provided.	No additional psychometric properties provided.

**Example: Graduation Test**

*General Info/Relationship to Objectives/Validity Evidence:* The current version (III) of the graduation test is a 100-item multiple choice test developed internally by the program’s faculty. The initial form of the test was piloted in 2003. Based on subsequent content analysis by two program faculty and psychometric analysis by the Center for Assessment and Research Studies, the test was re-vamped two times resulting in the current version (III), which has been administered since 2005. The items of the test were written expressly to correspond with objective 1 (identification of elements of 80’s culture – items 1 through 60) and objective 2 (interpretation of quantitative results – items 61-100). Two other faculty members reviewed the items and agreed that the items matched the objectives as intended. The

**Commented [KHF10]:** Element 3B on rubric focuses on type of measure. Here I indicate that this is a multiple choice test (a direct measure). In addition I provide some brief information about the test and its development.

**Commented [KHF11]:** Element 3A on rubric. I indicate how the instrument “maps on” to objective(s). The more specific, the better. Note that I indicated which items correspond to which objectives.

internal consistency (as estimated by Cronbach's alpha) of the identification subscale has ranged from .76 to .84 and for the methodology subscale has ranged from .74 to .81. According to measurement experts, these reliabilities are appropriate for group-level decisions. Furthermore, the scores on these two subscales have correlated from .22 to .40 with program GPA, which lends evidence that there is some overlap between students' performances in classes and their performances on the test.

*Data Collection:* All students in PCUL480 (our senior capstone) take the graduation test as part of Assessment Day in February. Because students are required to take this class, almost always their final semester, the results are reflective of graduating seniors. Ten percent of students' final grades in PCUL480 is based on their performance on this exam, hence students typically put forth good effort. Proctors report that over 95% of students finish the test within the allotted amount of time (90 minutes). This year (spring 2010), 89 out of 91 students who were currently enrolled in the class completed the test.

*Expected Results for Current Year (spring 2010):* Student performance has improved steadily since 2005. Last year's (2009) scores were 81% correct on the identification subscale and 60% correct on the methodology subscale. Faculty are satisfied with the performance on the identification subscale, so we continued with our desired result of 80% correct. On the other hand, the methodology scores were considered too low. Because we emphasized methodology in additional courses over the last several years, we upped our expectation for scores in the methodology section to 65%.

Below is an example for a performance assessment.

### **Example: Writing Rubric**

*General Info/Relationship to Objectives/Validity Evidence:* This is the 2<sup>nd</sup> year we will be using the 80s Pop Culture writing rubric, which corresponds with objective 3 (writing critically). We adapted this rubric from JMU's official writing rubric:

[www.jmu.edu/assessment/.../JMU\\_Final\\_Writing\\_Rubric\\_f08.pdf](http://www.jmu.edu/assessment/.../JMU_Final_Writing_Rubric_f08.pdf). The initial rubric was chosen as a starting point because it represented writing similarly to how we articulated it. Specifically its Complexity trait corresponded to our ideas of a "cogent argument"; the *Style and Usage & Mechanics* traits were similar to our conceptualization of "sentence-level mechanics that enhance readability"; and the *Organization* trait captured our ideas of "well organized." So, we kept those traits. However we felt the JMU rubric underrepresented "using references appropriately." Therefore, we added that element. So, in all we retained four elements from the JMU rubric and added a fifth (references) to best correspond with the definition of writing implied by our third objective.

We wanted to ensure that faculty were evaluating the writing assessment consistently. To do so, we consulted with [Program Assessment Support Services \(PASS\)](#), part of the Center for Assessment and Research Studies. PASS conducted generalizability (i.e., reliability) analyses of the results over the past two years. The first year yielded a phi coefficient of .54, which was unsatisfactorily low (i.e., low rater agreement). This year the phi coefficient is .68. We think this improvement may be due to better rater training. This year's aggregate writing scores correlated at .25 with the students' SAT Writing scores, providing some additional validity evidence for the our writing assessment.

**Commented [KHF12]:** Element 3E on rubric. I provide reliability information about the scores from the instrument and they are acceptable (i.e., > .60). Also, the scores are correlated with course grades. A positive correlation between the two provides additional validity evidence.

**Commented [KHF13]:** Element 3D on rubric. Note how the exact number of students who took the test is identified and the argument about how those students are representative to, in this case, graduating seniors. Information about student motivation and conditions of testing also presented.

**Commented [KHF14]:** Element 3C on rubric. What results did you want to see from the assessment and why? Providing a rationale is very important. You can justify your desired results in many ways, for example, using previous results as benchmarks, using a standard setting procedure for competency, using discipline specific standards, etc.

