WVU GEF Pilot Report and Action Plan

12-20-16

Pilot Goals and Expected Outcomes

The goal of this pilot is to determine a) the viability of adoption of the LEAP Outcomes for both GEF and university-level assessment b) the viability of adoption of the Value rubrics as scoring mechanisms c) the cycle, structure, and distribution of a system-wide GEF assessment plan

A. Pilot faculty were clearly able to identify a relevant LEAP Outcome for their courses. There is no concern that the LEAP Outcomes themselves are not viable for GEF assessment at WVU. However, the pilot group ended up consisting of nine participating faculty and eleven sections. Including the sub-outcomes for LEAP Outcome 2 and 3, there were fourteen possible choices for faculty to assess. The pilot group universally chose outcomes from LEAP Outcome 2 and, within Outcome 2, five of the eleven sections picked Critical Thinking and two picked Written Communication and Inquiry and Analysis each. This finding is not unexpected: meta-analyses of use of the LEAP Outcomes has long suggested the most frequently used outcomes were Critical Thinking and Written Communication.
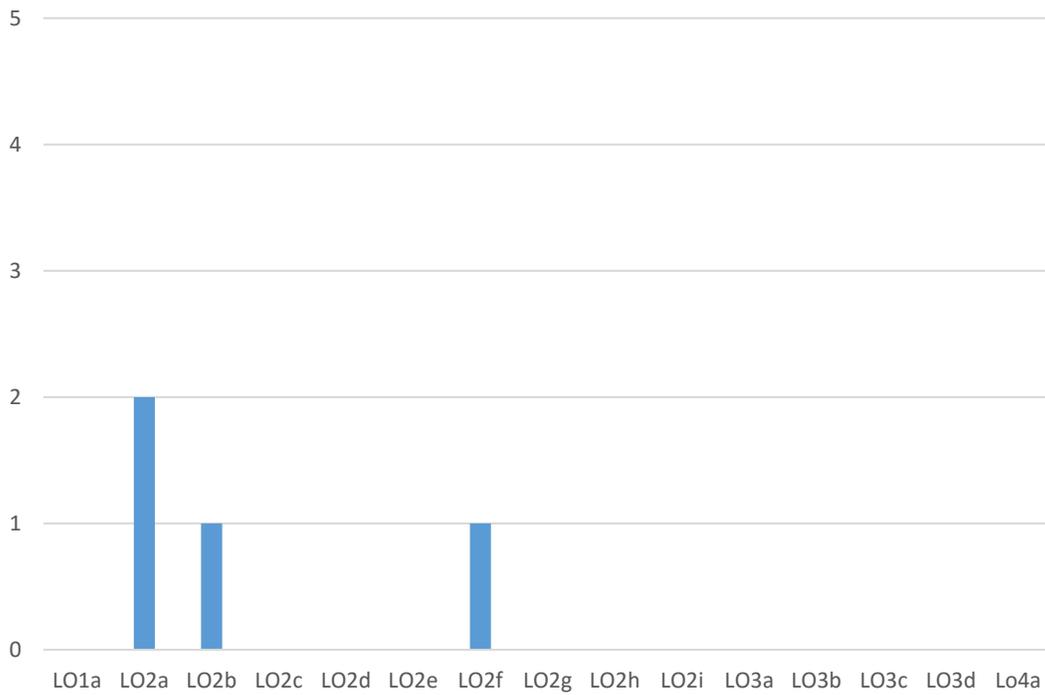
    Action plan: As the GEF areas proceed through the GEC to GEF process, the GEFCo is going to need to attend to the distribution of Learning Outcomes by area. There may be the need for realignment of the areas or a reassignment of learning outcomes.

