Abstract

Higher education has made impressive progress concerning student learning outcomes assessment practices. Yet – despite the assumption that better assessment would lead to better student learning - few examples of demonstrable student learning improvement exist at the academic degree or university levels. In 2014 Fulcher, Good, Coleman, and Smith addressed this concern in a NILOA Occasional Paper titled, "A Simple Model for Learning Improvement: Weigh Pig, Feed Pig, Weigh Pig". The "pig paper" elucidated basic steps for improvement: collect baseline data, intervene effectively with curriculum and pedagogy, and then re-assess to examine if learning did indeed improve. The current paper builds upon these steps by offering standards for learning improvement. We articulate these standards via a rubric and a hypothetical learning improvement report. These tools are intended to elevate learning improvement conversations, and increase the number of learning improvement examples across higher education.

AUTHORS

Keston H. Fulcher, Ph.D.

Kristen L. Smith, M.A.

Elizabeth R.H. Sanchez, M.A.

Allison J. Ames, Ph.D.

Cara Meixner, Ph.D.
James Madison University

Return of the Pig:
Standards for Learning Improvement

I have long been frustrated with hollow statements about assessment's ability to improve higher education. While I am not as cynical as Erik Gilbert – who penned the 2015 *Chronicle* article, "Does Assessment Make Colleges Better? Who knows?" – I get his point. Colleges across the world spend substantial amounts of time and money on assessment, but scant evidence exists to justify the resources (Suskie, 2010). The good news is I believe assessment's state of affairs can be changed. Academe can do better. Collaborators Kristen Smith, Elizabeth Sanchez, Allison Ames, and Cara Meixner join me in an important step toward shifting higher education's focus away from empty assessment practice to something more edifying. We propose a fundamental pair of resources: a) a rubric detailing standards for learning improvement, and b) a learning improvement report from a hypothetical program annotated according to the rubric.

In 2014, Fulcher and Smith contributed to a National Institute for Learning Outcomes Assessment (NILOA) Occasional Paper titled "A Simple Model for Learning Improvement: Weigh Pig, Feed Pig, Weigh Pig" (Fulcher, Good, Coleman, & Smith, 2014). The pig analogy is an extension of the old farm saying: a pig never fattened because it was weighed. Our consensus is that the same logic applies to higher education: merely assessing, repeatedly "weighing" students, will not improve their learning. The Weigh Pig, Feed Pig, Weigh Pig paper elucidated the relationship between assessment and learning, each step essential to a simple model of improvement:

CORRESPONDENCE

Email
fulchekh@jmu.edu

- collection of baseline data,

- integration of faculty training and development,

- use of evidence-based teaching strategies,

- effective modifications to the learning environment, *and*

- re-assessment to affirm efficacy and evidence improvement.

This contribution to RPA expands the simple model by providing more explicit learning improvement standards. While we are under no presumption that higher education will use the exact standards we provide or the very rubric we've created, eventually, universal standards should be adopted; a universal rubric should be used.

We've organized this paper as six parts. First, we provide context for how standards, including rubrics, have articulated best practices in assessment. Second, we examine the current and future state of affairs regarding learning improvement. Third and fourth, we provide learning improvement standards via a learning improvement rubric that more carefully articulates and elaborates the standards in behavioral terms. Fifth, we provide an annotated hypothetical example of a learning improvement report. Finally, we provide suggestions for how academe could use these learning improvement resources for maximum effect.

We propose a fundamental pair of resources: a) a rubric detailing standards for learning improvement, and b) a learning improvement report from a hypothetical program annotated according to the rubric.

Before Improvement: Standards and Meta-Assessment Rubrics Advance Assessment Best Practice

Many books provide fine-grained details of an assessment cycle (e.g., Bresciani, Gardner, & Hickmott, 2009; Erwin, 1991; Palomba & Banta, 1999; Suskie, 2009; Walvoord, 2010). But on the conceptual end, the most noteworthy guidance for best practices in assessment comes from the now defunct American Association for Higher Education (AAHE). Despite being over two decades old, Hutchings, Ewell, and Banta (2012) observed that AAHE's nine Principles of Good Practice for Assessing Student Learning have aged incredibly well. We agree. Indeed, many of the principles highlight an underlying improvement emphasis within assessment.

While we have included the first few lines of the nine principles, they are worth reading (or re-reading) in their entirety:

1. The assessment of student learning begins with educational values.

2. Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time.

3. **Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes. Assessment is a goal-oriented process.**

4. **Assessment requires attention to outcomes but also and equally to the experiences that lead to those outcomes.**

5. Assessment works best when it is ongoing not episodic.

6. **Assessment fosters wider improvement when representatives from across the educational community are involved.**

7. **Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about.**

8. **Assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change.**

9. Through assessment, educators meet responsibilities to students and to the public. (Hutchings et al., 2012, para. 6)

Note, we added post-hoc emphasis to principles 3, 4, 6, 7, and 8, which emphasize the key points of this article.

Since AAHE's nine principles were published, the emerging practice of meta-assessment has helped further clarify what constitutes good assessment (Fulcher, Coleman, & Sundre, 2016). Meta-assessment commonly involves using a rubric to evaluate the quality of a report by providing detailed feedback on assessment processes and characteristics such as student learning outcomes (SLOs), methodology, results, and use of results. Each characteristic (or rubric criterion) is described at various quality levels (e.g., beginning, developing, good, and excellent) in the rubric (Popham, 1997).

Prominent organizations such as the New Leadership Alliance (NLA), Voluntary System of Accountability (VSA), NILOA, the Association of American Colleges & Universities (AAC&U), American Association of State Colleges and Universities (AASCU), and the Association of Public & Land-Grant Universities (APLU) have designed rubrics to evaluate the quality of institution-level assessment. While the NLA initiative was short-lived and the VSA, NILOA, AAC&U, AASCU, and APLU initiative is in its infancy, they represent the first national-level efforts to use a rubric to articulate assessment quality and standards of practice. These institution-level rubrics heavily emphasize issues such as communication of results and administrative and structural supports for assessment.

Unfortunately, most meta-assessment at the institutional or organizational level inadequately gauges student learning improvement. In other words, improvement is more than just a byproduct of good assessment processes. Quality assessment results, in our experience, rarely prompt change that improves learning. We agree with those who believe that assessment should be subsumed within a larger learning improvement system (Lumina Foundation, 2016).

Current State of Affairs for Improvement: More Assessment ≠ More Learning Improvement

Again, advances in assessment practice differ from improvements in student learning. While books, standards, and rubrics have refined assessment methodologies, today's practice still fails to capture the learning improvement spirit underlying AAHE's 20-year-old principles.

It comes as no surprise that we have witnessed and demonstrated few improvements in student learning at the academic program level of our institution. Only one of 14 criterion in our own award-winning (Willard, 2015) meta-assessment rubric mentions using assessment results for learning improvement. Unfortunately, faculty at our university may not be aware of what a successful improvement initiative may look like or the specific steps they can take to evidence learning — this information is not provided in the meta-assessment rubric we use.

Over the past few years, a handful of influential scholars have voiced the noted lack of evidenced learning improvement in higher education contexts (Banta, Jones, & Black, 2009; Blaich & Wise, 2011). Banta, Jones, and Black (2009) found that only six percent of the best assessment reports across the nation demonstrated student learning improvement. When Banta and Blaich (2011) were asked by *Change Magazine* to provide current examples of learning improvement, they could not find a sufficient number to write the article. The authors proceeded instead to write an article about obstacles to learning improvement (Banta & Blaich, 2011).

Quality assessment results, in our experience, rarely prompt change that improves learning.

To some readers, the current state of affairs might seem surprising given higher education's intentions of assessment. Indeed, in a survey conducted by NILOA in 2013, provosts reported that "commitment to institutional improvement" was ranked third of the 13 most important reasons to conduct assessment, falling just behind regional accreditation and program accreditation. In a more pointed question, provosts were asked how assessment results were used; "curriculum modification" and "institutional improvement" were reported as uses "quite a bit." Further, curriculum modification and institutional improvement were endorsed as a use of assessment results fifth and seventh, respectively, out of 16 total possible uses (e.g., accreditation, program review, institutional benchmarking, etc.). On average, provosts reported that they used assessment results for curriculum modification "quite a bit" and for institutional improvement between "some" and "quite a bit" (Kuh, Jankowski, Ikenberry, & Kinzie, 2014).

One reason for the apparent discrepancy between NILOA's survey results and the national lack of demonstrable gains in student learning is the inconsistent and vague definition of *improvement*. Many rubrics (including our institution's meta-assessment rubric), assessment-related books, and assessment measures use the term improvement in an imprecise way, as a synonym for change or perhaps as any use of results (Smith, Good, Sanchez, & Fulcher, 2015).

During the creation of our new learning improvement rubric, we followed the Weigh Pig, Feed Pig, Weigh Pig article's definition of evidencing learning improvement: "making a change to a program and then re-assessing to determine that the change positively influenced student learning" (Fulcher, Good, Coleman, & Smith, 2014, p. 4). Using this definition, we believe that program and/or institutional learning improvement occurs much less frequently than "quite a bit."

Pivoting Higher Education toward Learning Improvement

At the national level, organizations such as the Southern Association of Colleges and Schools Commission on Colleges (SACSCOC) have taken a stance on improving student learning. According to SACSCOC, "the concept of quality enhancement is at the heart of the Commission's philosophy of accreditation" (SACSCOC, 2016, para. 3). More specifically, SACSCOC requires every institution seeking reaffirmation of accreditation to engage in a campus-wide initiative to enhance student learning (i.e., a Quality Enhancement Plan or QEP). Each QEP includes processes for identifying issues or needs that emerge from institutional assessment efforts. Then, the institution must create a plan to enhance student learning and/or the environments supporting student learning, which includes determining specific goals and assessment strategies.