**Commented [K15]:** PASS is available to aid you with all components of your assessment process. You may find them particularly helpful regarding data analysis and selecting instruments. [programassessment@jmu.edu](mailto:programassessment@jmu.edu) or 540-568-7962



*Data Collection:* All students in PCUL480 (our senior capstone) are required to complete a 10-page argumentative paper about how a political event in the 80s shaped pop culture. As this assignment is worth 25% of the course grade, students tend to give a good effort. Twenty papers, four from each section, are randomly selected. Two teams of two program faculty raters evaluate the papers, 10 papers per team. These four raters spend approximately an hour and a half on rater training at the outset to assist with inter-rater reliability.

Expected Results for Current Year (spring 2010): Each trait on the rubric is evaluated on a four-point scale (1 = Beginning; 2 = Developing; 3 = Competent; 4 = Advanced) with corresponding behavioral anchors. For example, for the *Usage and Mechanics* trait a 3 connotes writing that “Is generally free of errors in mechanics, usage, grammar, or sentence structure. Reads smoothly. Problems do not compromise meaning.” Given that these students are seniors, we endeavor for the average scores of this group to be at or higher than 3 for each writing trait, connoting competency or better.

### **Additional Notes per improving the methods section.**

#### *About...Relation between objectives and measures*

Because assessment should indicate the degree to which students have made progress on objectives, it’s important to make the case that your measures (instruments) are congruent with the objectives. The more detail you provide, the more convincing. For example, you could indicate which items correspond to which objective for a multiple choice test or how an element of a rubric matches to an objective for a performance assessment (like writing). Such detail indicates that the program gave serious consideration to the measure-to-objective match.

Note that one instrument, like a multiple choice test, could represent more than one objective. In other words, you need not have a separate instrument for each objective. Nevertheless, you would still need to make the argument about how that instrument corresponds to those multiple objectives.

#### *About...Type of measures*

Direct measures of student learning like tests, essays, and portfolios provide the most compelling assessment evidence. Indirect measures, like surveys, can be useful as a supplement to the direct measures. You can achieve a perfect score on this element of APT evaluation rubric by having a direct measure associated with each objective. The ideal situation would be to also include an indirect measure related to each of the objectives. In doing so, you could evaluate an objective from multiple perspectives (i.e., students’ actual performance and their self-reporting).

#### *About...Specification of desired results for objectives*

Results mean little without context. Therefore, providing a priori desired results allows programs greater context by which to talk about relative strengths and weaknesses. Desired results can be anything from percent of students meeting a standard, the average scores of students relative to a standard, degree of growth (requires pre-post testing), performance relative to national user-group norms, performance vs. previous year, etc. However, just providing a target number is not enough. One must provide a rationale for this number. If collecting data from an instrument for the first time, then “collecting baseline data” may be appropriate.

*About...Data collection and research design integrity*

Even with great objectives and measures, results will be compromised with poor data collection or research design. For example, if students do not give a good effort, or your sample is unrepresentative of your population of interest (e.g., all graduating seniors), then the results are less meaningful. The big elements here include information about your sample and how it's representative, the conditions under which the data are collected, and student motivation. Note that this rubric rewards programs for providing more information about their data collection process even if flawed. The rationale is that CARS can provide better help in subsequent assessment cycles to your program the more information you provide.

*About...Additional validity evidence*

To this point, every part of a program's APT contributes to building an argument that the upcoming results are meaningful. In the language of measurement specialists, you are providing evidence for the validity of your results. In addition to this information, reliability of your scores is critical. For example, if test scores had reliability of .30 (which is really poor), then the scores would have considerable error. It would be like a presidential poll where the 95% confidence interval for percentage of votes for a candidate was 50 plus or minus 30. Obviously, this level of imprecision would not allow us to make meaningful inferences about the presidential race. In the testing literature, a reliability coefficient of .80 is excellent within the context of program evaluation. For most commonly used types of reliability (e.g., internal consistency; test-retest), .60 is considered the bare minimum for acceptability. However, there are alternative ways of estimating reliability. If you have questions or concerns about this area, please contact PASS.

Section How To: 4. Results

This section should succinctly provide the results of the assessment as well as an interpretation of the results. This How-To segment provides the rubric elements that are used to evaluate this segment, then provides an annotated example of a results section including a table and interpretation of results. As with other sections, if you need assistance please contact PASS.