Distribution by Learning Outcome
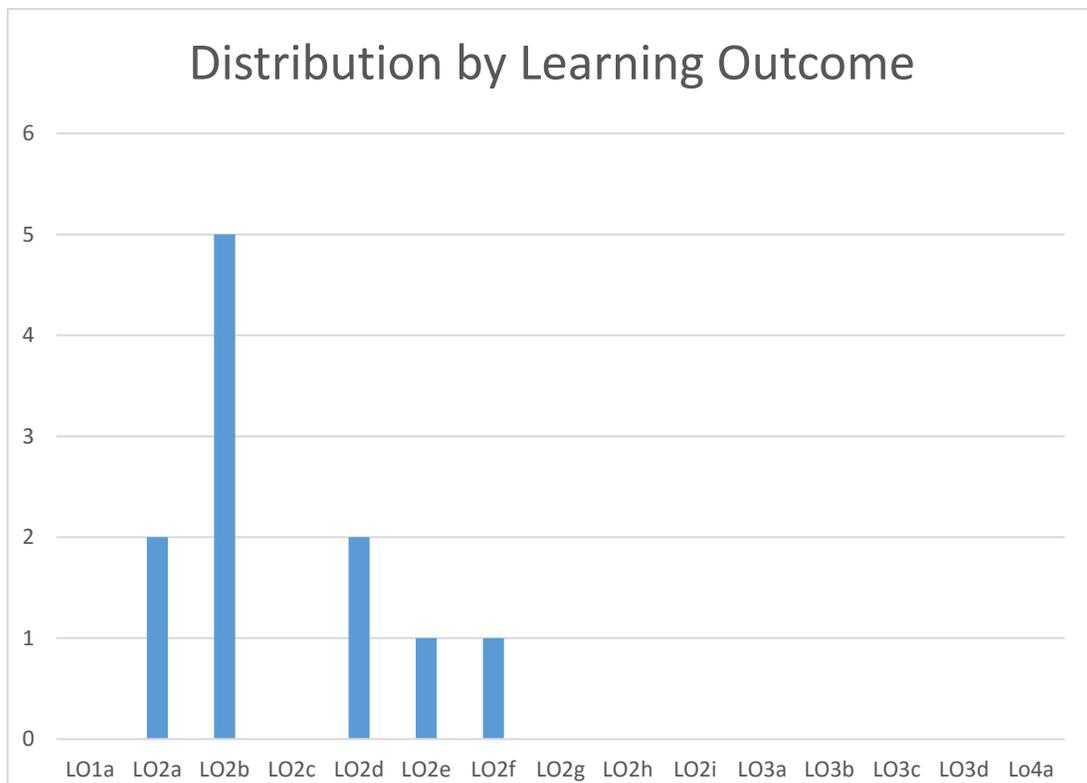
Distribution by Learning Outcome - Honors
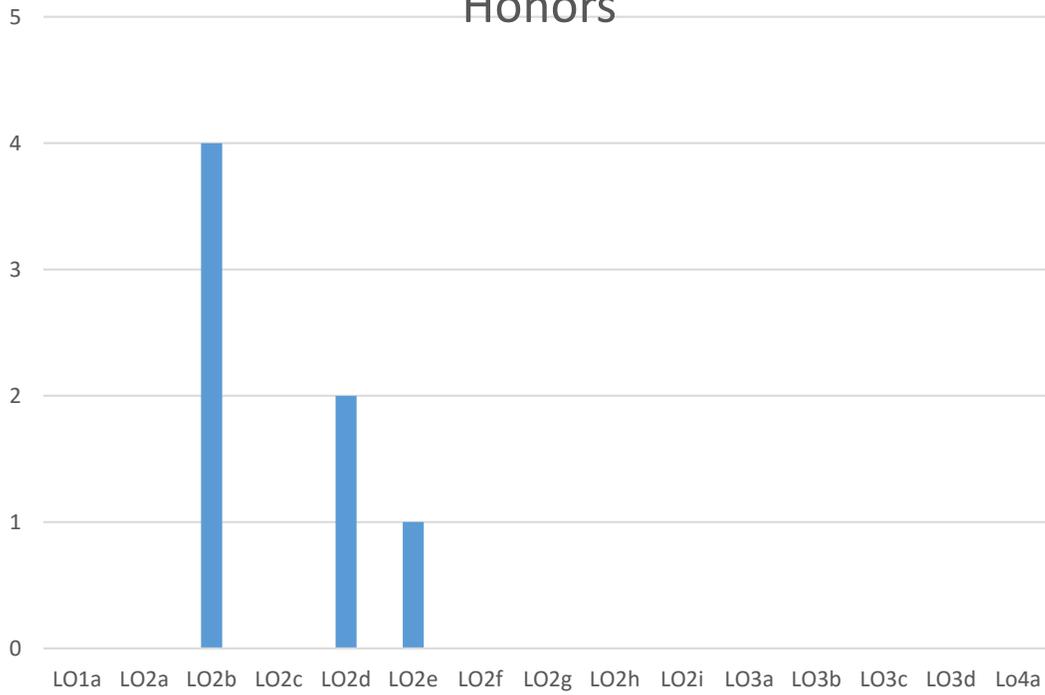
Distribution by Learning Outcome - PSC

B. Faculty were likewise able to identify a corresponding Value rubric for the LEAP Outcome. There is no concern that the Value rubrics themselves are not viable for GEF assessment at WVU.