Although isolated success stories can be identified at both the national and institutional levels, our aim is to further cultivate learning improvement examples and more generally to elevate the learning improvement conversation.

Another national organization, NILOA, offers advice regarding building strategies to intervene through assignment design (Hutchings, Jankowski, & Ewell, 2014). Other organizations, such as the State Higher Education Executive Officers Association (SHEEO) and AAC&U, developed the Multi-State Collaborative, in which colleges use rubrics to assess various learning areas and encourage processes that change curriculum and pedagogy in intentional ways (http://www.sheeo.org/projects/msc-multi-state-collaborative-advance-learning-outcomes-assessment). Charles Blaich and Kathy Wise, of the Center of Inquiry (http://www.liberalarts.wabash.edu/), consistently provide great contributions through forward-thought and institutional support and their efforts exemplify how faculty, administrators, and students can work together to use assessment data to influence improvement.

At the state level, institutions are preparing to incorporate models of learning improvement. In Virginia, the Virginia Assessment Group (VAG) and the State Council of Higher Education for Virginia (SCHEV) are encouraging such initiatives. Virginia Commonwealth University and Longwood University are beginning the piloting process. Meanwhile, in Georgia, Kennesaw State University has recently required improvement reports in addition to traditional assessment reports; programs must now re-assess to determine the efficacy of their changes. At James Madison University, we have piloted several learning improvement initiatives under the guide of the simple model (Fulcher et al., 2014) but using a more complex definition:

> 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 (Fulcher, Sundre, Russell, Good, & Smith, 2015, p. 3).

One academic program, Computer Information Systems, has already demonstrated large gains at the program level. A campus-wide project, The Madison Collaborative, has shown university-wide improvements in students' ethical reasoning skills. These success stories are attributed to partnerships between faculty leaders, assessment experts, faculty developers, and administrators who collectively worked to implement the simple model.

Although isolated success stories can be identified at both the national and institutional levels, our aim is to further cultivate learning improvement examples and more generally to elevate the learning improvement conversation. To do so, we provide learning improvement standards and the rationale for each in the spirit of the well-conceptualized AAHE principles. These standards for learning improvement are embodied and elaborated via a learning improvement rubric. This rubric can be used to guide and evaluate learning improvement initiatives. To show how the learning improvement rubric can be applied, we provide an annotated example of a report. We conclude with suggestions for how practitioners can use these two resources (i.e., the learning improvement rubric and learning improvement report) such that examples of demonstrable learning improvement become the norm for higher education, not the exception.

The Learning Improvement Rubric: Six Standards of Successful Learning Improvement

Our learning improvement rubric was crafted by Fulcher, Smith, and Sanchez throughout a semester-long independent course; vetted by assessment practitioners and faculty development experts; and is supported by extensive research and a few pilot initiatives. In our experience, all six standards detailed in the rubric are common in and necessary for successfully demonstrating program-level learning improvement.

Some terminology used in the rubric and example report will be new to most readers. Borrowing from our colleagues in computer information systems at JMU and faculty development, we have adopted the terms *as is* [*curriculum* or *environment*] and *to be* [*curriculum* or *environment*]. In this context, the *as is* curriculum or environment is what a program is or is not doing to meet the selected student learning outcome (SLO) before starting a learning improvement initiative; the *to be* curriculum or environment is the culmination of all proposed changes expected to improve student learning.

A. Faculty Involvement: Faculty are participating throughout the learning improvement initiative.

Ideally, improved learning at the program level means that all graduating students are better educated. Therefore, changes in individual courses must be connected and aligned by multiple faculty members.

Faculty contribution, buy-in, and engagement are crucial to the success of program-level improvement processes for several reasons (Shavelson, 2010). First, faculty are responsible for the success of their individual classes, which are part of a broader curriculum and academic program. Similarly, faculty serve as a vital bridge between students and knowledge/skill acquisition (Kuh, Ikenberry, Jankowski, Cain, Ewell, Hutchings, & Kinzie, 2015). They are the frontline facilitators and shapers of students' classroom learning experiences.

Ideally, improved learning at the program level means that all graduating students are better educated. Therefore, changes in individual courses must be connected and aligned by multiple faculty members. Such alignment necessitates buy-in and engagement from a dedicated cadre. For clarification, we don't mean that all faculty in a program have to work directly on the learning improvement project for it to be effective. However, the majority of those whose classes cover the selected SLO should be active, dedicated participants. When faculty involvement is maximized and championed, faculty can more effectively:

- create powerful teaching and learning strategies;
- determine how, when, and where changes to the program should be implemented;
- deliver new curricula to all students as intended to be delivered; and
- appropriately connect new learning experiences across different courses within the program in a way that facilitates learning.

Consider students in a hypothetical communications program at XYZ University who struggle to effectively deliver oral presentations with an engaging introduction, logical and fluid body, and smooth conclusion that reinforces the main ideas of the presentation (i.e., the selected SLO).

To remedy this learning deficit, faculty from the communications program could create an improvement initiative in hopes of bolstering their students' abilities to effectively deliver a presentation. If only two or three faculty in the program of 20 decide to get involved, planning program-level changes would be nearly impossible with so many faculty non-contributors. The chances are good that the oral presentation student learning outcome (SLO) is or would be covered in courses that are taught by faculty not participating in the improvement initiative. When the communications faculty conducts program-level assessments, any learning improvements of the few students who experienced a modified course would likely be washed out by the remaining students who were taught by faculty members who *did not* change their courses in efforts to improve oral presentation. Ultimately, program-level learning improvement cannot be achieved without a "program-level's worth" of faculty participation. Figure 1 depicts the faculty involvement criterion of the learning improvement rubric.

Any learning improvement initiative is a substantial undertaking. For this reason, we highly recommend focusing on one student learning outcome (SLO) at a time.

A. Faculty Involvement: Faculty are participating throughout the learning improvement initiative.					
	0 **Absent**	**1** **Beginning**	**2** **Developing**	**3** **Good**	**4** **Exemplary**
1. **Relevant Faculty Involvement**	No evidence of faculty involvement *or* degree of faculty involvement is unclear.	Evidence of ~50% of faculty involvement through *most* aspects of the learning improvement initiative.*	Evidence of 51-74% of faculty involvement through *most* aspects of the learning improvement initiative.*	Evidence of 75-89% of faculty involvement through *every* aspect of the learning improvement initiative.*	Clear evidence (e.g., faculty surveys, progress checks, etc.) of 90-100% of faculty involvement through *every* aspect of the learning improvement initiative.*
Aspects of the learning improvement initiative include: student learning outcome (SLO) selection, elaboration, and assessment alignment/match; baseline data collection; investigation of original program; program intervention; and re-assessment.					

Figure 1. Faculty involvement criterion of the learning improvement rubric.

B. Readiness: The program can provide an explanation of why the SLO was selected for improvement, what the outcome means, and how learning and improvement is and will be measured.

Any learning improvement initiative is a substantial undertaking. For this reason, we highly recommend focusing on one student learning outcome (SLO) at a time. Attempting to demonstrably improve more than one SLO in a given year may quickly become overwhelming and discouraging to faculty who are investing their time and expertise to the learning improvement initiative. Selecting an SLO that students are not achieving satisfactorily makes learning improvement manageable and will hopefully focus faculty contributions.

Before making any pedagogical or curricular changes, it is imperative that faculty elaborate the selected (*or targeted*) SLO; doing so will help faculty:

- gain a common, specific understanding of what needs improvement;
- tightly connect and align new pedagogy and curricula with the targeted SLO and assessment instrumentation; and
- provide a framework by which to evaluate the *as is* learning environment and the *to be* learning environment.

Without a clearly detailed or elaborated SLO, faculty will have only a very loose idea of the specific skills, knowledge, or abilities that they want to measure and improve. It may be easy for faculty to want to skip SLO elaboration in favor of immediately implementing a novel curricular change or modifying existing pedagogies. However, a nebulous or vague SLO is detrimental to the success of a student learning initiative because faculty might have different conceptualizations of what the SLO means, how it is best taught, and how it is appropriately measured. Figure 2 depicts the readiness criterion of the learning improvement rubric.

B. **Readiness: The program can provide an explanation of why the SLO was selected for improvement, what the outcome means, and how learning and improvement is and will be measured.**					
	0 **Absent**	**1** **Beginning**	**2** **Developing**	**3** **Good**	**4** **Exemplary**
1. *Program-level* **SLO Selection:** **Rationale with** **Evidence**	No rationale or evidence is provided for why the SLO was selected for improvement.	Rationale provided with *anecdotal* evidence *or* evidence doesn't clearly support the need for SLO improvement initiatives.	Rationale provided with evidence from *one* source* and clearly supports the need for SLO improvement initiatives.	Rationale provided with evidence from *two* sources* and clearly supports the need for SLO improvement initiatives.	Meets criteria for Good and cites literature suggesting a need to achieve the selected SLO that extends beyond the university. SLO must be a higher-order thinking skill (e.g., critical thinking).
***Sources may include: Direct measures (e.g., previous test scores); a survey of current students; alumni surveys; program reviews; accreditation reports; insights from experts in the field; or employer feedback.**					
2. **SLO** **Elaboration:** **Ascription of** **Meaning**	The SLO is not elaborated (e.g., Students will perform critical thinking.) *or* is not stated in student terms (e.g., Program faculty…).	The SLO is elaborated to a minor degree. (e.g., Students will perform critical thinking, which includes: analysis & evaluation).	The SLO is elaborated. Nevertheless, no explanation of the process is provided. (e.g., Students will perform critical thinking, which includes: analysis & evaluation. Analysis involves…and evaluation means…).	The SLO is elaborated and an explanation of the process is provided. (e.g., faculty were polled to determine how "analysis" should be defined and then…; the meaning of "evaluation" was ascribed through the process of…).	Meets criteria for Good and the elaboration is supported by external experts (e.g., definitions were vetted in job analysis interviews; meaning was supported in an extensive literature review, etc.).
3. Alignment/ **Match:** **Elaborated** **SLO and** **Assessment** **Measures**	No relationship between SLO and assessment is provided.	Vaguely references SLO and assessment alignment/match but it is unclear to what degree. (e.g., Faculty wrote items to match the definition of "analysis;" this test was chosen because of its match to our selected "evaluation" process).	Evidence of SLO and assessment alignment/ match. (e.g., a rubric or set of test items are provided that match elaborated SLO).	Meets criteria for Developing and is affirmed via backward design (i.e., provides a sample of course assignments that were created based on test items) or behavioral anchoring (qualitative examples of performance standards for rubric scores).	Meets criteria for Good and is affirmed by external experts through similar processes.