**IV. Results of program assessment**

1 – Beginning	2 – Developing	3 – Good	4 – Exemplary
<b>A. Presentation of results</b>			
No results presented	Results are present, but it is unclear how they relate to the objectives or the desired results for the objectives.	Results are present, and they directly relate to the objectives and the desired results for objectives but presentation is sloppy or difficult to follow. Statistical analysis may or may not be present.	Results are present, and they directly relate to objectives and the desired results for objectives, are clearly presented, and were derived by appropriate statistical analyses.
<b>B. History of results</b>			
No results presented	Only current year's results provided.	Past iteration(s) of results (e.g., last year's) provided for some assessments in addition to current year's.	Past iteration(s) of results (e.g., last year's) provided for majority of assessments in addition to current year's.
<b>C. Interpretation of Results</b>			
No interpretation attempted	Interpretation attempted, but the interpretation does not refer back to the objectives or desired results of objectives. Or, the interpretations are clearly not supported by the methodology and/or results.	Interpretations of results seem to be reasonable inferences given the objectives, desired results of objectives, and methodology.	Interpretations of results seem to be reasonable given the objectives, desired results of objectives, and methodology. Plus, multiple faculty interpreted results (not just one person). And, interpretation includes how classes/ activities might have affected results.

For illustrative purposes, the Center for Assessment and Research Studies has created a hypothetical degree program (a B.A. in 80s pop culture). The faculty of this program have written four objectives that drive the other steps of their assessment process. Provided below is an annotated example of how results could be presented. Note that a table is not required for a high rating, but for many programs it may be the most efficient way to describe multi-faceted results.

The results are set up to address two questions. To what degree did this cohort of students achieve the desired results and are this year's results different than the previous year (if applicable)?

**Commented [KHF16]:** If your results are set up to answer specific questions, framing those questions at the beginning will help readers with interpretation.

Page: 12 – For APT Assistance contact PASS at [programassessment@jmu.edu](mailto:programassessment@jmu.edu) or 540-568-7962

Scale or Subscale	Corresponding Objective(s)	2008 Results Mean	2009 Results Mean	*2010 Results Mean (sd)	Desired Result 2010	**2010 Different from 2009?
<b>GRADUATION TEST (n = 91): Subscales Scored in Percent Correct</b>						
Identification	1	76	81	79 (s = 8.2)	80%	no
Methods	2	62	60	68 (11.3)	65%	higher
<b>Writing Rubric (n = 25): 1 = Beginning; 2 = Developing; 3 = Competent; 4 = Advanced</b>						
Complexity	3	n/a	3.2	3.1 (.53)	3	no
Style	3	n/a	3	2.9 (.62)	3	no
Usage & Mechanics	3	n/a	3.4	3.2 (.58)	3	no
Organization	3	n/a	3.1	3.0 (.49)	3	no
References	3	n/a	2.7	2.6 (.75)	3	no
<b>Oral Communication Rubric (n=25) 1 = unsatisfactory, 2 = emerging, 3 = competent, 4 = highly competent</b>						
Delivery Skills	4	n/a	n/a	3.3 (.42)	3	n/a
Introduction	4	n/a	n/a	2.8 (.55)	3	n/a
Body	4	n/a	n/a	3.4 (.38)	3	n/a
Conclusion	4	n/a	n/a	2.8 (.49)	3	n/a
<b>Graduation Survey (n=91): 1 = no gain, 2 = small gain, 3 = moderate gain, 4 = large gain, 5 = tremendous gain</b>						
Identification	1	4.4	4.2	4.4 (.38)	3	no
Methodology	2	3.3	3.4	3.9 (.42)	3	higher
Writing	3	3.1	3.2	3.2 (.57)	3	no
Oral Comm	4	2.7	2.6	2.6 (.8)	3	no

\*The green color coding represents the degree to which the observed results were better than the desired result (the darker green, the better). The red coding is the degree worse than desired.

\*\*Based off of independent t-tests, using  $p < .01$  as significance level (lower alpha due to multiple comparisons).

**Commented [KHF17]:** Element 4A. Make the connection between objectives and results explicit.

**Commented [KHF18]:** Element 4B. Current results are presented with previous results. Current results and the previous year's results for most objectives is Exemplary.

**Commented [KHF19]:** By juxtaposing the current results with the desired results, one can quickly ascertain the degree to which an objective was accomplished.

**Commented [KHF20]:** If possible, provide scale from which scores are obtained. Greatly enhances interpretation.

**Commented [KHF21]:** Color coding results is, of course, not required but can quickly draw the reader's attention to salient findings. By using different shadings, one can convey that achieving desired results is a matter of degree.

