ASSESSMENT GUIDELINES for GENERAL EDUCATION COURSES
REVISED September 2017

For help with GE assessment: Erin Presley, Associate Professor of English
Bethany Miller, Assistant Director of Assessment

Submit assessment data via the link on the GE Website (Faculty → Assessment)

Appendix A: Schedule for assessing GE courses Fall 2014 - Spring 2018
Appendix B: Selecting samples of sections and student work for assessment
Appendix C: Example Data Report and Use of Data
Appendix D: Common questions about GE assessment

OVERVIEW OF ASSESSMENT PROCESS

- Faculty should collaborate to develop assessment items for each course.
- All sections of the same course must use the same assessment items and process.
- A scoring rubric for each GE Element guides the development of the assessment items, and Jennifer Wies is available to help as needed.
- Assessment items can be exam questions (objective or open-ended), written assignments, portfolios, papers, or oral presentations.
- Good assessment is not an "add on." Rather, it is an evaluation of student performance that provides faculty with information that **they would want regardless of whether assessment data were required.**
- It is best to evaluate students toward the end of the semester, or end of lessons focusing on a specific competency. We want to know what students have learned by the end of the course about specific competencies.
- Each course must be assessed twice in each four-year cycle. **(See Appendix A for assessment schedule.)**
- **Assessment instruments should be built into the course and used every semester, regardless of whether data are due to be reported to the University that year.**
- For courses with multiple sections: (1) assessment items should be given to students in all sections of a course, (2) all faculty should use the assessment items for grading purposes, but samples may be selected to be scored for assessment purposes **(See Appendix B).**
- Data should be aggregated across all sections chosen for assessment. UPAC should not see any reports that include faculty names or CRNs.
- Submit data report via online GE Data Link on GE website (**See Appendix C for example data report**).

EXPLANATION: THE "LEVELS" OF ACHIEVEMENT ON THE SCORING RUBRICS (Scoring Rubrics are on GE Website.)

When we evaluate students in our courses we can often categorize them into four levels: The students who "get it," (meet expectations) the students who don't "get it," (do not meet course expectations) and the students who fall in between. Occasionally, we have a student who exceeds course expectations.

These levels correspond to the levels on the Grading Rubrics developed for the assessment of General Education Courses. The levels on the Grading Rubrics ALWAYS refer to achieving the student learning outcomes for THAT SPECIFIC COURSE.

Students who meet course expectations are **Competent.** Students who do not meet course expectations are **Beginning.** Students who fall in between are **Developing,** and the occasional student who "exceeds expectations" is **Accomplished.** Please do not confuse "A" work with "Accomplished." "Accomplished" means "exceeds courses expectations."

HOW TO CATEGORIZE STUDENTS WHEN WRITTEN/ORAL ASSIGNMENTS ARE USED:

In order to have useful and valid assessment instruments it is strongly recommended that at least some competencies are evaluated with written/oral assignments.

Expectations are defined, in very general terms, on each scoring rubric. Faculty will need to make judgment calls about what expectations for student learning are appropriate for each level of accomplishment, given the type of course. **It is expected that faculty within each department will collaborate on this process to develop common expectations and common standards across all sections of the same course.**

ACCOMPLISHED: Students may be categorized as "**Accomplished**" on written/oral assignments if they exceed the expectations of the course for that competency. "Accomplished" level should be used ONLY in cases where the student clearly goes beyond the expectation. It is NOT just "A" work.

COMPETENT: Students may be categorized as "**Competent"** on written/oral assignments if they meet the expectations of the course for that competency.

DEVELOPING. Student's performance on assessment items does not fully meet the expectations for that competency, but is not totally "off base" either. (Performance is incomplete in meeting expectations for that competency.)

BEGINNING: Student's performance does not meet expectations for that competency.

HOW TO CATEGORIZE STUDENTS WHEN "OBJECTIVE" ASSESSMENT ITEMS ARE USED:

In order to have useful and valid assessment instruments it is strongly recommended that at least some competencies are evaluated with written/oral assignments. When "objective" items are included, please use the guidelines below for those items.

