Assessment Strategies for Competency-Based Learning—Lessons Learned

Darci Lammers, PhD
Senior Instructional Designer
University of Wisconsin-Extension
Division of Continuing Education, Outreach & E-Learning
darci.lammers@uwex.edu

Stephen Beers
Instructional Designer
University of Wisconsin-Extension
Division of Continuing Education, Outreach & E-Learning
stephen.beers@uwex.edu

Descriptors: Competency-Based Assessments, Assessment Strategies

Abstract

Competency-based education offerings are increasing in public and private colleges and universities across the nation. This shift in education is new to higher education, and with that, comes much ambiguity. The instructional design team within the University of Wisconsin-Extension’s Division of Continuing Education, Outreach, & E-Learning has been working for the past two years to develop a research-based, educationally sound model for designing and developing competency-based assessments. These efforts were a significant contribution to the first system-wide, competency-based initiative in the nation: the UW Flexible Option. This paper presents a synopsis of the instructional design team’s efforts, lessons learned, and recommendations.

Introduction

While competency-based education (CBE) is not new to higher education (Grant et al., 1979; Hodge, 2007), even in medicine, where it has been growing for the last 20 years (Frank et al., 2010; Morcke, Dornan, & Eika, 2013), how to define it on a national level remains ambiguous for practitioners and educators alike (Hodges & Lingard, 2012; Pijl-Zieber et al., 2014). One of the most significant roadblocks to implementation stems from asking instructors to think differently about content design (Hoogveld, Paas, & Jochems, 2005). This paper presents how the University of Wisconsin System is approaching CBE assessments on a system-wide scale, and the lessons learned as collaborative work among instructors, instructional designers, and administrators continues to shift thinking and practice toward competency.

Background

In January 2014, the system launched the UW Flexible Option program, a collaboration between the UW System campuses and UW-Extension’s department of Continuing Education, Outreach & E-Learning (CEOEL). The program allows students to demonstrate mastery of a subject area and earn college credit without adhering to a rigid course schedule. Since Flexible Option is meant to establish valid measures of ability regardless of study time or particular study materials, drafting requires an approach to assessment design fundamentally different from that used in traditional online courses. In fact, how assessments are created and aligned with instruction and resources is often considered the foundation of CBE (Biggs, 1996; Tillema, Kessels, & Meijers, 2000). As CEOEL instructional designers have partnered with UW faculty, difficulties and opportunities that come with thinking about assessment in different ways have been uncovered. In collaboration with faculty, CEOEL instructional designers are designing a research-based process that both identifies the characteristics of quality assessments and provides recommendations about when to focus on each of them.
The Foundation for Quality Assessments: Well-Written Competencies

A clear and concisely written set of competencies and outcomes is an essential foundation for developing authentic, well-planned assessments. CEOEL instructional designers, in collaboration with UW Flexible Option leadership, created a three-level competency hierarchy consisting of program competencies, assessment competencies, and outcomes (Figure 1).

Program Competencies

Faculty begin a competency-based assessment development by defining a set of competencies that are applicable at a program level. Program-level competencies are intentionally high-level, intended to capture the scope of a degree or certificate. The language used in program competencies defines what makes a student competent at a program level and should identify fundamental skills that are unlikely to change over time.

Assessment Competencies

Each program competency is categorized further into assessment-level competencies that define what is being assessed. Assessment competencies are likely to appear on transcripts with grades or other designations indicating mastery in the same way that course names and numbers are utilized today in a traditional course transcript. Instructional designers also use assessment competencies to verify and test the alignment of competencies and assessments within the UW Flexible Option program. Assessment competencies likely experience moderate change over time as each domain evolves and new trends are embraced.

Outcomes

Outcomes are written primarily to help students identify how they will demonstrate mastery of a particular assessment competency. They tell students what specifically to study as they explore learning resources. Of the three levels, outcomes are the most likely to change over time as resources and assessment methods evolve.

The three-tired competency hierarchy is critical to assessment design in at least two ways. First, it helps instructors and instructional designers define assessment scope. While program competencies are usually in place at the start of design, assessment authors may create the remaining two levels and assessments in tandem because of their push-pull nature. Each assessment is the concrete manifestation of assessment competencies and outcomes, and therefore sometimes helps to shape them. However, even if they are not completely fixed in advance, these levels
remain a boundary demarking what is and isn’t relevant assessment content. Second, the hierarchy provides a standard of alignment by which assessment authors can evaluate how well the skills, knowledge, and behaviors elicited by a given assessment method represent their intended targets. As upcoming sections of this paper discuss, alignment remains one of the most foundational, yet difficult, aspects of implementing competency-based assessments.

**Identifying Characteristics of Quality Competency-Based Assessments**

Working with a focus group of CBE experts, Baartman et al. (2006) created a 12-criteria “Wheel of Competency Assessment” that summarizes the qualities of high-quality assessments and depicts visually how they align with each other. The CEOEL instructional designers started with this list to ultimately develop a set of eight criteria particularly relevant to the Flexible Option program: usability, alignment, cognitive complexity, authenticity, fairness, consistency, fitness for self-assessment, and transparency (Table 1).

