Abstract

This article describes a faculty-led project to assess and revise institutional student learning outcomes at a small urban community college. The revision process involved four stages: (1) exploring stakeholders' explicit and implicit understandings through an experimental assessment; (2) using statistical tools to identify redundancies and opportunities for regrouping and revising the learning outcomes; (3) triangulating findings through focus group discussions and test assessments; and (4) drafting and refining the revised learning outcomes. By grounding revisions in stakeholders' explicit and implicit understandings of the existing outcomes, the school was able to streamline and significantly improve institutional student learning outcomes without starting completely from scratch.

AUTHORS

Forest Fisher, Ph.D.
City University
of New York

Tara Bahl, Ph.D.
City University
of New York

Nate Mickleson, Ph.D.
New York University

An Intentional Process for Revising Institutional Learning Outcomes

The AAC&U's (2009) VALUE Rubrics and the Lumina Foundation's (2014) Degree Qualifications Profile offer crucial frameworks for defining learning outcomes and using them at the assignment, course, program, and institutional levels. There is growing literature on the validity, affordances, and limitations of these frameworks (e.g., Colson et al., 2018; Stevenson et al., 2016). Despite increasing adoption of these sophisticated tools, it remains a challenge to ensure that assessment work is meaningful to faculty, staff, and students. Indeed, faculty and other stakeholders often experience learning outcomes assessment as an exercise in institutional box-checking that is irrelevant or even detrimental to their work with students (Stanny, 2018). As Schoepp & Tezcan-Unal (2017) have found, misperceptions about the purposes and uses of outcomes assessment can inhibit participation and limit an institution's ability to use assessment work to improve student outcomes. As Colson et al. (2018) have noted, faculty are also less likely to embrace and use learning outcomes frameworks they perceive to be unnecessarily complex.

This article describes a faculty-led project to assess and revise institutional student learning outcomes (ISLOs) at a small urban community college. We launched the project in spring 2019 and completed it in spring 2020 when the revised ISLOs were approved through the college governance. As we describe below, it proceeded in three general stages: gathering data and perspectives about the existing ISLOs from the college community; analyzing data, triangulating findings, and drafting revised ISLOs; and refining the revised ISLOs through consultation with the college community. The project began as a result of limitations we discovered through several cycles of assessment. These limitations motivated us to undertake the project we describe here to assess and revise our ISLOs. The

CORRESPONDENCE

Email
forest.fisher@guttman.cuny.edu

growing literature on outcomes assessment suggests several principles for ensuring the work is meaningful and effective:

- Develop clear outcomes that faculty, students, and other stakeholders perceive as directly aligned with curricula;
- Design collaborative processes that build from the bottom up, or from the classroom and co-curricular experience to the program and institutional levels;
- and Use evidence gathered through assessment to continually revise and refine outcomes and processes.

Effective assessment is not about checking boxes but about gathering evidence and using it to continually improve student learning outcomes.

The literature shows that distributions of labor and accountability within assessment processes play a key role in their effectiveness (Kinzie & Jankowski, 2015). Matuga & Turos (2018), for example, have found that misalignments can feed faculty disengagement and distrust. Engaging faculty insights and leadership in designing and enacting learning outcomes assessment is a key lever for transforming a "culture of assessment," with its bureaucratic overtones, into a "culture of improvement" that enables faculty, staff, administrators, and students alike to gather, reflect, and act on evidence of learning-in-progress (Stanny, 2018, p. 114; see also Roscoe, 2017).

We add to this literature in three ways. First, we explain how we used evidence from prior assessments as a foundation for improving our learning outcomes rather than starting from scratch. Second, we describe several methods we used to engage faculty and other stakeholders to build from the bottom up rather than from the top down. Third, we highlight the importance of taking time to examine and reflect on explicit and implicit understandings of existing ISLOs in order to ensure they are clear, concise, and aligned with curricula.

Context/Background

Outcomes

Our original ISLOs identified 24 skills in five categories: Broad Integrative Knowledge; Applied Learning; Specialized Knowledge; Intellectual Skills for Lifelong Learning; and Civic Engagement. The ISLOs' initial purpose was to provide a framework for us to follow in developing the college's curriculum. The Broad Integrative Knowledge and Intellectual Skills for Lifelong Learning outcomes informed the first-year experience and general education requirements and the Specialized Knowledge outcomes defined fundamental skills for the college's degree programs. The Applied Learning and Civic Engagement outcomes were intended to infuse the whole curriculum by orienting the work students might do toward engagement with the surrounding communities and preparation for emerging careers. The ISLOs were developed by a team of administrators and faculty through an iterative design process that drew from two national models: AAC&U's Essential Learning Outcomes and associated VALUE rubrics and The Lumina's Foundation's Degree Qualifications Profile (DQP). Each ISLO category has a corresponding rubric. Figure 1 below shows an excerpt from the Broad Integrative Knowledge rubric, an example we will revisit throughout this article.

Our framework shows the contrasting influences of these models. Some of the ISLOs correspond to stages in a student's progress toward their degree, similar to the schema outlined in the DQP. Other ISLOs describe broader areas of learning which mirror the approach embedded in AAC&U's Essential Learning Outcomes and VALUE rubrics. The project we describe in this article helped us disentangle these elements and develop revised ISLOs that are clear, concise, and aligned with our evolving curriculum.