Figure 2. Readiness criterion of the learning improvement rubric.

Like with many evolutions, the assessment practice evolution is best described as a change of emphasis as opposed to a radical revolution.

Consider, again, faculty members from a hypothetical communications program at XYZ University who engage in a learning improvement initiative to help their students effectively deliver an oral presentation (in addition to another four learning outcomes that have little to do with oral presentation delivery skills). Moreover, they did not detail the meaning of any of the SLOs that they targeted for their learning improvement initiative.

As communications faculty attempt to discuss current learning interventions in the as is curriculum, they may disagree and lack consensus. Indeed, the sheer amount of time it takes to discuss all of the SLOs becomes antithetical to advancing the learning improvement initiative. Furthermore, some faculty may think the SLOs have one meaning, while other faculty members conceptualize the SLOs in a completely different way. The discussions may eventually become circular because no one has defined or detailed what the SLOs really mean. All said, it may prove impossible to develop student knowledge, skills, and abilities when faculty have not elaborated the precise student knowledge, skills, and abilities they were trying to improve in the first place. They subsequently realize that the amount of time and resources

it takes to define and elaborate one SLO makes improving multiple SLOs at once unfeasible.

C. Baseline data collection and measurement quality: Student performance is measured before program-level changes are made; high quality measurement is ensured.

When making claims related to improvement or growth in educational contexts, rigorous methodology is a necessity. Perhaps the most often forgotten or neglected part of evidencing learning improvement is that faculty must collect baseline data *before* any pedagogy, curriculum, or course sequencing changes are made in the program. Collecting baseline data allows for a more meaningful comparison with post-learning improvement initiative data.

Evidencing improvement requires data on students' knowledge, skills, or abilities before and after changes are made. Faculty should ensure that they are measuring the targeted SLO in a way that yields reliable and valid scores. Note, the need for baseline data and adequate instrumentation are widely accepted characteristics of best practices in student learning outcomes assessment contexts. Having strong methodology will help faculty:

- demonstrate that students' assessment scores are reliable; and
- support claims of learning improvement (i.e., defend validity of improvement inference).

Without baseline data collection, programs and their partners in assessment and faculty development cannot empirically demonstrate that student learning has improved. Further, if the team does not use assessment tools that produce valid and reliable scores to measure the targeted SLO, they cannot capture requisite data on the specific and intended knowledge, skills, and abilities. Without using instruments that produce psychometrically sound data, the conclusions or inferences made from those assessment scores lack trustworthiness.

Imagine that faculty members from the hypothetical communications program at XYZ University started implementing new teaching strategies and targeted assignments before they collected data related to students' oral presentation abilities. Any data collected after the changes were made will not be at a true baseline – student learning may have already improved given the new course experiences. In other words, faculty will have a lack of data regarding how well their students were achieving the targeted SLO before they made changes to the curriculum; further, when collecting data after all new implementations are in place, the program cannot ascertain whether or to what degree students' oral presentation skills actually improved as a result of the learning improvement initiative.

Additionally, if the program measured students' oral presentation skills, but used a poorly designed instrument (e.g., an instrument that does not have desirable psychometric properties) there is no validity evidence to suggest that the assessment scores are meaningful. The inferences or conclusions of the initiative that communications program faculty attempt to make based on assessment scores will be severely compromised. Figure 3 depicts the baseline data collection and measurement section of the learning improvement rubric.

Certainly, universities with established assessment mechanics may have an easier time launching improvement initiatives. However, programs with fewer supports and resources should not feel discouraged. In fact, we recommend that institutions focus on one or two programs to pilot first. Having success at a small level will beget more successful learning improvement initiatives in the future.

D. Investigate curriculum and diagnose issues: An explanation of hypotheses for why what was originally being taught in the curriculum relative to the targeted SLO was ineffective.

A key component of demonstrable student learning is changing pedagogies and the curriculum—a learning *intervention* (Fulcher, Good, Coleman, & Smith, 2014). Before faculty can create meaningful, effective learning interventions, they must first investigate and

C. Baseline data collection and measurement quality: Student performance is measured before program-level changes are made; high quality measurement is ensured.				
	0 **Absent**	**1** **Beginning**	**2** **Developing**	**3** **Good**	**4** **Exemplary**
1. **Baseline Data Collection: Timing and Sampling**	No baseline data collected prior to intervention *or* assessment measures were indirect (e.g., course grades).	Baseline data collected using direct measures. However, the sample is unrepresentative (i.e., only a few students were assessed, students had low motivation to perform well, little effort was given on assessment, etc.) *or* not enough information is provided to determine veracity of the data collection process (i.e., the number or level of motivation of assessed students is unknown).	Baseline data collected using direct measures; However, the sample is somewhat unrepresentative (i.e., only some students were assessed, student motivation/effort varied when being assessed) *or* the data provided are not a true baseline (i.e., some intervention was already implemented at the time of assessment). However, steps are taken to address these issues (e.g., motivation filtering: students who showed little effort were not included in data sample; evidence provided that shows data sample is representative although sampling technique wasn't the best, etc.).	Baseline data collected using direct measures; sample is fairly representative and includes details (e.g., number of students assessed, motivation analysis, etc.). However, minor issues still exist (e.g., only 80% of students assessed).	Baseline data collected using direct measures before any intervention; The sample is representative; details provide evidence of good motivation (high-stakes testing) and defensible testing conditions (e.g., student questions about the assessment are answered, etc.).
2. **Psychometrics: Reliability and Validity**	No mention of reliability or the activities associated with reliability (i.e., using multiple raters for rubrics, training raters to use a rubric, etc.).	Evidence of activities associated with good reliability provided but reliability coefficients are absent or low *or* reliability analyses lack rigor (e.g., rater agreement is not exact).	Evidence of appropriate reliability that controls for chance provided. Marginal reliability rates are revealed.	Evidence of appropriate reliability that controls for chance provided. Passable reliability rates at the group level are revealed.	Evidence of appropriate reliability *and* validity (e.g., scores from assessment behave in predicted ways according to theories such as group studies or modeling).

Figure 3. Baseline data collection and measurement quality criterion of the learning improvement rubric.

understand the program as is – how and what information is being taught throughout the curriculum before any changes are made (Chaplot, Booth, & Johnstone, 2010).

Notions of "understanding about what is happening and what needs to happen to advance student success" and "defining problems and solutions" are emphasized in the Applied Inquiry Framework (Chaplot, Booth, & Johnstone, 2010, p. 10). Indeed, programs may find it difficult to make intentional, informed changes to the curriculum if faculty are unaware of why students are struggling to achieve an SLO.

There could be multiple reasons students are not achieving intended learning outcomes: A simple (and obvious) explanation is often that no faculty are covering the SLO material. More complexly, faculty may be covering the SLO material using ineffective pedagogies or the assignments and learning experiences are misaligned to the learning outcome's level (e.g., faculty are teaching students to *recall* facts when the SLO intends students to be able to *synthesize* the material). Indeed, sometimes the explanation for why students are not meeting

the learning outcome is more complicated. Perhaps students are taught the SLO material only in a lower-level course; the content is not reinforced or practiced later on. Over time, students' knowledge or skills deteriorate because content is not properly *scaffolded* (strategically covered) across courses within the major. Or, maybe the new content and teaching strategies in the classroom are not well implemented or received—the learning intended to take place never actually does. In such instances, qualitative data from students about their educational experience in the program can be invaluable to helping faculty make informed, meaningful changes. Investigating and diagnosing the curriculum, an often overlooked practice, is important because the process:

- allows faculty to become intimately familiar with the educational experience students are getting throughout the entire program, not just in the isolated courses or sections that they teach;

- helps faculty identify specific reasons students may not be achieving intended learning outcomes so that they can make more informed and intentional modifications;

- creates a space to discuss shortcomings in the delivery of the new curriculum (e.g., course sequencing or scaffolding, etc.);

- facilitates re-alignment of learning outcomes and assessment instruments; and

- encourages and values student involvement when faculty collect and use qualitative data from students regarding their educational experiences in the program.

Consider the outcome if programs do not investigate their *as is* curriculum and diagnose issues that might contribute to students' lack of success. Faculty may propose modifications to a curriculum based on personal hunches or their own perceptions, but student feedback can contribute to a more robust understanding of how to improve learning. Figure 4 depicts the learning improvement rubric criterion related to investigating curriculum and diagnosing issues.