Objective items have "right and wrong" answers, e.g., multiple choice, true/false, matching, fill-in-the-blank, short answer.

The scoring rubrics were designed for written/oral assignments that require faculty judgment; thus, the language within each "cell" is not useful for categorizing students with "objective" assessment items. The following guidelines should be used with "objective" assessment items:

A. For each competency on the rubric (Comprehension, Methods, etc.) faculty should write **at least 4 objective items** that students who are competent in that specific course should be able to answer correctly. **All sections of the same course must use the same assessment items**. Thus, faculty must collaborate when writing these items. **If only objectives items are used to evaluate a competency, it is strongly recommended that more than 4 objective items be used.**

B. Each competency on the rubric must be evaluated separately. Thus, a unique set of questions must be written for each competency. For each competency:
- Students who correctly answer **75--100%** of the items for that competency are **"Competent"** in that competency.
- Students who correctly answer **50--74%** of the items for that competency are **"Developing"** in that competency.
- Students who correctly answer **0—49%** of the items for that competency are **"Beginning"** in that competency.

Note. It is not acceptable to classify students as "Accomplished" when using "objective" items.

Appendix A
Assessment of General Education Courses: Fall 2014 – Spring 2018
Eastern Kentucky University

DATE	TASKS
Fall 2014—Spring 2015	Assess ALL GE COURSEs (f2014 or Sp2015).
Fall 2015—Spring 2016	Score student work, prepare data report, and meet with faculty to discuss use of data. Modify instruments, processes, and/or curricula as needed. Submit data report and use of data via online GE Data Link by May 1, 2016**.**
Summer 2016	Automatic reapproval for courses with accurate syllabi, two data reports, and meaningful use of data. (Courses are approved for four-year cycles, and reapproval depends on compliance with syllabus and assessment guidelines.)
Fall 2016—Spring 2017	Assess ALL Courses (f2016 or Sp2017).
Fall 2017—Spring 2018	Score student work, prepare data report, and meet with faculty to discuss use of data. Modify instruments, processes, and/or curricula as needed. Submit data report and use of data **BY APRIL 2, 2018.**

DATA SUBMISSION: GE website (gened.eku.edu). See Faculty → Assessment Information

Appendix B

Suggested Sampling Procedures for Multi-Section Courses

If a course has multiple sections, it is sometimes difficult to get data/assessment assignments from all sections. In an attempt to reduce the burden of getting all data, the following procedure may be used: *(Optional procedure: All sections may be used.)*

- 1—10 sections: Assessment data from all sections
- 11-20 sections: Assessment data from 10 sections*
- More than 20 sections: Assessment data from 1/2 of the sections*

*Please select sections to be included in data report by randomly selecting from the following groups: (a) Richmond campus; (b) Extended campuses; (c) On-line courses; (D) Part-time faculty; (E) Full-time faculty. *We understand that these groups are not mutually exclusive, but are provided as examples of populations that should be included in sample.*

Please note. The procedure of having all instructors in all sections collect the assessment assignment/data should be enforced, **i.e., all students should complete the assignment.** The only change is that the person responsible for gathering all the assignments/data now has the option of "nagging" only a subset of instructors to get those assignments/data at the end of the semester.

This new procedure is optional and is intended to help those departments who have problems getting data/assignments from every section.

Chairs. Chairs may want use the data to compare the performance of students in the various groups, e.g., on-line vs campus courses. However, data should be aggregated before submitting.

How many students' assignment do we need to grade? If the assignment is easy to grade (e.g. scantrons) then grading all of the students' work is desirable. If the assignment is more time-consuming to grade, then choose a random sample: (1) For 1—10 sections choose about 10% of the students' work, with a minimum of 30 students; (2) For more than 10 sections choose about 5% of the students' work. Fewer students may be chosen for the sample when the assignment is unusually difficult to evaluate, e.g., oral presentation videos, portfolios. Please contact Jennifer Wies for guidance.