It is quite common for schools to do what we did, adapting language from the DQP and the AAC&U's VALUE rubrics without raising questions about their construct validity. As Knekta et al. (2019) warns, "validity must be considered each time an instrument is used" since it may be valid for one population and purpose but not another (pp. 2). The VALUE rubrics were designed with a general population in mind and could not anticipate the many different, specific student populations with which they would be used. Even though we adapted these

Figure 1
Excerpt from Original Broad Integrative Knowledge ISLO

Broad, Integrative Knowledge: General Education *The outcomes in this category demonstrate that students can integrate learning from broad fields of general study and connect different academic disciplines and multiple perspectives.*				
Criteria or Domain	**Capstone 4**	**Milestones 3**	**Milestones 2**	**Benchmark 1**
a. Engages with issues that have contemporary, historical, scientific, economic, technological, or artistic significance	Applies new knowledge on an issue to academic and/or experiential contexts. Independently evaluates information from multiple sources. Can articulate multiple perspectives on an issue to others.	Situates an issue in a broader context to provide in-depth explanation. Independently gathers information from multiple sources. Can articulate own position on an issue.	Explores an issue with some depth by applying skills or presenting evidence provided in classes. Provides occasional insight and/or connection to self.	Explores issues at surface level, providing little insight and/or information beyond the basic facts. Can state ideas from other sources.
b. Exhibits an understanding of how different disciplines create knowledge and approach questions.	Synthesizes knowledge and approaches from at least two disciplines in planning and conducting research geared toward answering questions.	Considers that different disciplines ask and answer questions in different ways. Presents a rationale for following one disciplinary approach over another in specific cases.	Recognizes knowledge in a specific discipline. Asks and answers questions using general assumptions and approaches of one's own discipline.	Lists academic disciplines and expresses interest in one or more subject areas.
c. Evaluates multiple perspectives on key issues connected to societal concerns.	Synthesizes multiple perspectives through comprehensive evidence-based analysis of positions.	Analyzes multiple perspectives on a key issue connected to societal concerns. Provides some evidence to support an argument.	Acknowledges two sides of a key issue connected to societal concerns. Describes both perspectives by clarifying each position.	States a single perspective on a key issue connected to societal concerns with basic description.

The development of clear and concise ISLOs is crucial for any institution's curriculum development.

rubrics to our local context in designing our initial ISLOs, our assessments showed that some outcomes were less relevant than others to our students' experiences and the student work we assessed. As a result, some faculty members' "mental models" of assessment defined our institution-level work as being separate and distinct from assessments they use in their classrooms (Heinrich, 2017).

Structures

Our institution includes two structures designed to make the assessment process more collaborative: (1) a faculty-led Academic Assessment & Learning Committee charged with assessing student learning and recommending improvements and (2) dedicated Assessment Days at the beginning, middle, and end of the semester when no classes are held. Assessment Days are organized by the Assessment & Learning Committee and funded by the Office of Academic Affairs. These days provide time and space (and lunch!) for faculty, staff, and administrators to collaboratively assess student work, discuss and reflect on emerging evidence, and engage in curricular and professional development activities. Participation rates have been consistently high: a majority of the full-time faculty participate regularly along with a considerable number of staff from the Office of Student Engagement and other units. The Assessment & Learning Committee works in collaboration with two deans: one located in the Office of Academic Affairs and reporting to the Provost and the other located in the Office of Institutional Effectiveness and Strategic Planning and reporting to the President.

Assessment Plan

The college's assessment plan charges working groups co-chaired by two elected members of the Assessment & Learning Committee to assess each of the five ISLOs on staggered two-year cycles. Participants at the Assessment Days rate student work from first-

semester, first-year, and capstone courses for evidence of learning particular to each ISLO. The working groups then analyze, reflect, and report on the evidence and offer recommendations. Even with the working group structure and staggered assessment plan, assessing 24 skills across five ISLOs has proven difficult. When we looked across working group reports, a number of limitations arose repeatedly:

1. **A lack of consistency between levels in the rubrics.** For example, the benchmark (lowest level) for skill B on the Broad Integrative Knowledge rubric (see figure 1 above) states that a student "lists academic disciplines and expresses interest in one more subject area" while milestone 2 requires that a student "ask and answer questions using the general assumptions and approaches of one's own discipline." If a student asks and answers questions in their own discipline but does not list any other disciplines, should they receive a 1 or a 2 for this skill?
2. **Many outcomes appeared to measure similar skills.** For instance, on the Broad Integrative Knowledge rubric, the capstone (highest level) for both skills A and C is about synthesizing or engaging multiple perspectives.
3. **Some skills did not reflect the type of learning that was intended for the classroom.** For example, skill B on the Broad Integrative Knowledge rubric focuses on integrating different academic disciplines. Faculty wondered if this discipline-heavy language reflected the kind of integrated learning we envisioned for first-semester college students and instead suggested that we focus on integrating "perspectives" or "methods of inquiry."
4. **Difficulty finding examples of student work that were appropriate to assess with these rubrics.** One of the skills addressed by the ISLOs was "collaboration", and faculty found it difficult to assess collaboration through the end product of group work. "No Evidence" ratings ranged from 10% to 80% in our working groups' assessments.