<table>
<thead>
<tr>
<th>Assessment Consideration</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>1</td>
</tr>
<tr>
<td>Alignment</td>
<td>2</td>
</tr>
<tr>
<td>Cognitive Complexity</td>
<td>2</td>
</tr>
<tr>
<td>Authenticity</td>
<td>2</td>
</tr>
<tr>
<td>Fairness</td>
<td>2</td>
</tr>
<tr>
<td>Consistency</td>
<td>3</td>
</tr>
<tr>
<td>Fitness for Self-Assessment</td>
<td>3</td>
</tr>
<tr>
<td>Transparency</td>
<td>3</td>
</tr>
</tbody>
</table>


The smaller set of criteria came out of discussions that first used the lens of germane versus extraneous cognitive load (Sweller, van Merrienboer, & Paas, 1998) to identify potential arbitrary barriers to student success (unclear instructions, complex interface design, etc.), which can be particularly important for a largely self-directed CBE curriculum. CEOEL instructional designers also refined the list to use language familiar to faculty that addressed the most common questions received during the design process.

Drawing from existing research, CEOEL instructional designers further developed two criteria: authenticity and cognitive complexity. Authenticity is defined as fidelity to professional practice, but this is more straightforward in some domains than in others. For example, authenticity is easier to establish in a largely procedural domain like nursing or software coding, but more ambiguous in an area such as a freshman-level music appreciation course. Drawing on Gulikers, Bastiaens, and Kirschner’s five-dimensional framework for authentic assessment (2004), recommendations were developed for fidelity to professional thought, hypothesizing that, for example, while a multiple-choice question may not be a faithful representation of practice in some domains, it has the potential to elicit professional thought if crafted carefully. Drawing on research suggesting that authentic assessments encourage self-regulation through self-assessment (Olfos, 2007), instructional designers developed recommendations for tools such as small formative assessments, comparison with expert answers, rubrics, and practice assessments.

Cognitive complexity (Webb, 1997; Webb, 1999) is not unique to CBE, but it plays an important role when assessment takes precedence over instruction, as in the Flexible Option program. Drawing from Webb’s work and research on how even traditional assessment methods such as multiple-choice tests can elicit higher-order thinking (Burton et al., 1991), CEOEL instructional designers developed recommendations on how to create tasks and questions that require making multiple connections and synthesizing ideas. This was a familiar idea to faculty used to Bloom’s cognitive domain taxonomies (Bloom et al., 1956). The instructional designers simply provide strategies and examples that can be incorporated into instructors’ chosen assessment formats.
Assessment Development Toolkit and the Drafting Process

After the instructional designers finalized the eight criteria and started using them, assessment authors became frustrated almost immediately; trying to balance all eight criteria at once can be cognitively overwhelming and lead to haphazard results that are strong in one criterion but not another. Therefore, instructional designers created an assessment toolkit that not only explains the criteria but also sequences when to focus on each through a three-stage assessment drafting process.

Stage 1: Initial Planning

The first stage, Initial Planning, involves deciding on all assessment methods for a program in order to establish project scope. This stage focuses primarily on the usability criteria, asking assessment authors to identify the assessment methods that are most useful, scalable, and maintainable over time. Assessing usability ensures that the benefits of the assessment outweigh the time and resources needed to create them.

Designers encourage authors to consider several questions during the initial planning stage such as: (1) What method(s) will be used for assessments? (2) In formats that involve a certain number of questions or tasks, how many questions or tasks will each assessment have? (3) How many attempts will be allowed for each assessment? (4) Will rubrics be used for grading purposes and if so, will they be student-facing, faculty-facing, or both? and (5) Will practice assessments be utilized?

Initial alignment of competencies and assessments remains an enduring consideration that starts during Stage 1. After several months of trial and research, instructional designers at CEOEL determined it best to assess competencies using a one-to-one or one-to-many relationship. During Stage 1, it is important to decide whether each assessment competency will be assessed on a one-to-one basis, where one assessment competency is assessed using one assessment, or on a one-to-many relationship, where several assessments are used to assess one competency.

Once these questions are considered, instructional designers and authors collaboratively begin developing a map to ensure that planned assessments align with competencies and are usable (Table 2).

<table>
<thead>
<tr>
<th>Assessment Competency</th>
<th>Learning Outcomes</th>
<th>Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyze the effect nutrition has on chronic disease and overall wellness</td>
<td>Analyze behaviors that limit the progression of chronic disease</td>
<td>Case study:</td>
</tr>
<tr>
<td></td>
<td>Evaluate resultant health outcomes based on nutritional choices</td>
<td>• Chronic disease analysis report</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Analysis of current diet and issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Create diet for the individual based on his or her likes and dislikes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Summary of specific health benefits of the new diet</td>
</tr>
<tr>
<td>(Format: dropbox submission)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Example of initial assessment competency, outcomes, and assessments map.