The second matter to be evaluated under the use of the Value rubrics is whether or not the brief training session held for faculty members in the pilot produced sufficient norming in the scoring. To evaluate the norming, courses were aggregated by course level and the distribution of scores is shown below.

Faculty Scoring Distribution

Faculty Scoring by Percent

Faculty Scoring Distribution by Pilot Group
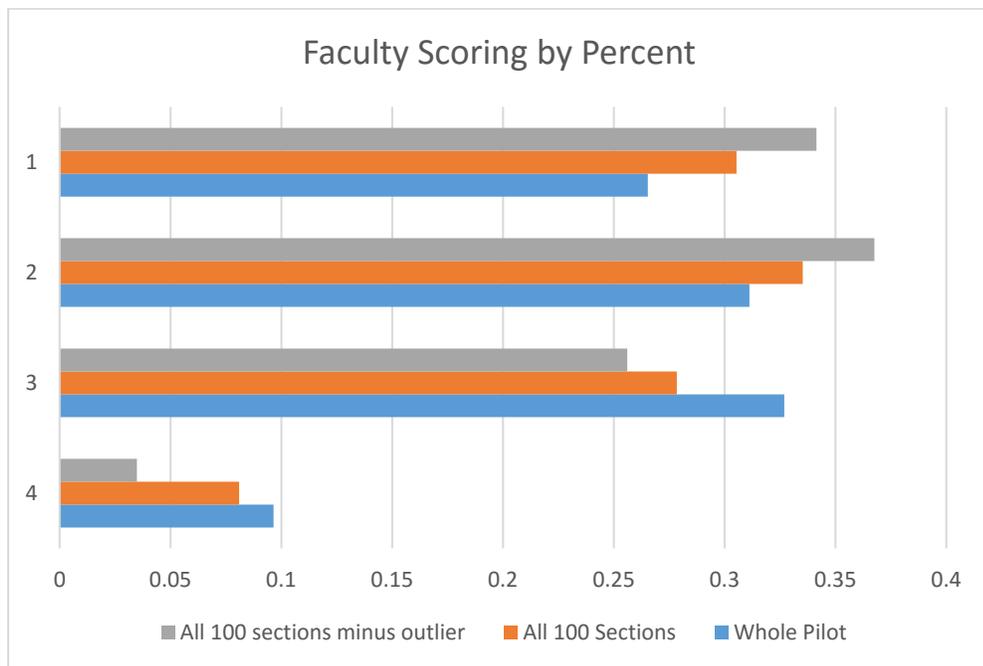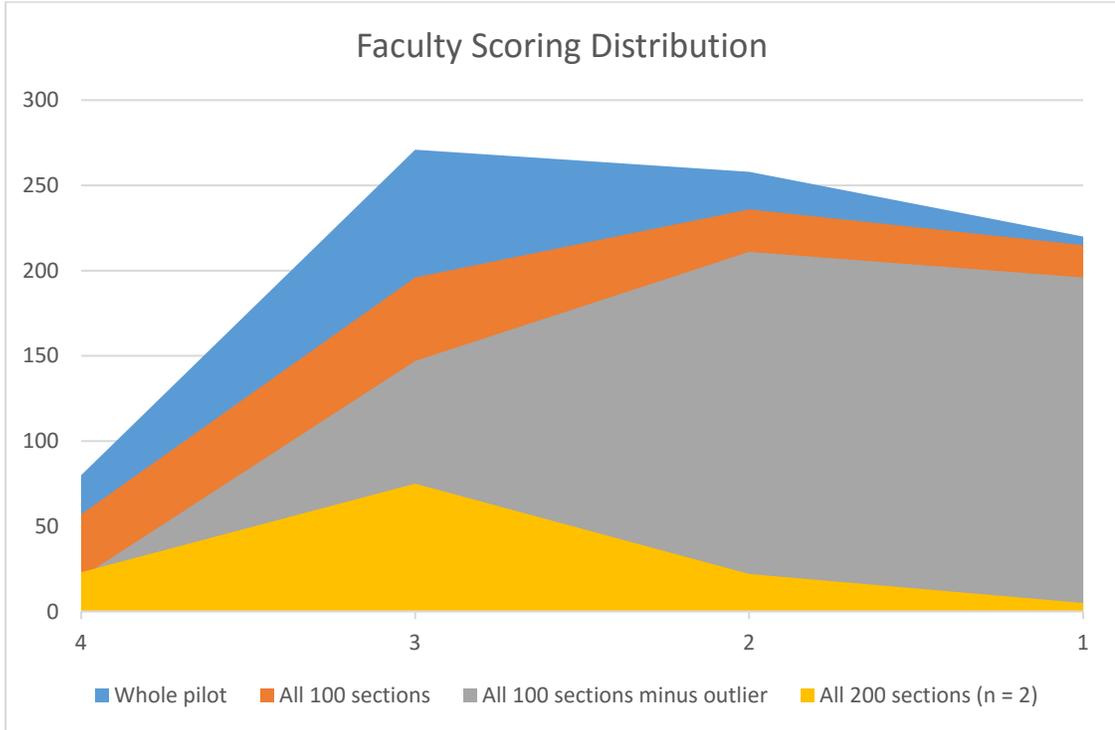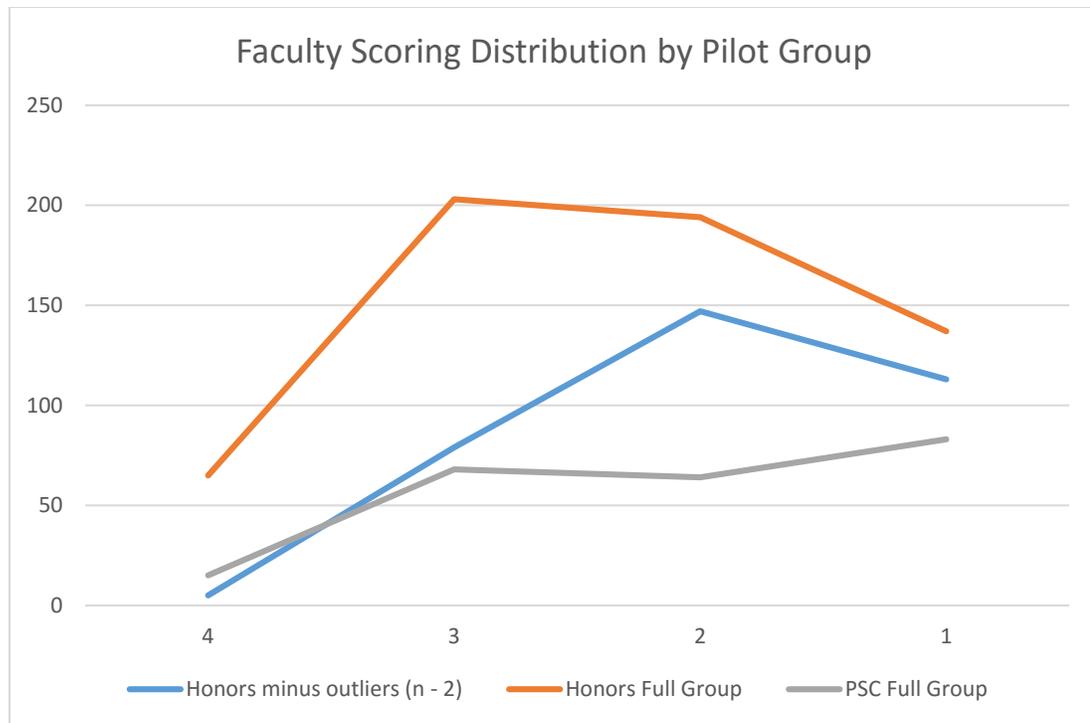
Taken as a whole, the pilot group performed quite well, assigning approximately four times as many 1s as 4s; baseline expectation within a general education program is roughly 40% at the 1, 40% at 2, 15% at 3, 5% at 4. The percentage chart confirms that this expected baseline was more or less met with there being only a slightly higher percentage of 3s expected, pulling approximately 10 to 17.5% more than expected from all three other scores. Overall, for such a brief norming session, the data is well within tolerable limits, especially when eliminating the outliers from the data.