As a result of these challenges, the working groups consistently observed **low inter-rater reliability** in their assessment. For example, when evaluating skill B on the Broad Integrative Knowledge rubric, only 43% of participants agreed in their ratings. We measure agreement when raters' scores are within 1 or when they agree that the work offers no assessable evidence of a given skill.

The Assessment & Learning Committee considered rewriting the outcomes from scratch to resolve these limitations. However, since our assessment work had already taught us quite a bit about the ISLOs, we determined that starting from scratch would mean introducing a whole new set of unknown issues. Our goal was to fix known issues, not introduce new unknown issues. Therefore, the committee decided to take the approach of identifying which skills from the original ISLOs might be eliminated, which might be maintained, and which might be combined. We designed an assessment project with four stages: (1) exploring stakeholders' explicit/implicit understandings through an experimental assessment; (2) using statistical tools to identify redundancies and opportunities for regrouping and revising the ISLOs; (3) triangulating findings through focus group discussions and test assessments; and (4) drafting and refining the revised ISLOs.

Even with the working group structure and staggered assessment plan, assessing 24 skills across five ISLOs has proven difficult.

Explicit/Implicit Understandings of ISLOs (Spring 2019 - Summer 2019)

In this section, we describe the process we used to revise the ISLOs and explain how different steps in that process explored stakeholder's implicit and explicit understandings of the ISLOs. As a first step, we collected two different types of data about stakeholder understandings of the ISLOs. The first type of data measured faculty and staff's *explicit understandings* of the learning outcomes in relation to one another. We revised the language of the learning outcomes and their rubrics following the recommendations of previous working groups. We then printed each outcome on a separate piece of paper and gave copies of these outcomes to groups of faculty and staff at one of the Assessment Days. We asked each group to discuss the outcomes and reorganize them in the way that they felt made the most sense.

This process produced nine different potential ways of reorganizing the learning outcomes. For instance, group 1 organized the outcomes into four categories: Communication, Cultural Background and Identity, Problem Posing, and Knowledge of the Field or Program of Study while group 5 used five: Critical Thinking & Practice or Applied Learning, Research Process, Disciplinary Fluency, Self-Aware Learning, Civic and Community Engagement. We looked for commonalities across groupings that would allow us to better understand how faculty and staff explicitly envisioned the learning outcomes in relation to one another. For instance, both groups 1 and 5 put skills 3B ("Connections to Experience") and 4E ("Cultural Background and Identity") together, group 1 in the category Cultural Background and Identity, and group 5 in the category Critical Thinking & Practice or Applied Learning. This suggested an overlap in what our assessment of these skills might measure.

We found a conflict between faculty and staff's explicit and implicit understandings of ISLOs, highlighting potential inaccuracies in assessment practices.

We also surveyed a small group of faculty to understand better their experiences using the rubrics. The survey asked participants three questions: "What was clear, effective, or useful about using the rubric?", "What was confusing, ineffective, or difficult about using the rubric?", and "In what ways would you recommend revising this rubric?" Some survey results corroborated findings from other types of data. For instance, one survey respondent observed that skill "1C [on collaboration] was difficult [to assess] because the assignments that I read did not specify group or individual work." Likewise, another respondent observed that skill 4B ("Synthesize Multiple Perspectives") was "confusing because of the term 'discipline-specific issues', which only shows up in Level 2." Overall, respondents noted "the progression [of skills] didn't seem logical."

The results also suggested that faculty members' *explicit* descriptions of their understandings of ISLOs may differ from how they use them in practice. For example, a faculty member may say that two outcomes are related, but in practice, they may actually score these two outcomes very differently. This suggests a conflict between explicit and implicit understandings of the outcomes. To identify these potential inaccuracies, we decided to collect a second type of data that would help us assess *implicit understandings* of the learning outcomes. We asked faculty and staff to assess student work using the rubrics and then used a statistical technique called Exploratory Factor Analysis to identify groups of outcomes that tended to receive similar ratings.

To do this, we selected 80 samples of student work from recent course and program-level assessments: 40 from courses in the first-year core curriculum and 40 from courses in the programs of study. Half of the samples in each selection pool had received lower ratings in previous assessments (an average rating of 2 or below on a scale of 1-4) and half had received higher ratings (an average above 2 on a scale of 1-4). Figure 2 provides more detailed information about the sample.

Figure 2
Sample for Re-Assessment

Course	Total sample	Student work with average ratings 1-2	Student work with average ratings above 2
First-year Social Science Course	20	10	10
Interdisciplinary Freshman Seminar	20	10	10
Business Administration Capstone	10	5	5
Human Services Capstone	10	5	5
Liberal Arts & Sciences Capstone	10	5	5
Urban Studies Capstone	10	5	5
Total	**80**	**40**	**40**

To re-assess these pieces of student work, we designed test rubrics that divided the 24 skills listed in our ISLOs across four rubrics containing six skills each. The groupings did not align with our existing categories. In fact, we made an explicit effort to place similar or related

skills on different test rubrics because we worried that faculty might give side-by-side skills similar ratings just because they appeared superficially similar. For example, we placed the skills "Quantitative Data Analysis" and "Quantitative Problem Solving" on different test rubrics because we worried faculty might rate them similarly just because they both have the word "quantitative" in their titles.