This initial map provides an opportunity to visually consider whether the planned methods of assessment are feasible considering faculty time and resources, scalability, and maintainability prior to investing time in assessment drafts.

Stage 2: First Draft and Feedback

During the second stage, authors create assessments and focus on the next four assessment criteria: alignment, cognitive complexity, authenticity, and fairness. Alignment evaluates how well assessments represent competencies, outcomes, curriculum, and content. Cognitive complexity challenges faculty to consider the inclusion of tasks that require higher cognitive skills. Instructional designers challenge faculty to give all assessment questions the “Search Engine Test.” If the answer to an assessment question can be found simply by typing the question in an common Internet search engine, it is likely not cognitively complex enough to evaluate a student’s true ability. Authenticity of assessments is important in CBE so that students can demonstrate mastery of their skills using real-world situations and challenges. Finally, fairness ascertains how well an assessment’s challenge lies solely in its target skills and knowledge, rather than factors like primary language, financial circumstances, access to technology, or other considerations that create a potentially unequal playing field.
This stage is where the bulk of the work in the assessment creation process takes place. All assessments, including practice assessments, and grading rubrics and schemes are created during Stage 2. Practice has shown instructional designers and authors the importance of not underestimating the amount of time necessary to create a quality assessment drafts. Instructional designers and faculty estimate the amount of time they think is needed to complete all assessments, rubrics, and grading schemes, and then double that amount. This is especially important if this is the first time either is creating competency-based assessments. Instructional designers and faculty will spend the bulk of their assessment creation time in stage two and should plan to collaborate often during this time.

As in Stage 1, instructional designers work with faculty to consider: (1) How do assessments truly measure targeted competencies? (2) How can assessments be designed to require more than memorization and repetition, where appropriate? (3) How can assessments be created to mimic professional thought and practice in some way? and (4) How can we ensure that assessments are accessible to students of different locations, financial statuses, languages, and cultures? As a result of these considerations, instructional designers and faculty often revisit some of the decisions made in Stage 1.

Stage 3: Additional Drafts
After completing first drafts, instructional designers and faculty collaborate to provide feedback and make revisions, thereby leading to the final stage. The focus of the additional drafting stage is on the remaining three assessment considerations: consistency across assessments, fitness for self-assessment, and transparency. Consistency across assessments ensures that established evaluation methods are consistent across assessors, students, and time. Fitness for self-assessment promotes student-centered, self-regulated learning. Lastly, evaluating the transparency of each assessment certifies that the instructions are clear and understandable to all participants, regardless of any external factors.

During Stage 3, faculty and instructional designers work closely to refine the initial assessment as well as any assessments intended for additional attempts (e.g., assessments students will complete if they do not achieve mastery on the first attempt). This is the time to put the finishing touches on all assessment-related materials.

Assessment Development Toolkit
The toolkit developed by the CEOEL instructional design team supports faculty by outlining this three-stage sequence and allowing content authors to drill down into specific criteria to see general steps on how to proceed, supported by examples taken from Flexible Option offerings. Designers can use all or parts of the toolkit to structure discussions with faculty members, who can then reference it as they develop assessments. The toolkit is agnostic toward content and format. Therefore, it can be used in both one-on-one and group settings across content areas. A criterion such as “authenticity” applies equally to a multiple-choice exam, portfolio, and simulated interview.

During Stages 2 and 3, instructional designers work with faculty authors to complete a competency/assessment alignment template more detailed than that which is shown in Table 1. The updated template provides a granular level of detail regarding each assessment and how it aligns with its intended competency. An abbreviated example of this tool is shown in Figure 2.

![Figure 2. Example of detailed competency and assessment alignment mapping tool.](image-url)

Some assessment authors have found the level of detail in this tool overwhelming, suggesting the tension between tracking the necessarily large amount of detail needed to craft and maintain a competency-based assessment and presenting that detail in a way that is clear and actionable. The CEOEL instructional design team is currently evaluating how an updated version might best serve faculty in the future.
Conclusion

At the time of this writing, two iterations of the assessment drafting process have been completed across multiple campuses. The CEOEL instructional design team is drafting a third iteration and organizing assessment creation by using the toolkit as a touch point between faculty and instructional designers. The goal of our presentation is to openly share both what seems to be working—such as focusing on different criteria at different times—and where we still have questions, such as those surrounding the use of a standard template to capture all assessment information.

We continue to evaluate assessments based on student feedback, student completion rates, and faculty feedback. Above all, we are learning the importance of challenging how both designers and instructors think about assessments as a prerequisite for effectively using resources like the toolkit. We share both the challenges we have encountered and our emerging solutions in the hope that we can provide insights for other organizations pursuing competency-based assessments, while also learning from those organizations.

References