Let us return once more to the faculty members from the hypothetical communications program at XYZ University who engaged in a learning improvement initiative to help their students effectively deliver an oral presentation. The communications faculty decide to use a different textbook to teach the oral presentation material because the textbook they were using in the past was "dated" and "students needed more modern examples of how to effectively deliver an oral presentation with an engaging introduction, logical and fluid body, and smooth conclusion that reinforces the main ideas of the presentation."

Had the faculty conducted a qualitative study or otherwise collected feedback, they would have learned from their students that the way students are taught to deliver a presentation, craft engaging introductions, and reinforce main ideas in the conclusion is mostly effective but it is more practice, in more courses, that students need. Also, students may have explained that they needed more information on how to deliver a presentation in different contexts and settings, as well as more detailed feedback and a clearer explanation of what "high quality" oral presentations entail. Without knowing that students need more demonstrations of high quality speeches, assistance in making speeches in different context, and time to practice and develop their skills throughout the program, crafting new content and teaching strategies to fit the areas of improvement is nearly impossible.

E. Learning Intervention: The program establishes an appropriate timeline for faculty development, instrument development, and intervention; the intervention is implemented with fidelity; necessary adjustments are made.

Through two learning improvement pilot projects, we have found that successful learning improvement initiatives take multiple years and long-term planning. For instance, one of the pilot programs spent approximately two years before they could evidence initial student learning improvements. Faculty from this program must wait an additional year before

Perhaps the most often forgotten or neglected part of evidencing learning improvement is that faculty must collect baseline data *before* any pedagogy, curriculum, or course sequencing changes are made in the program. Collecting baseline data allows for a more meaningful comparison with post-learning improvement initiative data.

D. Investigate curriculum and diagnose issues: An explanation of hypotheses for why what was originally being taught in the curriculum relative to the targeted SLO was ineffective.					
	0 **Absent**	**1** **Beginning**	**2** **Developing**	**3** **Good**	**4** **Exemplary**
1. Investigation: Program-level Curriculum Relative to SLO	No information provided regarding if, or to what degree, the selected SLO is covered in the "as is" curriculum.	A program-level curriculum map is provided; the map shows the courses in which the SLO is covered and indicates the theoretical intensity or degree of coverage (i.e., SLO content is primarily taught in course X, time spent covering content is X, assignments related to SLO are...).*	Meets criteria for Beginning *and* provides a study of faculty to determine if they agree, at the program-level, that the theoretical curriculum matches the actual curriculum.*	Meets criteria for Developing *and* a scaffolding of the curriculum is provided (i.e., the typical student is taught the SLO content in the following courses to the following degrees…).*	Meets criteria for Good *and* feedback from a representative sample of students about their experiences in the curriculum in regards to the SLO is provided (e.g., through a survey and/or focus group, we found that…).*
*Note: If investigation accurately reveals little or no original content coverage for the selected SLO, programs should receive a score of 3 (Good) for this criterion.					
2. Investigation: Individual Course-level Coverage of SLO Content	No information provided regarding if, or to what degree, specific courses cover SLO content.	Faculty participating in the intervention provide some details regarding course-level learning experiences** covering the SLO content *but* investigation is cursory (i.e., doesn't include enough sections/classes to be representative).	Faculty participating in the intervention provide comprehensive, qualitative description of the learning experiences** and how they align with the SLO content.	Meets criteria for Developing *and* includes approximately how much time students spend with the identified learning experiences.*	Meets criteria for Good *and* provides evidence that faculty talked to a representative sample of students about the effectiveness of their learning experiences regarding the SLO.**
**Course-level learning experiences can be identified as/through assessments, class activities, teaching styles, syllabi, etc.					
3. Investigation Conclusions: Logical Insights about Why Efforts are not as Effective as Intended	No insights provided.	Insights provided but do not flow logically from the investigation.	Informed, logical insights provided that point to strengths (if there are any) and weaknesses of how SLO is addressed at *either* the program *or* the course level, but not both.	Informed, logical insights provided that point to strengths (if there are any) and weaknesses of how SLO is addressed at *both* the program *and* the course level.	Meets criteria for Good *and* the insights have been vetted through students *and* external experts or stakeholders.

Figure 4. Investigating curriculum and diagnosing issues criterion of the learning improvement rubric.
they can fully implement their learning initiative and hopefully demonstrate the learning

they can fully implement their learning initiative and hopefully demonstrate the learning improvements they had planned.. A thorough timeline is critical to keeping track of the learning improvement initiative's route across multiple semesters.

Making changes to a curriculum can include modifying the techniques or strategies faculty use to teach SLO material, introducing new material, changing when/where in the program students learn the content (e.g., course scaffolding), and more. To increase the chances of these changes becoming improvements, faculty participating in the initiative should be well-supported and prepared. Although it is certainly possible for faculty to change their own teaching pedagogies, redesign their courses, and better align course content with the targeted SLO, most will need some assistance before doing so (Fink, 2003). Thus, before faculty can be more effective in delivering SLO material, they would benefit from professional development. Understanding different kinds of pedagogical techniques, learning theories, and course designs

may help equip faculty participating in the initiative. Many peers, faculty development centers, conferences, and online workshops are valuable resources. It is important to remember that changes to both the faculty and student experiences in the program are necessary to evidence learning improvement; that baseline data collection from a well-developed assessment instrument(s) needs to occur before changes are made to the program; and that changing an entire curriculum is a time consuming and intricate endeavor. Figure 5 depicts the learning improvement rubric criterion related to learning interventions.

Through two learning improvement pilot projects, we have found that successful learning improvement initiatives take multiple years and long-term planning.

E1. Learning Intervention: The program establishes an appropriate timeline for faculty development, instrument development, and intervention; the intervention is implemented with fidelity; necessary adjustments are made.				
	0 **Absent**	**1** **Beginning**	**2** **Developing**	**3** **Good**	**4** **Exemplary**
1. **Percentage of Students in Program Affected by Intervention**	No evidence of what % of students affected *or* only 0-25% of students affected.	26-49% of students affected.	50-74% of students affected.	75-89% of students affected.	90-100% of students affected.
2. **Program-Level Intervention**	No evidence of intervention at the program-level.	Some evidence that intervention is being implemented at the program-level but it is unclear which courses are targeted or at what intensity.	A revised curriculum map is provided that shows where critical elements of the SLO are newly covered and at what intensity.	Meets criteria for Developing *and* provides compelling explanation for scaffolding (e.g., explains where and why elements of SLO are introduced, reinforced, and mastered).	Meets criteria for Good *and* is validated by external experts and students.
3. **Individual Course-Level Intervention**	No evidence of intervention at the individual course level.	Some evidence that intervention is being implemented in individual courses but one or more of the following problems exist: the % of courses that are affected is unclear, 50% or fewer of targeted courses are affected *or* course-level implementation does not match program-level plan.	51%-74% of faculty participating in the intervention have course student learning outcomes (C-SLO) aligned with the appropriate degree of coverage of the selected SLO; learning experiences for each C-SLO in every course is provided; alignments are made clear in syllabi.	Meets criteria for Developing *except* 75%-89% or more of faculty participating in the intervention have course student learning outcomes (C-SLOs) aligned with the appropriate level of the selected SLO.	Meets criteria for Good *except* 90%-100% of faculty participating in the intervention have C-SLOs aligned with the appropriate level of the selected SLO.
4. **Faculty Development for Intervention**	No evidence of faculty development or preparation for intervention.	Vague references to faculty preparation (e.g., we discussed implementation in a meeting, materials were distributed to faculty).	Description of faculty development processes are clear, yet, the process is insufficient relative to what is needed to effectively intervene (i.e., "faculty spent 1 hour in training session;" clearly not enough time for most program-level initiatives).	Faculty development processes are clear and sufficient; faculty member strengths are drawn upon (e.g., program disseminates teaching/ pedagogy knowledge; consults with *either* an SLO expert or curriculum /pedagogical expert).	Meets criteria for Good *and* faculty consulted with outside experts related to the SLO as well as curriculum/ pedagogical design experts.

Figure 5. Learning intervention criterion of the learning improvement rubric.

After an appropriate assessment instrument is developed or found (a complex and involved process in-and-of itself) and baseline data are collected, faculty can begin to make changes to the program. In order to systematically evaluate how successfully the changes are being implemented, programs can collect what assessment practitioners call implementation fidelity data. That is, the program can measure the extent to which what is delivered in the classroom differs from what was planned or intended (Fisher, Smith, Finney, & Pinder, 2014; Gerstner & Finney, 2013; O'Donnell, 2008). Implementation fidelity is an advanced technique

that is an incredibly helpful tool yet, grossly underused in higher education. Figure 6 depicts the learning improvement rubric criterion related to interventions.

Creating a timeline, preparing faculty participants, and evaluating progress through implementation fidelity data collection will help programs:

- establish realistic expectations about when and how student learning can improve;

- keep a schedule of data collection and implementation;

- organize efforts to ensure that changes are made at appropriate times and places within the program;

- discover different methods to implementing new pedagogy and curricula;

- make inferences from assessment scores about student learning improvements; and

- identify what changes or modifications can be made to the planned initiative.