Eight faculty were paid a small stipend through an internal grant to use the test rubrics to assess student work over two days during the summer. On the first day, each participant assessed roughly 20 pieces of student work using one of the four test rubrics. On the second day, they used a different test rubric to assess roughly 20 more pieces of student work. In this way, each individual participant was responsible for assessing student work for 12 of the 24 ISLO skills. In total, we assessed 72 pieces of student work for all 24 skills. While we could have generated more robust data by asking each participant to assess each piece of student work for all 24 skills, we determined this would have been too cognitively taxing.

[The results suggest] a conflict between explicit and implicit understandings of the outcomes.

Exploratory Factor Analysis (Fall 2019)

Next, we used Exploratory Factor Analysis (EFA) to analyze this data. EFA is a statistical technique used to extract underlying latent variables that might characterize a data set (Tabachnick et al., 2007). Social scientists frequently use it to measure the construct validity of different statistical instruments like a survey (Knetka et al., 2019). The core idea is to look for correlations in the responses to different survey questions. A researcher might find that respondents tend to answer three questions similarly suggesting that there is one underlying factor explaining the answers to all three of these questions. In this way, EFA allows researchers to identify a smaller number of factors that might explain their survey responses and then determine whether or not these factors align with constructs they are studying. As explained below, we decided that descriptive statistics were more appropriate for our small data set, but we have included a short discussion of EFA here because it offers an innovative approach that other institutions with larger, more robust data sets might want to consider.

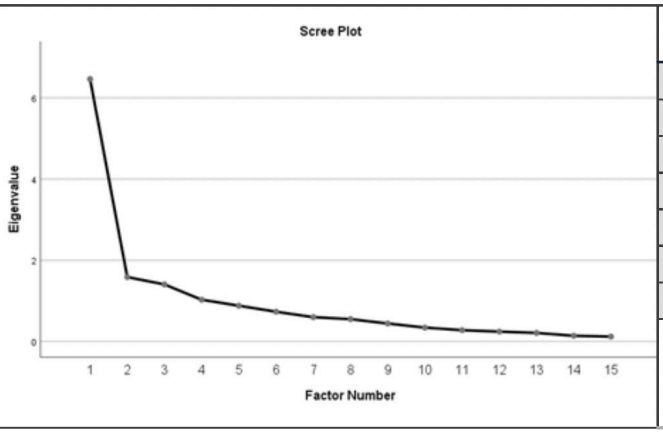
The EFA results are shown in figures 3 and 4 below. Figure 3 shows the Scree Plot and eigenvalues of the factor analysis. These measures are used to determine the appropriate number of factors to extract. A best practice is to consider only factors with an eigenvalue of 1 or higher (4 factors in our case); but since our data are messy and our sample size small, we opted to consider six factors. The Scree Plot provides a way of visualizing this information by plotting the eigenvalues on the vertical axis and the factors on the horizontal axis. It is also common to include only factors that appear before the "knee" of the Scree Plot where the graph levels out.

Figure 4 shows the "rotated factor matrix" for our EFA. The columns along the top list the factors indicated by the Scree Plot and the rows list the skills that these factors "load onto." The values inside the matrix range from 0 to 1, indicating the extent to which each factor loads onto the corresponding skill. The closer this number is to 1, the more that this skill is a determining component of the factor. For example, the first cell in the matrix has a value of 0.888, indicating that skill 3D ("Analysis of Ideas") is a deciding piece of the first factor. As Knetka et al. (2019) observe, there is "no clear rule for when an item has a factor loading that is too low to be included" (pp. 11). However, it is common to omit values less than 0.3 and we follow this convention in figure 4.

The rotated factor matrix in figure 4 was created using an orthogonal rotation algorithm (Varimax). Orthogonal rotations produce uncorrelated factors whereas oblique rotations (like Promax or Direct Oblimin) allow the factors to correlate. We found similar results using Oblimin. We chose to include the Varimax rotation here because orthogonal rotations tend to produce fewer factors and uncorrelated factors tend to be easier to interpret. This six factor Varimax model explained 70.5% of the variance across the 15 variables.

EFA is used to *explore* a data set and look for patterns. There is a related technique called Confirmatory Factor Analysis (CFA) which is used to confirm that data support a previously hypothesized model. Best practice is to conduct an EFA with one sample and then confirm these results with a different sample using CFA. As Knekta et al. (2019) explain, "This confirmation should never be conducted on the same sample as the initial EFA. Doing so does not provide generalizable information, as the CFA will be (essentially) repeating many of the relationships

Figure 3
Scree Plot from Exploratory Factor Analysis



Factor	Eigenvalue
1	6.464
2	1.585
3	1.404
4	1.026
5	.878
6	.731
7	.598

Figure 4
Rotated Factor Matrix

	Factor 1	2	3	4	5	6
skill_3D	.888					
skill_3E	.817					
skill_3C	.811					
skill_3F	.623					
skill_4F		.787				
skill_4A		.646				
skill_4D	.317	.644		.373		
skill_1F			.805			
skill_1A			.661		.353	
skill_1B			.637	.378		
skill_2B				.575		
skill_2E				.538		
skill_4B		.443			.730	
skill_1E				.314	.449	
skill_2F				.357		.853

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

…we made an explicit effort to place similar or related skills on different test rubrics because we worried that faculty might give side-by-side skills similar ratings just because they appeared superficially similar.

that were established through the EFA. Additionally, there could be something nuanced about the way the particular sample responds to items that might not be found in a second sample" (pp. 8). We did not confirm our results with CFA because we did not have resources to conduct a second test assessment. Without CFA confirmation, we worried about the reliability of our EFA results.