You do NOT need to have the same sample size for each competency. For example, if your assessment instrument includes multiple choice items to measure "Comprehension," and an essay to measure "Integration," then you might report all students' scores for "Comprehension," and a sample of 30 students for integration.

Appendix C

EXAMPLE OF DATA SUMMARY & USE OF DATA

Course: PST 101
Data from: Spring 2015
Report prepared by: Mary Chang

Number of sections assessed: 4

Assessment Task: *Comprehension* and *Methods* were assessed via 60 multiple-choice items on the final exam. *Clarity of Expression, Formal Analysis, Theoretical Application,* and *Integration* were assessed via a paper.

Sample: Four sections of the course were offered in spring 2011 (total 100 students). All students' exams were assessed for *Comprehension* and *Methods.* A random sample of 10 student papers from each section were selected and graded by a committee for *Clarity of Expression, Formal Analysis, Theoretical Application* and *Integration.*

Data Summary (NUMBER of students in each category)

PST 101: Spring 2011

Competency	**4-Accomplished**	**3-Competent**	**2-Developing**	**1-Beginning**
Comprehension	0	40	30	30
Clarity of Expression	0	20	20	0
Formal/Structural Analysis	5	10	20	15
Theoretical Application	5	20	15	10
Methods	N/A	90	10	0
Integration	0	35	5	0

EXAMPLE: USE OF DATA

Comprehension and Methods. The distribution of scores for *Comprehension* appears reasonable. However, there were 5 comprehension items that few students answered correctly. The faculty judged all 5 items to be important, and have discussed how those concepts could be emphasized and illustrated more in future classes. We have agreed to include the *PST Paper #3c* as an assigned reading in each class, as it emphasizes the concepts that we think are important. Furthermore, we held a professional development session to share teaching strategies we each use to introduce, reinforce, and assessment the essential concepts.

The distribution for *Methods* does not appear reasonable because 90% of the students were assessed as "competent." This does not reflect the students' performance on other assignments in the course. The faculty discussed the assessment items for methods and determined that they are challenging, but that faculty were more-or-less teaching to the test by using the same examples in class as are on the exam. We agreed to NOT use the same variables and concepts in any class examples that appear on the exam. We agreed on a set of examples that could be used in class and will check this item on the next assessment to see how students perform.

Clarity of Expression, Formal Analysis, Theoretical Application & Integration. 50% of students were "developing" on *Clarity of Expression.* Because most students in this course have not completed ENG 101 and 102, this seems like a realistic distribution. Faculty have discussed ways to help students write with more clarity, and next semester will begin some practice writing, with feedback, in several class sessions. We have agreed to use the *TCAC Book* that was written by a team of EKU faculty involved in the TCAC program. The book has several lesson plans and suggestions for helping students become better writers. We also agreed to use the Noel Studio to present short lessons to students, and to encourage students to visit the Noel Studio often. Faculty agreed to focus on feedback of writing, rather than the amount of writing. After providing feedback students will be given the chance to revise their work, and to reflect, in writing, what they learned from the feedback and revision that they will use in future writing.

The distribution of scores for *Integration* does not appear reasonable. Faculty perceptions are that students do not integrate material well, but assessment scores suggest that a majority of students are competent. Faculty re-visited the integration part of the paper and realized that the assignment did not require students to integrate material on their own. Rather, in class, faculty tended to integrate the material for the students, and in the papers, students reiterated what the instructors had previously integrated. Faculty have agreed to modify their teaching strategies to demonstrate integration of material in class, but not use the same material that students are later required to integrate in the papers. We discussed several integration examples that we could present in class that do not use the variables students are expected to integrate on the assessment exam.