The 100 level sections performed well within tolerable limits; it would seem from that population sample that the pilot's level of scoring norming was a success. When looked at further with the data from one section which was an outlier in the number of 3s and 4s awarded, the data at the 100 level looks well-distributed; again, this underscores the success of the scoring norming delivered to the faculty at the training sessions.

The 200 level data is difficult to interpret, primarily because it came from an n of 2. Clearly, the distribution of scoring is the reverse of expected.

Action plans: 1) Formalize scoring norming 2) Develop a method for delivering that training to larger groups of faculty and at regular, repeated intervals 3) Watch the longitudinal performance of those faculty not attending regular norming sessions 4) Develop a method of identifying and following up with individual faculty members whose scoring suggests they are out of normal scoring practices.
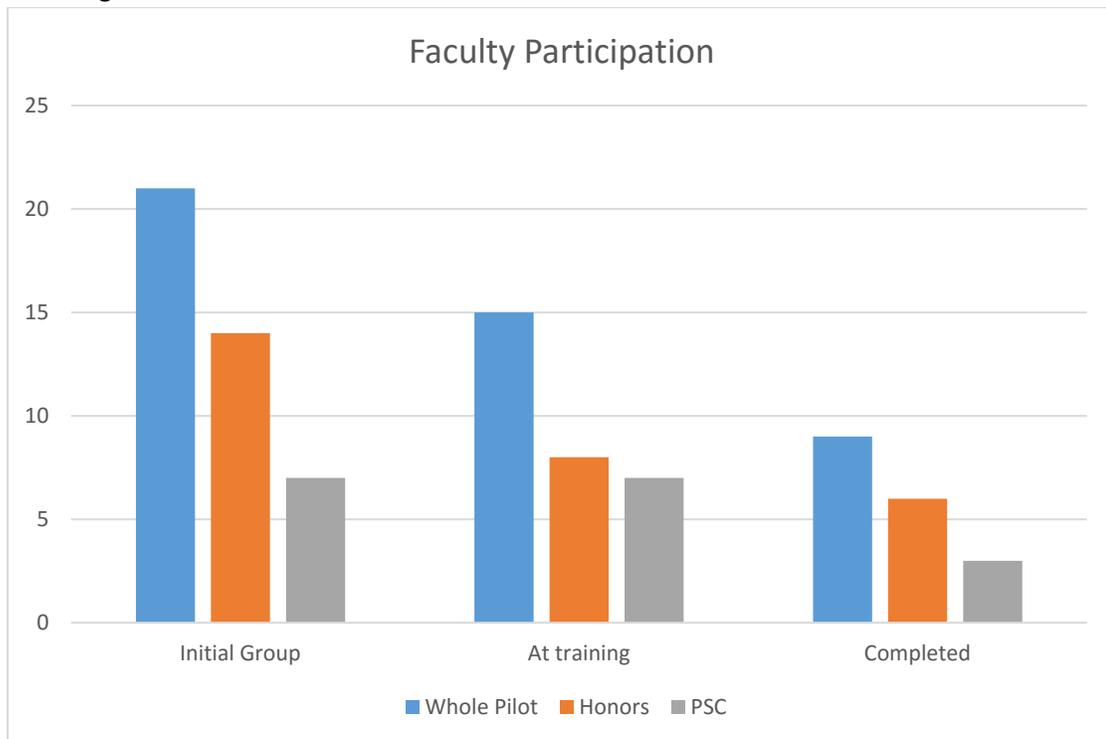
C. One goal was to determine the rate at which faculty could be expected to a) attend training sessions and then b) complete and report the assessment data. ***There were no expected rates for either.*** The goal was to begin forming a baseline of expected participation that could be used
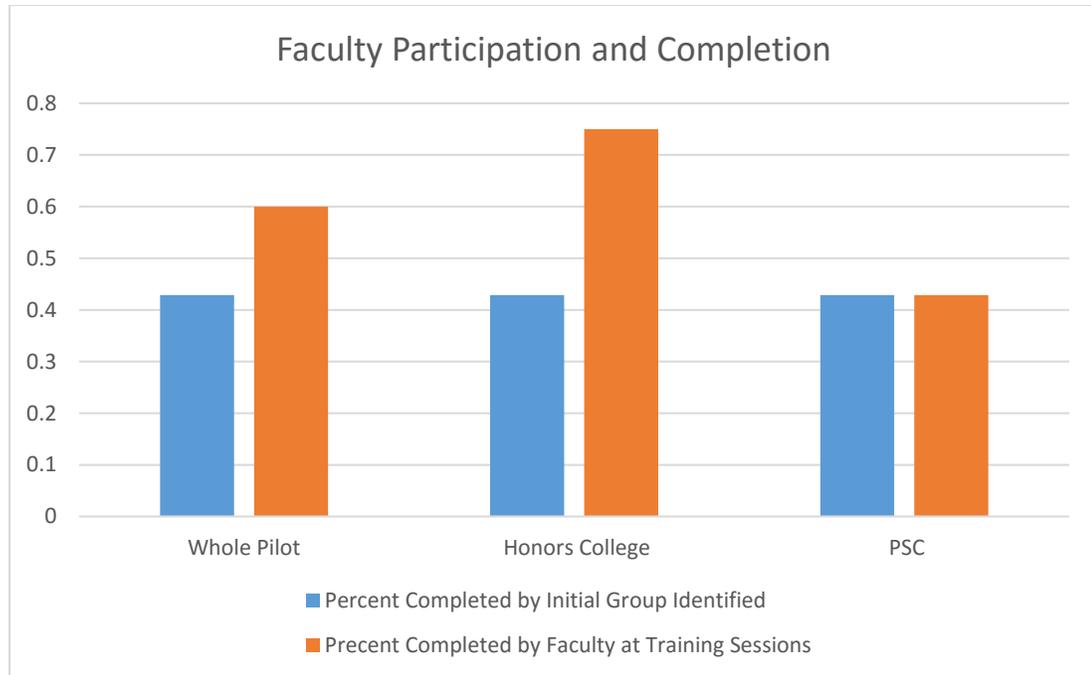
by assessment professionals to determine appropriate population sizes in order to get a usable sample size at the end of the assessment activity.