Our small sample presented several additional challenges. Wolf et al. (2013) observe that in some situations a sample as small as 30 observations will suffice for EFA, but in general, researchers recommend 150-300 observations and a minimum of 5-10 observations per variable (MacCallum et al., 1999). With limited resources, we assessed only 72 pieces of student work for 24 skills (or variables) which amounts to only three observations per variable. However, nine of the 24 skills were not included in the EFA because of missing values. (i.e., a very small portion of the student work in our sample proved appropriate for assessing these skills). Fifteen variables put our data right on the margins of an appropriate sample size. However, researchers (Tabachnick et al., 2007) also typically recommend no more than one factor for every three variables, meaning our data should only produce about five factors. It was not our intention to collapse the 24 skills to only five skills. With these considerations in mind, we ultimately favored descriptive statistical results (the correlation matrix in Appendix 1) over the results of the EFA. Nonetheless, the EFA still proved instructive.

For example, enterprising readers will notice that the first four factors align closely with the skills grouped together in the test rubrics. Participants tended to rate items on the

same rubric similarly even though the test rubrics were designed to combine skills that appeared to have nothing in common. This pattern also appears in the correlation matrix described below. Additionally, the values reported for the fifth factor suggest there is a latent variable that partially explains both skill 4B ("Synthesize Multiple Perspectives") and skill 1E ("Interdisciplinary Knowledge"). In surveys and focus groups, faculty repeatedly told us that they felt these two skills were redundant; factor five presents one place where our implicit and explicit understandings of the learning outcomes seem to align.

To make better use of our data, we decided to focus our analysis on the correlation matrix shown in Appendix 1. The values in this matrix are the Pearson correlation coefficients which range from -1 to 1. Values closer to 0 indicate that the two skills are mostly unrelated. Values closer to 1 indicate that these skills tend to be rated in a similar manner, and values closer to -1 indicate that these skills tend to be rated in an opposite manner. We have omitted values less than 0.4 for ease of readability. Since the correlation matrix is a descriptive tool rather than an inferential tool like the EFA, we were able to include data for 22 of the 24 skills (instead of 15). Two of the skills, 1C ("Collaboration") and 4C, ("Connections to the First Year Experience") were still omitted because fewer than 10 of the 72 examples of student work provided evidence of these skills. We decided that if these skills produced so little evidence from student work, they could not be assessed using the test rubrics and should be eliminated or addressed elsewhere in the college's assessment plan.

Triangulating Findings - A first draft of new ISLOs (Fall 2019)

At this point, we had accumulated multiple types of data about our ISLOs. Some of the data (the sorting activity and the survey results) gave us insights into how faculty and staff explicitly understood our learning outcomes while other data (the correlation matrix and EFA results) examined implicit understandings of these outcomes. To make sense of all this data, the Assessment and Learning Committee formed subgroups to review each piece of data in turn. The subgroups identified learning outcomes that should be eliminated or combined based on the data they were reviewing. Recommendations were collected on a white board and discussed with the larger group.

We looked for similar findings across the different types of data and revised the learning outcomes accordingly. For example, we noted skills 3A ("Reflection on Learning") and 3B ("Connections to Experience") had a high correlation (0.79) in the correlation matrix. Likewise, in the sorting activity, many groups combined skill 3B ("Connection to Experience") with skill 4E ("Cultural Background and Identity"). We thus decided to collapse all three of these skills into one learning outcome with a rubric based on skill 3B since participants felt this rubric was the clearest. Similarly, we decided to eliminate skill 1C ("Collaboration"). We found very few items in the student work that could be used to assess this learning outcome; and during the sorting activity, many groups suggested eliminating it. By cross-referencing different types of data, we were able to make informed decisions about how to reduce the number of learning outcomes.

Finalizing the revised ISLOs (Fall 2019 - Spring 2020)

After analyzing explicit and implicit understandings of the ISLOs and using statistical tools to analyze redundancies and opportunities for revision, we presented a draft of the revised ISLOs at an Assessment Day in December 2019. We wanted to maintain the collaborative, participatory process that guides our assessment work, so we presented data from the EFA and correlation matrix alongside the draft revised ISLOs and asked faculty and staff to use the revised rubrics to assess examples of student work. Our goal was not to collect assessment data but rather to gain a better understanding of the process and experience using the new draft ISLOs. Roughly 35 faculty and staff participated. After engaging with the draft rubrics, participants completed an interactive survey that asked them to reflect on their experience using the rubric to evaluate a piece of student work.

We wanted to maintain the collaborative, participatory process that guides our assessment work . . .