**Commented [KHF22]:** Element 4A. When appropriate, using statistics will help determine if differences are due to chance or reproducible.

**Interpretation:** Multiple faculty members discussed and interpreted the assessment results at a program meeting (see Dissemination). On most scales or subscales our students' scores either surpassed or approached the desired results. FYI: We consider light red, light green, or white as approximately hitting the target. Nevertheless, a few results are noteworthy. On the positive side, the subscore for methodology (associated with objective 2) was higher than last year's score and exceeded our expectations for desired results. Furthermore, our students' scores on the methodology section of the graduate survey also surpassed our expectations and were statistically significantly better than 2009's results. Given that this cohort was the first to receive a more deliberate and intensive curriculum on methodology (i.e., extra emphasis in several courses), these results may indicate that this new curriculum is more effective than earlier iterations.

Regarding weaknesses, it appears that students' scores, on average, are rated below competent in using references in their papers (a component of the 3<sup>rd</sup> objective). This interpretation is reinforced given that the relatively low scores have been observed over two cohorts. According to the faculty who rated the papers, errors in citations were fairly common. Specifically, students failed to use the correct punctuation within citations and the citations within text often seemed forced.

Additionally, performance on the oral communication assessment (corresponding to objective 4) appears to be an area of weakness. The introduction and conclusion sections of the capstone presentation were rated slightly below competent, and students reported that they made the least improvement in this area. Furthermore, one has to question the meaning of the ratings associated with the rubric. As stated in the methodology section, the reliability of the oral comm ratings was poor.

**Commented [KHF23]:** Note that it is NOT necessary to interpret every result in detail. Rather, summarize and provide analysis of the salient findings.

**Commented [KHF24]:** Indicate the positive findings from the assessment, even if they are just relative positives.

**Commented [KHF25]:** Element 4C. Note that the objectives are consistently referenced in the interpretation.

**Commented [KHF26]:** Element 4C. Note how the curriculum is integrated into the interpretation.

**Commented [KHF27]:** Procuring feedback from several faculty on the results can help support (or raise skepticism) about the findings from the assessment. Either way, it is important to the interpretation.

**Commented [KHF28]:** Often the interpretation of results is hindered by problems in the methodology. You do not need to identify every flaw, but comment on the most critical areas of the assessment process that should be improved.

Section How To: 5. Dissemination

Element V of the Assessment Progress Template Evaluation Rubric (below) is perhaps the simplest to write up. Nevertheless, it represents a critical component to the assessment process: how results are shared with stakeholders. Below is the rubric element associated with this section followed by an annotated example.

**V. Documents how results are shared with faculty/stakeholders**

1 – Beginning	2 – Developing	3 – Good	4 – Exemplary
<b>5. Documents how results are shared with faculty/stakeholders</b>			
No evidence of communication	Information provided to limited number of faculty or communication process unclear.	Information provided to all faculty, mode (e.g. program meetings, e-mails) and details of communication clear.	Information provided to all faculty, mode and details of communication clear. In addition, information shared with others such as advisory committees, other stakeholders, or to conference attendees.

The results included in this APT are shared with all program faculty at the end-of-the-year program meeting. At which time, the program assessment coordinator highlights the strengths and weaknesses and elicits the rest of the faculty for additional details that would facilitate interpretation. Next, based on our interpretation of the results, we identify actions to take in the upcoming year to improve the program and, if necessary, to improve components of the assessment process.

When the final APT is compiled, we submit via e-mail copies to all of the faculty and our program advisory board, which includes a student representative. Finally, we typically make two or three presentations on assessment at conferences. In 2009-2010, we made two such presentations at the Virginia Assessment Group and at the National Association for Pop Culture Conference.

**Commented [KHF29]:** If a program provided this level of detail, it would receive a 3 "Good." All faculty are included and the dissemination process is clear.

**Commented [KHF30]:** Sharing assessment information with other relevant stakeholders bumps this section up to a 4 "Exemplary."

Section How To: 6. Use of Results

The most important purpose for assessment is to inform a program’s strategy to improve. This section allows programs to explain how they have used results to inform improvement, the most important purpose for assessment. In addition, it allows programs to indicate how they have or will improve their assessment processes. Below are the rubric elements associated with this section followed by an annotated hypothetical example.