E2. Intervention: The program establishes an appropriate timeline for faculty development, instrument development, and intervention; the intervention is implemented with fidelity; necessary adjustments are made.					
	0 **Absent**	**1** **Beginning**	**2** **Developing**	**3** **Good**	**4** **Exemplary**
5. Intervention Timeline	No timeline for assessment and intervention is provided.	References assessment and intervention. Not enough details provided to discern sequencing of interventions and assessments *or* the sequencing departs from assess, intervene, re-asses.	Notes pre-assessment, in what courses the intervention occurs, and post-assessment; the intervention strength (degree of implementation), however, is not laid out.	Notes pre-assessment, in what courses the intervention occurs, and post-assessment; clearly indicates degree of implementation throughout; *but* time-sequencing may be too ambitious (i.e., does not sufficiently account for instrument construction or faculty development).	Notes pre-assessment, in what courses the intervention occurs, and post-assessment; indicates degree of implementation throughout; clear and well laid out; do-able (i.e., accounts for time to ensure instrument is well-developed; time for faculty development).
6. Intervention Implementa-tion Fidelity: Quality and Adjustments	No context or updates regarding implementa-tion of intervention provided.	Cursory information provided; not enough to make inferences about implementation quality. For example, "the program was implemented according to the timeline" with no further explanation *or* the implemented intervention differs substantially from the planned intervention.	Attempts to investigate the quality of implementation, occasionally bringing up insights of what is changing and where. Nevertheless, the process of checking fidelity is not strong enough to give a full picture of what is actually happening at the program-level.	Systematically investigates program-level fidelity (e.g., surveys of students) throughout intervention implementation. Provides details regarding the match between the planned and actual implementation, yearly. In addition, where issues of implementation arise, program makes or states adjustments for subsequent years.	Systematically investigates course (e.g., auditing) *and* program-level fidelity throughout intervention implementation. Provides details regarding the match between the planned and actual implementation, yearly. In addition, where issues of implementation or process arise, program makes or states adjustments for subsequent years.

Figure 6. Intervention criterion of the learning improvement rubric.

Again, using the faculty from the communications program at XYZ University who want to improve their students' abilities to effectively deliver an oral presentation as an example, we can demonstrate the importance of allocating time for faculty preparation and instrument development. The hypothetical communications faculty are committed to enhancing their students' oral presentation skills—each member is taking ownership over a well-specified content area and some are researching new ways to teach the material.

However, as the semester begins and the faculty workload becomes more demanding, many of the faculty participants revert back to their original teaching styles and material. Without time allocated for faculty development, few faculty were able to identify and integrate more effective pedagogical techniques for teaching oral presentation skills. As schedules fill up, the faculty meet less frequently. The absence of an implementation timeline contributes to efforts that are inconsistent and uncoordinated, changes that are implemented before baseline data are collected, and progress that is not tracked. As a result, no demonstrable learning improvement exists. The communications faculty have every right to feel disheartened by a lack of success.

F. Re-assess: The impact of the intervention is measured; program-level changes contribute to improvements in student learning.

The final step in the learning improvement process is re-assessment (Fulcher, Good, Coleman, & Smith, 2014; Walvoord, 2010). To demonstrate learning improvement, students must be assessed both before and after they have experienced changes that were made to pedagogy and curricula. Note, in order to demonstrate program-level improvement, re-assessment must take place at the program-level using the same instruments and methodology used to collect baseline data (e.g., standard C). Figure 7 depicts the learning improvement rubric criterion related to re-assessment. Re-assessing students allows faculty to:

- empirically demonstrate that student learning has improved via outcomes assessment data;

- precisely articulate how much or to what extent student learning has improved from pre-intervention (e.g., baseline) to post-intervention;

- integrate outcomes assessment with implementation fidelity data to further tweak or refine any aspects of the learning intervention; and

- use fidelity data to further investigate potential (in)efficiencies of specific features of the learning intervention.

Unfortunately, programs often fail to re-assess student learning after changes are made. Faculty and assessment practitioners alike often mistakenly think that because changes are made to the curriculum or assessment instrument, student learning improves. As Fulcher and colleagues (2014) describe:

> They [faculty] make statements like, "We made x, y, and z improvements to the program." But they really mean that they made x, y, and z changes. A change is only an improvement when one can demonstrate its positive effect on student learning. (p. 4).

Imagine the faculty from the hypothetical communications program at XYZ University made changes to their curriculum and pedagogy in an attempt to improve their students' abilities to effectively deliver an oral presentation. Imagine, they also collected baseline data, implemented their agreed upon learning intervention, and now claim students are better at effectively delivering an oral presentation. This could be the case, but because they did not re-assess students' oral presentation abilities after all of the learning modifications or interventions were implemented, faculty have no data or empirical evidence of learning improvements. They have little idea of how well (or poorly) their learning improvement efforts paid off.

We've tried to provide a detailed explanation of our six learning improvement standards. We've also included an example learning improvement report (see Appendix) for the same hypothetical communications program. In the learning improvement report, six

F. Re-assess: The impact of the intervention is measured; program-level changes contribute to improvements in student learning.					
	0 **Absent**	**1** **Beginning**	**2** **Developing**	**3** **Good**	**4** **Exemplary**
1. Re-assess SLO: Same Assessment Measures Used	No pre-intervention data *and/or* no post-intervention data were collected; therefore, re-assessment is impossible.	Pre-intervention and post-intervention data reported; nevertheless, not enough information is provided to evaluate the veracity of the results (e.g., the methodology from pre- to post- data collection changed to the degree that comparisons are meaningless: different sampling schemes, etc.) *or* a different assessment measure is used *or* (test/rubric) items within the same measure change moderately.	Pre-intervention and post-intervention data reported; methodology changed moderately but steps were taken to mitigate those changes (e.g., using SAT scores as a covariate to adjust for group differences). Comparison of pre- and post- data may make some sense but the integrity is compromised.	Pre-intervention and post-intervention data reported. The collection process was reasonable and the method stayed faithful (which could include using psychometrically-verified equivalent test forms).	Post-intervention data reported; evidence that the assessment measures were consistent and data collection processes (i.e., sampling) enable "apples to apples" comparison between pre- (potentially mid-) and post-intervention cohorts.
2. Magnitude of Student Learning Improvement: Statistical Gains	No information regarding student learning improvement is provided.	Student learning improvement (gain) is provided but not in a standardized way *or* the gain is not statistically significant *or* the effect size is negligible: Cohen's d of < .15.	Gain is statistically significant and the effect is small: Cohen's d of ~.3.	Gain is statistically significant and the effect size is moderate: Cohen's d of ~.5.	Gain is statistically significant and the effect size is large: Cohen's d of .8 or greater.

Figure 7. Reassess criterion of the learning improvement rubric.

> To demonstrate learning improvement, students must be assessed both before and after they have experienced changes that were made to pedagogy and curricula.

faculty members teaching four courses in the communications program agreed that it was a disservice to allow graduating students to earn a degree without being able to give an effective oral presentation. Furthermore, they decided to use the standards outlined in the learning improvement rubric as a way to guide and evaluate their learning improvement initiative. The example report is annotated with comments that explicitly link the standards included in the learning improvement rubric.

Conclusion: Where does learning improvement go from here?

Learning improvement has been central to the spirt of assessment for decades as evidenced by AAHE's principles of good practice. Twenty years later, there are signs that higher education is slowly pivoting toward demonstrating learning improvement. Calls for evidencing learning at institutions of higher education are being answered with meta-assessment rubrics; experts are vocalizing the need to do better; pilot programs and initiatives are surfacing in several national organizations and independent colleges and universities.

The standards we present via the learning improvement rubric and the example learning improvement report are attempts to advance the conversation of student learning. These resources highlight many of the components that must be in place for learning improvement to succeed.

We have a few suggestions to continue this discussion: some ideas target institutions of higher learning and others are meant for groups representing many states and regions.

These suggestions are not intended to be an ivory tower wish list. Instead, they are intended to be practical, actionable steps.

At an institutional level:

- Adopt assess, intervene, re-assess (weigh pig, feed pig, weigh pig) as the simple model for improvement;

- Think carefully about and provide resources for academic programs wanting to discuss learning improvement, such resources could include a modified learning improvement rubric and example report that meets internal needs;

- Pilot learning improvement initiatives – attempt to find one program that is ready for improvement (i.e., faculty agree on some SLO they want to improve and have buy-in) and try it out using the learning improvement rubric as a guide;

- After an initial success, specify how many programs should show improvement and make it a goal in the university's strategic plan.

For leading higher education groups at a state, regional, and national level:

- Expand many potentially successful initiatives by providing more specifics of what learning improvement is and looks like;

- Pull together resources for actual improvement. (NILOA has a good start but more examples are needed). Again, be more specific regarding advice to practitioners on the ground; advocate for reassessment.

- Give designations for great examples of programmatic learning improvement. Do something similar to what the Excellence in Assessment designation is doing but evaluate related to a more specific definition of learning improvement – as opposed to mere quality of assessment practice.

- Provide grants and awards to universities that show improvement at the program and university levels.

We note that this article is published in an assessment journal, RPA, which has extensive reach in assessment circles. While one of our authors (Meixner) is the Executive Director of our Center for Faculty Innovation, this piece is primarily written from the perspective of assessment experts attempting to gather support from fellow assessment experts. Nevertheless, the influence of non-assessment perspectives does not have to stay limited.

We hope other groups pick up and develop program-level learning improvement. Those in faculty development, for example, could provide much more commentary on enhancing pedagogy and curriculum. High-level administrators could provide suggestions for strategically organizing resources. Informal faculty leaders – the ones who make things happen within programs – could provide insight in developing internal support from colleagues. Precocious college students could also contribute. They are, after all, the ones whose learning we are trying to improve. We postulate that the movement's progress will accelerate markedly faster with collaboration from all of these groups and we need to strategize about how we get them to the table.

On a fun note – and bringing us full circle to the title's pig reference – Bellarmine University started what we hope becomes a trend in higher education. They liked the concept of improvement so much that they presented the ideas to their faculty and gave each of them a squeezable pig as a reminder of the weigh pig, feed pig, weigh pig model. We at JMU quickly followed suit; buying and distributing our own squeezable pigs. In fact, JMU and Bellarmine swapped pigs at the December 2015 SACSCOC conference as a gesture of porcine unity. Shortly thereafter a Berea College student leader brought the idea back to his school. Administrators promptly bought and distributed pigs.