This activity helped to corroborate evidence from prior assessments, the results of our statistical analysis, and insights from a broad range of stakeholders about the draft revised ISLOs. In addition, it helped us determine that some of the revised ISLOs required additional attention and a more serious overhaul. For example, we determined that the outcomes related

to technology needed further development. The framing of our original ISLOs in this area no longer reflected the kinds of work the college was asking students to do. Similarly, the outcomes related to civic engagement and global learning required a deeper dive. To revise these ISLOs, we met with faculty and staff with expertise and leadership roles in related areas - technology, science, and civic engagement/global learning. During these meetings, we reviewed the second draft of the revised ISLOs and discussed how they might better reflect student work in these areas.

We developed a second, "semi-final" draft of the revised ISLOs by meticulously incorporating data and insights collected from the activities and conversations described above. We presented this draft during a college Assessment Day in June 2020, six months after we presented the first revised draft. As part of the presentation, we again asked participants to assess student work using the draft rubrics and collected survey data about participants' experiences using the rubrics. This semi-final draft did not provide names for any of the ISLO skills. We additionally asked survey participants to suggest names for these skills to get a sense of whether the correct ideas were coming across in the rubrics. Thirty-two faculty and staff participated in this activity. We collected assessment data during these activities in addition to survey feedback. The data and survey feedback confirmed that our semi-final draft ISLOs were aligned with the explicit and implicit understandings of faculty and other stakeholders and with the work students were currently doing in our classes. We presented the revised ISLOs for ratification through the college governance process in fall 2020.

The subgroups identified learning outcomes that should be eliminated or combined based on the data they were reviewing.

The New ISLO Framework

Figure 5 shows an example rubric from the revised ISLOs we developed. The new framework addresses several of the challenges described above:

1. **Consistency between levels in the rubrics.** Since the new ISLOs were adapted from the original set (as opposed to starting from scratch), we were able to respond to stakeholder feedback about inconsistencies in the rubrics. The new ISLOs use more consistent language. For example, the lowest level of skill B on the original Broad Integrative Knowledge rubric (see figure 1) asked students to "list academic disciplines" in "one or more subject areas" while subsequent levels only required them to ask/answer questions based on a single discipline. Stakeholders told us that compiling a list of disciplines was not a compelling way for students to demonstrate integrative learning, and it seemed misguided to give students credit for integrative learning if they only asked or answered questions using a single discipline. We eliminated this language. The lowest level on the new rubric (see "Synthesizing Methodologies" in figure 5) requires students to attempt to "ask and answer questions using the general assumptions and approaches of two or more disciplines / methodologies."
2. **Redundant outcomes.** The new framework reduces the number of skills from 24 to 15 by consolidating redundancies identified in the process described above. For instance, skills B and C from the original Broad Integrative Knowledge rubric were both consolidated into the "Synthesizing Methodologies" skill on the new rubric since multiple forms of data suggested these two skills were measuring similar things.
3. **Some skills did not reflect the type of learning that was happening in the classroom.** For example, skill B from the original Broad Integrative Knowledge rubric focused on integrating academic "disciplines." One of our first-year courses asked students to conduct research using both qualitative and quantitative data. Faculty felt that this class required students to synthesize different approaches, but those different techniques were not necessarily representative of specific academic disciplines. We thus refined the rubric to refer to "methodologies" rather than disciplines.

1. **Difficulty finding examples of student work that were appropriate to assess with these rubrics.** The "Collaboration" skill was removed from our ISLOs. Multiple pieces of data suggested it was difficult to assess collaboration using the type of student work our assessment system provided. We still think collaboration is an important skill for students to learn, but we believe its assessment belongs elsewhere.

Figure 5
New Integrative Knowledge Rubric

Integrative Knowledge

*Integrative learning is the process of making connections between ideas and experiences from different contexts in order to leverage knowledge in new and more meaningful ways. This rubric, especially skill D1, is informed by Veronica Boix-Mansilla's notion of "integrative leverage" which suggests that quality work integrates different disciplines/methodologies "to generate a new and preferred understanding." Expert practitioners of these skills will integrate knowledge and modes of thinking from multiple disciplines or perspectives. They will situate issues in broader contexts and relate them to their own lived experiences. In particular, integrative knowledge is not exclusive to curricular experiences; it also applies to co-curricular experiences like student leadership, peer mentoring, tutoring, etc. In this rubric, we use the word **perspectives** to refer to perspectives of specific cultures or stakeholders as opposed to disciplinary perspectives. We use the word **methodologies** to refer to the approaches that different fields use to ask or answer questions.*

Skill	Level 1	Level 2	Level 3	Level 4
Synthesize Methodologies	Attempts to ask and answer questions using the general assumptions and approaches of two or more disciplines / methodologies, but does so ineffectively.	Effectively asks and answers questions using the general assumptions and approaches of two or more disciplines / methodologies, but does not integrate these approaches.	Integrates knowledge and approaches from at least two different disciplines / methodologies in planning and conducting research.	Integrates knowledge and approaches from at least two different disciplines / methodologies in planning and conducting research, and critically compares these different approaches.
Connections to Personal Experience	Identifies connections between one's own life experiences and/or prior knowledge to academic texts/ ideas.	Explains connections between one's own life experiences and/or prior knowledge to academic texts/ideas using basic examples, facts, or theories.	Explains connections between one's own life experiences and/or prior knowledge to academic texts/ideas using multiple, rich examples, facts, or theories.	Connects examples of one's own life experiences and/or prior knowledge to academic texts/ideas to illustrate concepts from multiple perspectives.
Contextualize an Issue	Explores an issue at the surface level, providing little insight and/or information beyond the basic facts.	Moves beyond basic facts to demonstrate an awareness of multiple perspectives on an issue.	Provides some historical/social context around an issue to explain how different perspectives relate to one another.	Situates an issue in a broader historical/ social context to demonstrate an understanding of the issue from multiple perspectives.