**VI. Documents the use of results for improvement**

1 – Beginning	2 – Developing	3 – Good	4 – Exemplary	Cusp of National Model for Learning Improvement	National Model for Learning Improvement
<b>6. Documents the use of results for improvement</b>					
<b>A. Program modification and improvement regarding student learning and development</b>					
No mention of any modifications.	Examples of modifications documented but the link between them and the assessment findings is not clear.	Examples of modifications. (or plans to modify) documented and directly related to findings of assessment. However, the modifications lack specificity.	Examples of modifications (or plans to modify) documented and directly related to findings of assessment. These modifications are very specific (e.g., approximate dates of implementation and where in curriculum they will occur.)	Evidence, from direct measures, suggesting learning improvement due to program modifications. This program responded to previous assessment results, made curricular and/or pedagogical modifications, RE-assessed, and found that student learning improved. Lack of clarity regarding the interventions or methodological issues (unrepresentative sampling, concerns regarding student motivation, etc.) leave legitimate questions regarding the improvement interpretation.	Strong evidence, from direct measures, supporting substantive learning improvement due to program modifications. This program responded to previous assessment results, made curricular and/or pedagogical modifications, RE-assessed, and found that student learning improved. The rationale and explanation of the modifications leading to the change are clearly laid out. The methodology is of sufficient strength that



					most reasonable alternative hypotheses can be ruled out (e.g., sampling concerns, validity issues with instrument or student motivation). In essence, the improvement interpretation can withstand reasonable critique from faculty, curriculum experts, assessment experts, and external stakeholders.
<b>B. Improvement of assessment process</b>					
No mention of how this iteration of assessment is improved from past administrations.	Some critical evaluation of past and current assessment, including acknowledgement of flaws, but no evidence of improving upon past assessment or making plans to improve assessment in future iterations.	Critical evaluation of past and current assessment, including acknowledgement of flaws; Plus evidence of some moderate revision, or general plans for improvement of assessment process.	Critical evaluation of past and current assessment, including acknowledgement of flaws; both present improvements and intended improvements are provided; for both, specific details are given. Either present improvements or intended improvements must encompass a major revision.	N/A	N/A

Regarding using the results for improvement, we would again like to reiterate the improvement on objective 2 (methodology) we observed on both the graduation test and the graduate survey. We

believe this improvement is due to extra emphasis in methodology we implemented over the last two years across several courses. These changes were initiated as a result of earlier assessments.

For this year, the results point to two areas of concern: using references appropriately (part of Objective 3) and concerns about Objective 4 (oral communication). I'll begin with references. This area has been identified as a weakness since we've used the current rubric (2-yrs) and the program faculty have also confirmed that this finding resonates with what they have observed in class. At the program's end-of-the-year meeting, the faculty agreed on a plan to address this problem, which will be implemented in the fall of 2010. Specifically, the instructors of the two classes where writing is heavily emphasized - PCUL401 (80s Politics and Culture) and PCUL404 (The 80s and Today) – will

1. Share the results of the past writing assessment with students, emphasizing that references is a concern.
2. Provide poor and good examples of incorporating references into papers. Note: Dr. C. Lauper has agreed to pull together these examples for the other faculty.
3. Evaluate references explicitly (using that component of the writing rubric) on papers in their classes.

Given that some students take these classes as juniors and others as seniors, the full effect of this intervention will not likely show up in students' scores until spring 2011.

The conclusion that our students are not performing well in oral communication is not as strong. Indeed the graduate survey responses indicate that students feel they have made less gains in this area than other areas, but our direct measure (evaluation of presentations) is not at a stage yet where the results are reliable. Therefore, we will concentrate next year on improving this assessment. The problem is that raters do not agree closely with one another. Perhaps this should not be a surprise given that raters received no training prior to rating. Dr. D. L. Roth has agreed to work with the Center for Assessment and Research Studies to create a training session for the ratings for the spring 2011. This training should improve the reliability of the raters, thus giving us more confidence in the results.

In sum, students in our program appear to be learning the objectives. We are pleased with the improvement with objective 2 (Methodology), but note some minor areas to improve. Students exhibit difficulty using references appropriately in paper and our program's assessment of oral communication is not where it needs to be to provide meaningful results. Therefore, our program is providing resources to address both of these concerns.

**Commented [KHF31]:** Mature programs may be able to document the effectiveness of the changes they incorporated from past assessment. This is truly "closing the loop."

**Commented [KHF32]:** Element VIA. Note references back to objectives.

**Commented [KHF33]:** Element 6A. Note the level of specificity here. It's obvious that this program has carefully evaluated the weakness, and assembled a logical set of interventions to address it. Generally, it is better to concentrate on a smaller number of weaknesses and engage them with depth as opposed to trying to tackle every issue (unless the program supports making a complete curricular overhaul).

**Commented [KHF34]:** Element 6B. Acknowledges problem with assessment and provides a clear strategy to address it.