Appendix D

COMMON QUESTIONS: ASSESSMENT of GENERAL EDUCATION COURSES

1. WHAT AM I ASSESSING?

You are assessing the General Education Competencies that the course addresses. These Competencies are listed on the **scoring rubric** for the course (rubrics are created for each GE Element). The Competencies align with the broad GE Goals that are the foundation of the class. Your specific student learning outcomes (as stated on the course syllabus) should align with the rubric competencies.

2. WHY NOT JUST USE COURSE GRADES?

Course grades are based on a combination of various course assignments. A student who earns a "B" in a course may have earned an "A" on three exams, and a "D" on a research paper. The global grade of "B" provides no information regarding in which parts of the course the student performed better or worse. On the other hand, assessment results for this student might indicate that he/she performed well only on lower-level thinking tasks, such as comprehension of information, which is often evaluated with exams. Assessment results might show that this same student performed poorly on higher-level thinking tasks, such as integration of material, which might be evaluated via a written assignment, or essay questions on an exam.

Thus, assessment allows us to tease apart the various learning outcomes for the course and determine student learning separately for each outcome. When learning is strong for an outcome, faculty can feel confident that the current curriculum (e.g., lesson plans, assignments) could be maintained. On the other hand, when learning is weak, faculty can tweak just that part of the curriculum in order to help students better learn specific outcomes.

3. HOW DO WE KNOW ASSESSMENT ITEMS ARE VALID?

The issue of validity (do items measure the construct accurately?) is no different for an assessment instrument than it is for other types of student-learning evaluation (e.g., exams, papers). One type of validity that is important for academic tests is *content validity*, which refers to whether assessment items (or test items) reflect a specific domain of content. Thus, assessment items for a math course might include decisions regarding which formulas to use, performance of calculations, and interpretations of results. By creating assessment items that fit each competency on the rubric, assessment items are likely to have better content validity than many tests. Ultimately, the validity of the assessment instrument reflects that knowledge base of the people who created it. This is why assessment instruments should be created by faculty experts. When a group of experts collaborate on creating assessment items, the validity is likely to be better than when only one expert creates items.

4. HOW DO WE DEVELOP AN ASSESSMENT INSTRUMENT?

First, departments should strive to involve as many faculty as possible in the creation of an assessment instrument. Faculty should discuss what they perceive as the important aspects of a course, note the commonalities, and agree on a "core set" of course material. Faculty should then collaborate on developing assessment items that (a) measure this core set of material, and (b) fit the competencies on the appropriate rubric developed for GE.

Second, faculty should discuss the procedure for administering the assessment items. The assessment items should be embedded in an exam or assignment that is given later in the semester in order to test what students have learned in the course. All faculty should give the assessment at a similar time and in a similar context. Students should be told the same thing about the assessment – specifically, NOTHING that you would not normally say about an exam.

5. HOW DO WE GRADE THE ASSESSMENT ITEMS?

Individual faculty should grade their students' assessment items, use the score as part of the students' grade, and use the overall results to determine potential strengths and weaknesses of his/her individual course. This individual grading may or may not be standardized. For example, if the assessment involves a research paper, one faculty member might decide to weigh the literature review 30%, while another faculty member may weigh the literature review only 10%. However, these scores are for the individual faculty members' use; they are not the scores that are reported for assessment purposes.

For purposes of reporting assessment results, data from all sections should be aggregated. Thus, instructors might submit to a committee all the assessment exams/papers. That committee might select a random sample (See Appendix A) and grade (or re-grade) the items according the GE Element-specific Scoring Rubric. All faculty members should grade several assignments to check for reliability in grading procedures. If grading appears to be reliable for several assignments, then one faculty person could grade each assignment, while collaborating with other graders when necessary.

Scoring student work in a group provides opportunities for faculty to discuss common errors, and strengths and weakness of students, and provides opportunities to discuss potential changes to the assessment instrument, process, instructions to students, and/or course lesson plans. **This is the preferred procedure and allows faculty to discuss appropriate ways to USE the data from assessments. Ultimately, using the data in an attempt to improve student learning is the goal of GE assessment.**

.