The new framework organizes the skills in ISLOs categories that allow students (and faculty and staff) to see at a glance the range of work they will be asked to do across their degrees. The skill "Synthesize Methodologies," for example, is listed under the Integrative Knowledge ISLO alongside two others: "Connections to Personal Experience" and "Contextualize an Issue." These three skills describe three distinct ways we expect students to integrate knowledge and methodologies they are studying in their classes. Taken together, the three skills in the Integrative Knowledge ISLO signal to students that we conceive of the "knowledge" we expect them to construct across their careers holistically, as combining lived experience, academic disciplines, and social contexts. The predecessor skills were listed in a more amorphous and curriculum-focused category, "Broad, Integrative Knowledge: General Education," as if they were relevant primarily to the student's general education classes rather than to both those classes and their program of study. This framing was helpful to us as we designed the curriculum in the college's early years; but, as our assessment results and this research project shows, it has hindered our ability to define and communicate our overall expectations for student learning.

The new framework reduces the number of skills from 24 to 15 by consolidating redundancies identified in the process described above.

Discussion

Looking back on the years-long project we undertook to assess and revise our ISLOs, three critical features emerged that might transfer to other institutional settings. These features were central in making the process unfold in an effective and inclusive manner and ensuring there were consistent spaces for reflection. We offer them up as "best practices."

Start from where you are, rather than starting over

When the revision process began, there was a strong push from many faculty, staff, and administration to start from scratch. The limitations associated with our original ISLOs had caused significant frustration; as a result, many stakeholders wanted to develop completely new ISLOs. However, the college had collected several cycles of assessment data and the working groups had issued detailed reports about how individual outcomes could be revised to reflect the kinds of teaching, learning, and values that underpin the college. These reports provided a trove of information and insights about our current ISLOs and allowed us to use what we already knew to begin the process of revising them.

The new framework organizes the skills in ISLOs categories that allow students (and faculty and staff) to see at a glance the range of work they will be asked to do across their degrees.

In other words, we approached the revision process with the goal of addressing challenges and issues we already knew existed. Rather than introduce new, unknown challenges by completely rewriting the ISLOs, we used a large body of existing knowledge to rebuild them. This approach also helped to continually propel the work forward, rather than getting caught up in new, unfamiliar challenges and issues.

Ensure faculty and staff remain at the center throughout the process

The process outlined above prioritized faculty and staff agency. Faculty and staff lead the revision process and constantly came back to the larger college community not only to reflect on the data but also to help generate new data that guided the process. This iterative, inclusive process maximized participation and increased buy-in from faculty, staff, and administration as we shared the draft revised ISLOs. The bottom-up approach we used is markedly different from an approach spearheaded by administrators or a small group of faculty/staff.

One example of how this approach worked in practice is the way we used Assessment Days throughout the year to keep stakeholders engaged. We used activities during these days to do much of the revising work. For instance, we invited participants to regroup the ISLOs and evaluate our test rubrics rather than simply presenting our findings. During several Assessment Days, we asked faculty and staff to use draft ISLO rubrics to assess student work and then provide feedback on their experiences. This iterative and inclusive process helped capitalize on the expertise of the practitioners who will use these rubrics.

A holistic approach to gain insight from stakeholders

Earlier, we outlined the process of accessing faculty and staff members' *explicit*, as well as *implicit*, understandings of our learning outcomes. For instance, we facilitated an activity during an Assessment Day that invited members of the college community to regroup existing ISLOs in order to understand how they think about them. Additionally, faculty, staff, and administrators engaged in a structured discussion about larger institutional values that we identify as critical to our college. Activities along this vein provided us with data and insights about how community members explicitly think and feel about ISLOs. We also accessed their implicit understandings by way of analyzing assessment data of student work with Exploratory Factor Analysis. These activities provided us with data and insight about how community members use the ISLOs and rubrics in practice. We were then able to compare these different types of understandings to identify redundancies in our ISLOs.

Rather than pursuing one over the other, bringing together these different strands of assessment data and analysis provided us with a more comprehensive snapshot of our ISLOs. Comparing these two data also produced evidence of potential inconsistencies between how community members explicitly think and talk about ISLOs and how they make use of them to assess student work. For example, faculty and staff explicitly identified "Quantitative Data Analysis" and "Quantitative Problem Solving" as referring to the same skill. However, our

EFA indicated that these skills were not assessed similarly and did not overlap. This holistic approach provided us with more nuanced, inclusive perspectives on current ISLOs and enabled us to make targeted revisions.