Twenty years later, there are signs that higher education is slowly pivoting toward demonstrating learning improvement.

Of course, following the assessment cycle and distributing squeezable pigs won't magically foster student learning improvement. Pigs, however, do serve as a fun way to start the conversation by reminding us that learning doesn't improve by merely assessing it. If you would like to be part of this initiative, an easy first step is joining in on this discussion. Come look for us at an assessment conference near you. We'll happily swap pigs with you…and strategies for improvement!

AUTHOR'S NOTE:

As we mentioned, the bulk of the development of the learning improvement rubric was conducted as an independent study with students Smith and Sanchez based on past research on meta-assessment and applied work with learning improvement projects piloted at our home institution. After the independent study concluded and a draft of the rubric was created, Megan Good, Natasha Jankowski, Carol Hurney, and Meixner provided excellent suggestions, especially regarding rubric framing. We acknowledge and extend our gratitude to these colleagues for their contributions to this work.

References

Banta, T. W., & Blaich, C. F. (2011). Closing the assessment loop. *Change: The Magazine of Higher Learning*, *43*(1), 22-27.

Banta, T. W., Jones, E. A., & Black, K. E. (2009). *Designing effective assessment: Principles and profiles of good practice*. San Francisco, CA: Jossey-Bass.

Blaich, C. F., & Wise, K. S. (2011, January). *From gathering to using assessment results: Lessons from the Wabash National Study* (NILOA Occasional Paper No. 8). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Bresciani, M. J., Gardner, M. M., & Hickmott, J. (2009). *Demonstrating student success: A practical guide to outcomes-based assessment of learning and development in student affairs*. Sterling, VA: Stylus Publishing.

Chaplot, P., Booth, K., & Johnstone, R. (2010). *Building a culture of inquiry: Using a cycle of exploring research and data to improve student success.*

Erwin. T. D. (1991). *Assessing student learning and development: A guide to the principles, goals, and methods of determining college outcome*. San Francisco, CA: Jossey-Bass.

Fink, L. D. (2003). *Creating significant learning experiences: An integrated approach to designing college courses*. San Francisco, CA: Jossey-Bass.

Fisher, R., Smith, K. L., Finney, S. J., & Pinder, K. (2014). The importance of implementation fidelity data for evaluating program effectiveness. *About Campus*, *19*, 28-32.

Fulcher, K. H., Good, M. R., Coleman, C. M., & Smith, K. L. (2014, December). *A Simple Model for learning improvement: Weigh pig, feed pig, weigh pig*. (NILOA Occasional Paper No. 23). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Fulcher, K. H., Coleman, C. M., & Sundre, D. L., (2016). Twelve tips: Building high-quality assessment through peer review, *Assessment Update*, *28*(4), 1-2, 14-16.

Fulcher, K. H., Sundre, D. L., Russell, J. A., Good, M. R., Smith, K. L. (2015). Assessment progress report template (APT): Revised rubric draft.

Gerstner, J. J., & Finney, S. J. (2013). Measuring the implementation fidelity of student affairs programs: A critical component of the outcomes assessment cycle. *Research & Practice in Assessment*, *8*, 15-28.

Hutchings, P., Ewell, P. T., and Banta, T. W. (2012) *AAHE principles of good practice: Aging nicely*.

Hutchings, P., Jankowski, N. A., & Ewell, P. T. (2014, November). *Catalyzing assessment design activity on your campus: Lessons from NILOA's assignment library initiative*. Urbana, IL. National Institute for Learning Outcomes Assessment and Institute for Evidence-Based Changes.

Kuh, G. D., Ikenberry, S. O., Jankowski, N. A., Cain, T. R., Ewell, P. T., Hutchings, P., & Kinzie, J. (2015). *Using evidence of student learning to improve higher education*. San Francisco, CA: Jossey-Bass.

Kuh, G. D., Jankowski, N. A., Ikenberry, S. O, & Kinzie, J. (2014, January). *Knowing what students know and can do: The current state of student learning outcomes assessment in U.S. colleges and universities*.

Lumina Foundation. (2016). *The emerging learning system: Report on the recent convening and new directions for action*.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, 78(1), 33–84.

Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco, CA: Jossey-Bass.

Popham, W. J. (1997). What's wrong – and what's right – with rubrics. *Educational Leadership* 55(2), 72-75.

Shavelson, R. J. (2010). *Measuring college learning responsibly*: *Accountability in a new era*. Stanford, CA: Stanford University Press.

Smith, K. L., Good, M. R., Sanchez, E. H., & Fulcher, K. H. (2015). Communication is key: Unpacking "use of assessment results to improve student learning." *Research & Practice in Assessment*, *10*, 15-29.

Southern Association of Colleges and Schools Commission on Colleges. 2016. *General information on the reaffirmation process*.

Suskie, L. (2009). *Assessing student learning: A common sense guide* (2nd ed.). San Francisco, CA: Jossey-Bass Publishers.

Suskie, L. (2010). Why are we assessing? *Inside Higher Education*.

Walvoord, B. E. (2010). *Assessment clear and simple: A practical guide for institutions, departments, and general education* (2nd ed.). San Francisco, CA: Jossey-Bass.

Willard, T. (2015). *CHEA news release: Winners of the 2015 CHEA award for outstanding institutional practice in student learning outcomes announced*.

Appendix

Learning Improvement Report Annotated to Learning Improvement Rubric

Selecting the Targeted SLO

The Communications program endeavors to improve students' oral presentation skills; this goal is articulated through program Objective 4:

> *Students graduating from the BA program in Communications will (A) effectively deliver a presentation with an (B) engaging introduction, (C) logical and fluid body, and (D) smooth conclusion that reinforces the main ideas of the presentation.*

Why are these skills important? According to our alumni survey results, our students often pursue marketing jobs in which oral presentation skills are critical. Additionally, the *Journal of Effective Communications Education* cited oral presentation as the second most important skill for graduate students in the field.

Evidence to Support the Learning Improvement Initiative

While students are performing well on most objectives, they have struggled with oral presentation. Graduating students' skills have fallen below faculty standards in areas A, B, and D listed previously (i.e., effective delivery skills, introduction, and conclusion, respectively) of our oral presentation rubric for the last several years (see Table 1A). Additionally, students self-report their lowest gains in oral presentation (see Table 1B for a summary of these results).

Commented [A1]: Recall the **definition of learning improvement** as rationale for these sections and guiding framework for the learning improvement program presented here. The definition is provided again, here, so that the reader can see clearly the link between the definition, rubric, and example.

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."

Commented [A2]: Provided rationale for why program selected oral presentation skills as its learning improvement project. This corresponds to **Element/Standard B.1** of the Learning Improvement Rubric: *Program-level* SLO Selection: Rationale with Evidence.

Commented [A3]: Through collaboration with the institution's Assessment Center, the oral presentation rubric was found to be an effective measure at capturing Communications communication skills. It is important to ensure the measurement instrument is valid and reliable such that gains being seen, or not seen, are captured effectively by the instrument. This corresponds to **Element/Standard C.2** of the Learning Improvement Rubric: Psychometrics: Reliability and Validity.

Table 1A. *Objective 4: Oral Presentation Capstone Assessment Results of Three Cohorts, Oral presentation Rubric*

Oral presentation Rubric	2011 Results Mean	2012 Results Mean	2013 Results Mean (SD) Desired Mean = 3	Score Difference*** 2012 – 2013
A. Delivery Skills	*2.8*	*2.5*	*2.6 (.42)**	No
B. Introduction	*2.7*	*2.9*	*2.8 (.55)**	No
C. Body	*3.1***	*2.9*	*3.0 (.38)***	No
D. Conclusion	*2.9*	*2.7*	*2.7 (.49)**	No

Note. Oral presentation Rubric (n = 25): 1 = unsatisfactory, 2 = emerging, **3 = competent, 4 = highly competent*
*Orange coding indicates the degree to which the observed results were worse than desired.
**Blue coding indicates the degree to which the observed results were better than the desired result.
***Based on independent t-tests, using p < .01 as significance level (lower alpha due to multiple comparisons).

Table 1B. *Objective 4: Oral Presentation Self-Report, Graduation Survey*

Graduation Survey	2011 Results Mean	2012 Results Mean	2013 Results Mean (SD) Desired Mean = 3	Score Difference*** 2012 – 2013
Oral Comm. Skills	2.7	2.6	*2.6 (.8)**	No

Note. Graduation Survey (n = 91): 1 = no gain, 2 = small gain, **3 = moderate gain, 4 = large gain...*
*Orange coding indicates the degree to which the observed results were worse than desired.
***Based on independent t-tests, using p < .01 as significance level (lower alpha due to multiple comparisons).

Explaining the Course Modification Process: Investigating Oral presentation Skill Development

In order to begin changing the way we provide oral presentation education, Communications faculty first investigated how and where we taught these skills across the program curriculum. To do so, we looked at faculty-submitted syllabi and schedules. The curriculum map (see Table 2) lists our required courses along with the degree to which each of our program objectives were *theoretically* covered before this learning improvement project. The

> **Commented [A4]:** Here, we see a collection of baseline data. This corresponds to **Element/Standard C.1** of the Learning Improvement Rubric: Baseline Data Collection: Timing and Sampling.

> **Commented [A5]:** This section details how the program investigated the targeted objective BEFORE the modifications were designed and implemented. The purpose of this section is that, too often, programs jump to new modifications without an understanding of current efforts. This corresponds to **Element/Standard D** of the Learning Improvement Rubric: Investigate Curriculum and Diagnose Issues.

more instruction time allotted to developing oral presentation skills, the higher the degree of coverage. Oral presentation skills correspond to program Objective 4.