The initial pilot group identified 21 total potential participants between both the Honors College and Potomac State. 15 of those agreed to participate in the pilot once contacted. 9 of those who agreed to participate forwarded data. That's a 42.85% participation rate of all the faculty identified which isn't entirely unexpected; the better news is that 60% of those who actually attended the training session completed the assessments which is a workable share.

The Honors College completed 42.85% based upon the initially identified group (42.85% was a magic number in this pilot). Here again the good news was around those who attended the training sessions – 75% of those who attended the training sessions completed the assessments which is a great sign moving forward.

Potomac State faculty completed 42.85% based upon the initially identified group as well as for those who attended the sessions. The culture here is clearly different than at Honors College where there was a large initial group that opted out and thus a higher completion rate for those who attended the sessions. At PSC, everyone in the initially identified group attended the training session but then the final completion rate was the same overall as Honors but lower for those attending the sessions. In other words, the only real difference here is that it may be harder to gauge who will participate ultimately at PSC. The Honors group was much more self-selecting.

Faculty Participation — bar chart showing Whole Pilot, Honors, and PSC across Initial Group, At training, and Completed categories.

Faculty Participation and Completion

As far as delivery of faculty data to the Director of Academic Excellence and Assessment goes, the results were positive. Only one set of data was incorrectly reported which again underscores the effectiveness of the rather brief training session provided.

Faculty were slightly more likely to email their results than to use Sharepoint to upload their results. Moreover, the Sharepoint site saw remarkably little traffic and use. Resources were almost entirely unused while email requests for clarification or for resources were moderately frequent.

Action plan: 1) Collect another semester of data to confirm that faculty scoring norming is as successful as demonstrated here. 2) Collect another semester of data to establish a firm baseline of expected participation at both Morgantown and PSC. 3) Rethink use of Sharepoint and email communications to best facilitate and support the process. 4) Develop a Qualtrics reporting form to test.