Conclusion

Our college started using the revised ISLOs in the 2020-2021 academic year. Preliminary feedback has generally been positive. This article provides a deep reflection on our revision process, lifting up key themes and considerations we identify as recommendations to any institution grappling with developing or revising ISLOs. As we note above, three principles emerged through our work that seem particularly salient for colleges embarking on similar outcomes revision projects: starting our revision process from where we were rather than replacing our existing ISLOs wholesale; striving to ensure faculty and staff stakeholders played leading roles throughout the process; and using multiple approaches, including inferential and descriptive statistics, pilot assessments, surveys, and small and large group discussions to develop a holistic understanding of our existing ISLOs and possible revisions.

We… found that these consultations increased buy-in and ensured that each subsequent draft of the rubrics better reflected our values as a community.

These principles helped us mitigate challenges identified by other practitioners related to stakeholder misperceptions of purposes and uses of ISLOs (see Colson et al., 2018; Schoepp & Tezcan-Unal, 2017). They also helped us negotiate the limitations of our own work, including the restricted sample we used for our test assessment and the well-intentioned complexities of our existing ISLOs while we worked with our colleagues to construct what Stanny (2018) has described as a "culture of improvement" (p. 114). As a concluding point, we note that we plan to continue consulting with the college community as we roll out the revised ISLOs. We did this during each step in the revision process and found that these consultations increased buy-in and ensured that each subsequent draft of the rubrics better reflected our values as a community.

References

Association of American Colleges and Universities. (2009). *Valid assessment of learning in undergraduate education (VALUE)*. https://www.aacu.org/initiatives/value

Colson, T., Berg, B., Hunt, T., & Mitchell, Z. (2018). Simple, transparent, and less burdensome: Re-envisioning core assessment at a regional public university. *Journal of Assessment and Institutional Effectiveness*, *7*(1-2), 92-114. https://doi.org/10.5325/jasseinsteffe.7.1-2.0092

Heinrich, W.F. (2017). Toward ideal enacted mental models of learning outcomes assessment in higher education. *Journal of Applied Research in Higher Education*, *9*(4), 490-508. https://doi.org/10.1108/JARHE-10-2016-0064

Kinzie, J., and Jankowski, N. A. (2015). Making assessment consequential: Organizing to yield results. In G. D. Kuh, S. O. Ikenberry, N. A. Jankowski, T. R. Cain, P. T. Ewell, P. Hutchings & J. Kinzie (Eds.), *Using Evidence of Student Learning to Improve Higher Education* (pp. 74-93). Jossey-Bass.

Knekta, E., Runyon, C., & Eddy, S. (2019). One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE—Life Sciences Education*, *18*(1), 1-17. https://doi.org/10.1187/cbe.18-04-0064

Lumina Foundation. (2014). *The degree qualifications profile*. https://www.luminafoundation.org/resource/dqp/

Matuga, J. M., & Turos, J. M. (2018). Infrastructure support for using assessment data for continuous improvement. *New Directions for Teaching & Learning*, *2018*(155), 81-88. https://doi.org/10.1002/tl.20306

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, *4*(1), 84. https://doi.org/10.1037/1082-989X.4.1.84

Roscoe, D. D. (2017). Toward an improvement paradigm for academic quality. *Liberal Education*, *103*(1), 14-21.

Schoepp, K., & Tezcan-Unal, B. (2017). Examining the effectiveness of a learning outcomes assessment program: A four frames perspective. *Innovative Higher Education*, *42*(4), 305-319. https://doi.org/10.1007/s10755-016-9384-5

Stanny, C. J. (2018). Putting assessment into action: Evolving from a culture of assessment to a culture of improvement. *New Directions for Teaching & Learning*, *2018*(155), 113-116. https://doi.org/10.1002/tl.20310

Stevenson, J. F., Hicks, S. J., & Hubbard, A. (2016). Evaluating a general education program in transition. *New Directions for Evaluation*, *2016*(151), 37. https://doi.org/10.1002/ev.20197

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5, pp. 481-498). Pearson.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and psychological measurement*, *73*(6), 913-934. https://doi.org/10.1177/0013164413495237

Appendix 1 – Correlation Matrix

Note: Positive correlation coefficients with values less than 0.4 have been omitted for ease of reading. Higher values have been highlighted in progressively darker shades of gray to emphasize where the strongest correlations exist. There were very few negative correlations, but these have been italicized to help distinguish them from the positive values.

	1A	1B	1D	1E	1F	2A	2B	2D	2E	2F	3A	3B	3C	3D	3E	3F	4A	4B	4D	4E	4F
1A	1																				
1B	0.54	1																			
1D	0.44		1																		
1E			0.47	1																	
1F	0.68	0.63			1																
2A	0.40				0.44	1															
2B			0.40		0.40	0.49	1														
2D								1													
2E						0.77	0.44		1												
2F					0.47	0.54	0.48	*-0.22*	0.41	1											
3A				*-0.15*						0.44	1										
3B										0.50	0.79	1									
3C		0.42			0.52						0.77	0.79	1								
3D					0.45						0.73	0.72	0.83	1							
3E					0.43						0.77	0.69	0.73	0.79	1						
3F		0.43				0.41					0.52	0.66	0.61	0.62	0.61	1					
4A		0.44											0.50	0.47			1				
4B	0.50		0.51	0.46										0.41			0.56	1			
4D	0.45	0.49					0.46			0.49			0.51	0.50	0.40	0.45	0.69	0.57	1		
4E			*-0.20*														0.40			1	
4F																	0.55	0.46	0.54		1