Table 2. *Curriculum Map of Communications Program (Obj. 4 is Oral presentation)*

Course/Learning Experiences	Obj. 1	Obj. 2	Obj. 3	**Obj. 4** **Oral presentation**
COMM201 (Introduction to Communication Theories)	3	0	1	**0**
COMM301 (Research Methods in Communication Studies)	3	0	1	**2**
COMM302 (Rhetorical and Scientific Communication)	3	0	1	**0**
COMM303 (Public Relations)	3	0	0	**2**
COMM304 (News and Journalism)	3	1	1	**0**
COMM361(Interpersonal and Organizational Communication)	0	3	1	**0**
COMM401 (Digital Media and Social Communication)	1	1	3	**0**
COMM402 (Presentational Speaking)	1	0	1	**3**
COMM403 (Policy and Campaign Communication)	2	0	0	**0**
COMM404(Marketing and Corporate Communication)	0	2	3	**0**
COMM480 (Capstone)	0	2	2	**2**

Note. Syllabi coverage of Objective 4: 0 = no coverage; 1 = slight coverage; 2 = moderate coverage; 3 = major coverage

As is shown in the curriculum map, four course syllabi addressed the oral presentation learning objective (Objective 4); these courses have been shaded purple in Table 2. In three courses, oral presentation was *theoretically* covered at a moderate level (e.g., a few assignments related to oral presentation skills). One course *theoretically* covered oral presentation at a major

level (e.g., multiple weeks were dedicated to developing oral presentation skills). On paper, it would seem, students should have ample opportunity to learn how to deliver an effective presentation.

Nevertheless, evidence collected from our oral presentation rubric during the capstone course assessment clearly indicated that students are not as proficient in oral presentation as program faculty expect (e.g., an average of 3 on the communication rubric). Low self-reported gains were also cause for us to change how and to what degree oral presentation skills were taught.

The six faculty members who teach sections in courses with oral presentation objectives (i.e., COMM 301, 303, 402, and 480) met in hopes to discover, as a program, why student scores were falling short of meeting expectations. The meetings were facilitated by our program coordinator and were held three times in the month of March. What follows is a summary of our discussions:

> **Commented [A6]:** This section corresponds to **Element/Standard A** of the Learning Improvement Rubric: Faculty Involvement.

- Indeed, students did verbally present in each of the four courses that had oral presentation course objectives. However, some faculty noticed that students were not taking the assignments seriously. Several faculty members took an informal poll of students in the days after the capstone presentation. Very few students raised their hand when the professor asked if they had practiced the entire presentation at least twice.

- Although the oral presentation rubric was used for COMM480, the capstone course, professors teaching other courses were unaware of the rubric's existence. Many said that the rubric may be helpful in guiding feedback to give to students in their classes.

- Across the four courses, oral presentation assignments varied greatly by instructor and section. More often than not, oral presentation assignments and evaluations were more heavily weighted toward course content rather than developing communication skills. One professor characterized this trend as follows:

  If the presentation was reasonably accurate, the student received an "A," despite lackluster oral presentation skills. I would make comments on the feedback sheet, 'seemed a bit nervous, spoke too quickly…,' but that was about it. I provided more specific feedback regarding the accuracy of the presenter's content. Nevertheless, the presentation quality was far, far away from what would be considered professional or polished.

- Several of the faculty revealed that they did not feel comfortable providing feedback on students' oral presentation skills. Although we acknowledged the importance and necessity of the objective, we had received little or no training regarding how to provide effective feedback.

Course Modifications: Learning Interventions

After meeting with the program coordinator, a course modification plan was devised and supported by **all six faculty** who teach program courses with an oral presentation component. Given that all four courses are required for all majors, **100% of students in the program will be affected**. Note that **each of the four** courses were modified to some degree; however, the most extensive modifications were in COMM402: Presentational Speaking and COMM480: Capstone.

Commented [A7]: This section corresponds to **Element/Standard D** of the Learning Improvement Rubric: Hypotheses for why what was originally being taught in the curriculum relative to the targeted SLO was ineffective.

Commented [A8]: Note that ALL SIX of the faculty who teach the oral presentation courses in the program were part of the project. Program-level modifications require coordination across faculty in courses targeted for the modifications. This corresponds to **Element/Standard A** of the Learning Improvement Rubric: Relevant Faculty Involvement.

Commented [A9]: States what percent of program students are affected. Programmatic modifications must affect a large percentage of students. This corresponds to **Element/Standard E2** of the Learning Improvement Rubric: Intervene.

What follows is a short description of each course modification (i.e., learning intervention). An overview of these modifications can be found in Table 3.

Modification 1: Explaining Importance and Clarifying Expectations. One of the first required courses in the major is COMM402 (Presentational Speaking). In this class, students present their final project at the end of the semester. The three faculty who teach COMM402 will explain both the *importance of oral presentation* and the *expectations of program faculty*. Instructors will communicate that this has, in general, been an area of weakness for graduates; furthermore, likely employers and graduate schools desire students who are competent of such skills. Both faculty and students will need to work hard to develop oral presentation skills. Students will watch videos of the three best capstone presentations from previous years. Faculty will describe to their students how each of the presentations were evaluated on the oral presentation rubric.

Modification 2: Using the Oral Presentation Rubric. Oral presentations in each course (COMM 301, 303, 402, and 480) will be evaluated on content (70%) and oral presentation skills (30%). Each faculty member will use the oral presentation rubric to score the 30% of the students' presentations.

Modification 3: Emphasizing Practice. In all courses with an oral presentation component, faculty will urge students to practice their presentations *at least* four times before the in-class performance. Every student will be given a copy of the rubric, provided time in class to practice with other students, and encouraged to tape and review their practice efforts.
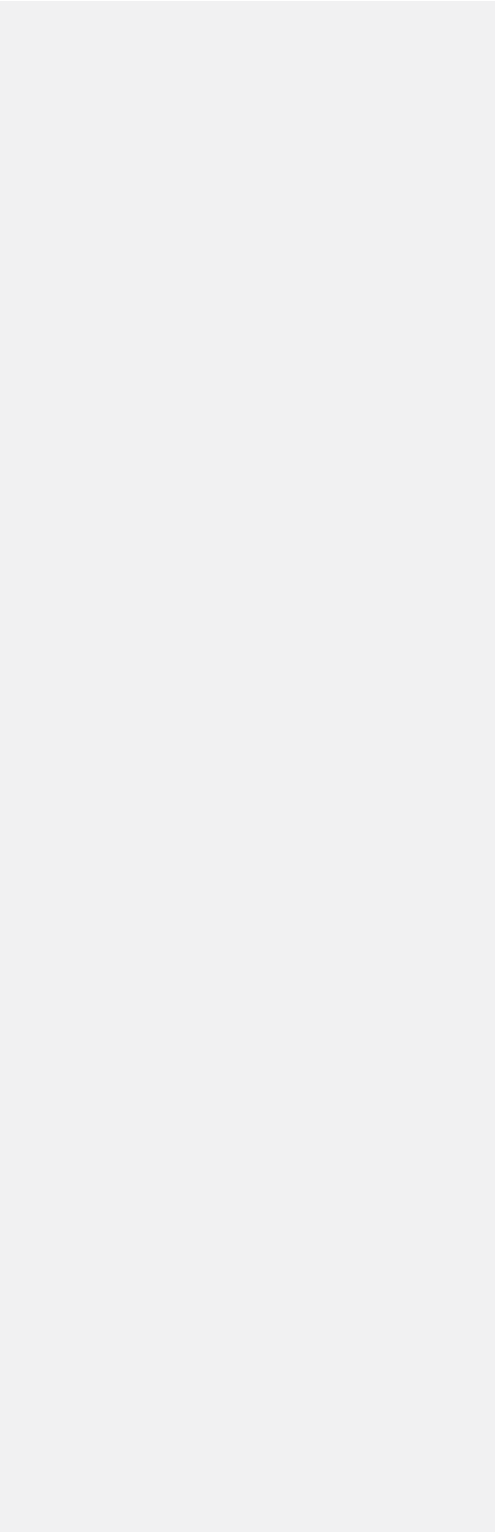
Modification 4: Upping the Stakes of Capstone Presentations. For the capstone, the ante will be raised. The final oral presentation will be open to all program faculty and to all

> **Commented [A10]:** Explain program-level modifications in detail. This corresponds to **Element/Standard E1.2** of the Learning Improvement Rubric: Program-Level Intervention.

majors; it will also be recorded. The three capstone professors will convey to students that this

presentation demonstrates not only what students have learned in the program but also how well-

prepared they are for jobs or graduate school.

Note, we (the six faculty members) collectively spent five days in a workshop prior to the

first week of Fall classes. The campus Center for Faculty Development, Teaching, and

Innovation facilitated discussion and strategy of how to incorporate the listed modifications into

our Communications courses.

Table 3. *Curriculum Map and Modifications*

Course/Learning Experiences	Modifications tied to Course/Learning Experiences
COMM301 (Research Methods in Communication Studies)	• Faculty will explain the *importance of oral presentation* and the *expectations of program faculty*. • Instructors will communicate that this has, in general, been an area of weakness for graduates and employers and graduate schools desire students who are competent of such skills. • Students will watch videos of the three best capstone research presentations from previous years. • Faculty will describe to their students how each of the research presentations were evaluated on the oral presentation rubric and present students with the rubric to clearly articulate the expectations. • Oral presentations of final research projects will be evaluated on content (70%) and oral presentation skills (30%) & the oral presentation rubric will be used to score the 30% of the students' presentations. • Faculty will urge students to practice their research presentations *at least* four times before their final presentation to the class.
COMM303 (Public Relations)	• Public relations presentations will be given orally and evaluated on content (70%) and oral presentation skills (30%) & the oral presentation rubric will be used to score the 30% of the students' presentations. • Faculty will urge students to practice their public relations oral presentations *at least* four times.
COMM402 (Presentational Speaking)	• Oral presentations will be evaluated on content (70%) and oral presentation skills (30%) & the oral presentation rubric will be used to score the 30% of the students' presentations.

	• Faculty will urge students to practice their presentations *at least* four times.
COMM480 (Capstone)	• Oral presentations of the Capstone project will be evaluated on content (70%) and oral presentation skills (30%) & the oral presentation rubric will be used to score the 30% of the students' presentations.
	• Faculty will urge students to practice their presentations *at least* four times.
	• The final oral presentation of the Capstone project will be open to all program faculty and to all majors; it will also be recorded. The three capstone professors will convey to students that this presentation demonstrates not only what students have learned in the program but also how well-prepared they are for jobs or graduate school.

Because modifications 1-4 take place in several courses that span multiple semesters of the program, the total effect of the course modifications was not realized or evidenced/captured for several years. In order to provide this final report, we gradually modified courses and evaluated students. In 2014, we evaluated students giving their capstone presentations who had not taken any courses with new oral presentation assignments or instruction. This time point serves as our baseline data point.

In 2015, we again evaluated students giving their capstone presentations using the oral presentation rubric. Because course modifications were made to two courses during both semesters of the 2014-2015 year, students had taken some courses with new oral presentation assignments and instruction. We consider this a "partial modification" time point.

By the time students gave their capstone presentations in 2016, all four courses that we planned to modify were indeed changed. Students graduating in 2016 and 2017, having taken all four modified courses, were evaluated using the oral presentation rubric. This is considered a "full modification." See Table 4 for details.

Table 4. *Planned Course Modification and Data Collection Sequencing for Oral presentation in the Communications Program*

Planned Course Modifications	2014	2015**			2016**			2017**		
COMM301 (Research Methods in Communication Studies)	0	0			**1**	**2**	**3**	1	2	3
COMM303 (Public Relations)	0	0			**2**		**3**	2		3
COMM402 (Presentational Speaking)	0	**2**		**3**	2		3	2		3
COMM480 (Capstone)*	0	**2**	**3**	**4**	2	3	4	2	3	4

	This year serves as our baseline data.	This year serves as our partial modification (intervention) data.	This year serves as our full modification (intervention) data.	This year serves as another full modification (intervention) data.
Intervention Implementation Status	Students graduating in 2014 took **no courses** with oral presentation modifications.	Students graduating in 2015 took **two courses** with oral presentation modifications.	Students graduating in 2016 took all **four courses** with oral presentation modifications.	Students graduating in 2017 took all **four courses** with oral presentation modifications. Also, had added benefit of individual instruction & assignment tweaks.

Note. Data collected during COMM480; the course is only offered during the Spring semester. **Faculty received training on rubric use before Fall semesters. 0 = No modification; 1 = Explaining importance and clarifying expectations; 2= Using the oral presentation rubric; 3= Emphasizing Practice; 4= Upping the stakes of capstone presentations

Conclusion (2016). Clear evidence, provided by scores from a rubric used to evaluate student presentations in a capstone course, suggested that multiple cohorts of graduates were, in fact, failing to meet faculty expectations. Responses from graduation surveys reinforced the need for better oral presentation education. To determine which courses could be modified, and how, to help students learn, the six faculty in this example met several times. Discussions were insightful and illuminating: some faculty commented that few students took the capstone presentation seriously, many discussed how few opportunities students had to practice their speeches and receive feedback, several faculty teaching lower level courses were unaware of the oral presentation rubric used in the capstone course, and still others noted that they had no training providing feedback on student presentations. A set of course modifications emerged through these meetings. The Communications program did not make changes to all four courses with oral presentation objectives right away. Instead, the program faculty developed a learning improvement timeline. The following example Learning Improvement Report is a complete report that documents *four years* of gradual course modifications and improving student presentation scores on the oral presentation rubric.

Evidence in this report clearly shows (i.e, statistical significance and a large effect size) that the students in this program have improved in giving an oral presentation. Figure 1 illustrates these improvements. Because the cohorts have remained relatively similar throughout the past decade, the Communications faculty can attribute this improvement to the four-course modification student learning improvement initiative. Validity and reliability evidence is provided to support these claims.
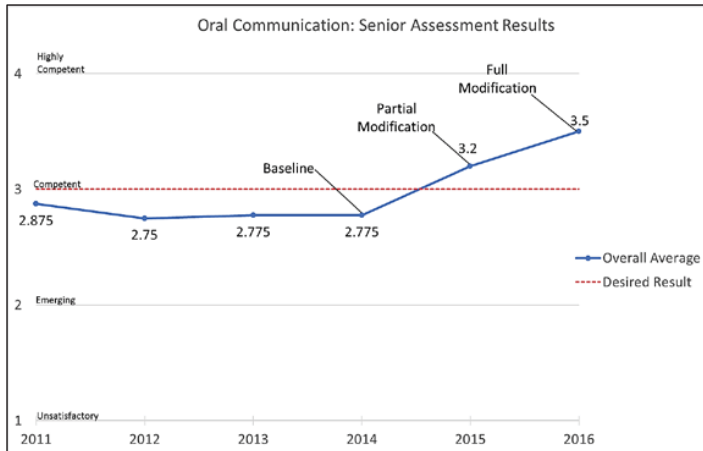
Figure 1. *Objective 4: Oral presentation Capstone Assessment Results of Three Cohorts, Oral presentation Rubric Average Scores*

Conclusion (Updated 2017). After a year of planning and two years of modifying our courses, we are happy to say that students are learning oral presentation skills and meeting faculty standards. As is shown in Table 5, students' oral presentation proficiency in areas A, B, and D of the program objective and oral presentation rubric improved from somewhat below competent (a 3 on the rubric) to midway between 3 and 4 (highly competent): a statistically significant change of over 1 standard deviation (SD) gain (a large effect). See Tables 5A and 5B.

Table 5A. *Objective 4 Update: Oral Presentation Capstone Results of Three Cohorts*

Oral presentation Rubric	2014 Results Mean	2015 Results Mean	2016 Results Mean (SD) Desired Mean = 3	Score Difference*** 2014 – 2016
A. Delivery Skills	2.6 (.42)*	3.1	3.5	Yes
B. Introduction	2.8 (.55)*	3.2	3.4	Yes
C. Body	3.0 (.38)**	3.2	3.5	Yes

Commented [A12]: Magnitude of learning improvement provided: The gain is statistically significant and the effect size is at least "moderate" Cohen's D is > 0.5. This corresponds to **Element/Standard F** of the Learning Improvement Rubric: Impact of the Intervention.

Commented [A13]: Direct measure used to evaluate improvement. This corresponds to **Element/Standard F.1**. of the Learning Improvement Rubric: Impact of the Intervention.

Commented [A14]: Note that the assessment results bookend the modifications. Assess – intervene – re-assess. This corresponds to **Element/Standard F** of the Learning Improvement Rubric: Impact of the Intervention

D. Conclusion	2.7 (.49)*	3.3	3.6	Yes

*Oral presentation Rubric (n = 25): 1 = unsatisfactory, 2 = emerging, **3 = competent**, 4 = highly competent*
*Orange coding indicates the degree to which the observed results were worse than desired.
**Blue coding indicates the degree to which the observed results were better than the desired result.
***Based on independent t-tests, using p < .01 as significance level (lower alpha due to multiple comparisons).

Table 5B. *Objective 4 Update: Oral presentation Self-Report, Graduation Survey*

Graduation Survey	2014 Results Mean	2015 Results Mean	2016 Results Mean (SD) Desired Mean = 3	Score Difference*** 2014 – 2016
Oral Comm. Skills	2.6 (.8)*	3.3	4.2	Yes

Note. Graduation Survey (n = 91): 1 = no gain, 2 = small gain, **3 = moderate gain, 4 = large gain, 5=tremendous gain*
*Orange coding indicates the degree to which the observed results were worse than desired.
**Blue coding indicates the degree to which the observed results were better than the desired result.
*** Based on independent t-tests, using p < .01 as significance level (lower alpha due to multiple comparisons).

Please note: In the years before our learning improvement initiative, oral presentation scores on the program's rubric had remained relatively low (e.g., below the desired result of 3, *Competent*). After a thorough investigation, there are no indicators that more recently graduated cohorts of students would be naturally better at giving oral presentations. Our incoming and graduating student cohorts have relatively similar academic demographics (e.g., SAT score averages across the cohort are roughly equivalent). For this reason, we can say that our students improved their oral presentation abilities because they received better instruction, practice, and feedback through a program-level curricular modification.

We took several additional steps in order to ensure that the results documented in this report can be trusted and that learning improvement gains can be linked to the program-level curricular modification:

Commented [A15]: Note that the program summarizes validity evidence of their methodology showing that it is robust. This corresponds to **Element/Standard C.2** of the Learning Improvement Rubric: Psychometrics: Reliability and Validity.

- We carefully selected the oral presentation rubric relative to the oral presentation program objective (content validity);
- we kept the same rubric throughout the entire learning improvement project;
- before non-capstone instructors used the rubric, we had training sessions;
- the Phi Coefficient, an indicator of reliability, ranged from .61 - .78 over the years reported (this is an acceptable range for performance assessment);
- over time, the rubric scores correlated with survey scores regarding oral presentation improvement, providing some concurrent validity evidence; and
- more detail regarding the methodology can be found in the program's assessment report (APT).