Cost-Eff ective Strategies for Developing Formative Assessments in Online Workplace Training

I. INTRODUCTION

Both formative and summative assessments play an essential role in workplace training. Summative assessments are typically used to validate that knowledge has been gained; most often, these come in the form of end-of-unit or end-of-course tests. Such assessments are undeniably important to learning, yet they are often perceived as the product of unfortunate necessity rather than an opportunity to enhance the learning experience. Formative assessments, by contrast, support learning in a non-threatening way, offering constructive feedback throughout a course and allowing learners to build upon their strengths and address their weaknesses in a more timely fashion. Particularly in online learning, formative assessments are an essential way to help learners gain and retain new knowledge, skills, and abilities (KSAs).

Many constructive, engaging, and low-cost strategies exist for allowing learners to assess their progress and accomplishments before the risks of decreased motivation, confidence, and understanding become unfortunate realities that negatively impact instructional outcomes, on-the-job performance and online learning ROI.

II. FORMATIVE ASSESSMENTS AS A PARADIGM SHIFT

The most traditional form of assessment, summative assessment, is intended to measure the sum of knowledge, skills, and abilities acquired in a unit or course. It contributes to the course grade and/or successful completion of the course. Formative assessments play a very different role: they usually do not contribute to the grade of the course, and they may or may not be required for course completion. Instead, formative assessments focus on providing useful practice and feedback throughout a course to assist the learner in acquiring and applying the new KSAs.

Since the early 1990s, there has been a clear shift toward adding formative assessments—assessments that further the learning experience without the intimidating fallout of summative assessments. Where these are combined with summative assessments, they can improve the learner’s performance on the end-of-unit and end-of-course assessments and reduce test anxiety for all types of assessment in the course.

As early as 1994, Gipps noted that assessment is “undergoing a paradigm shift, from psychometrics to a broader model of educational assessment, from a testing and examination culture to an assessment culture” [1]. Newer forms of assessment—formative, performance-based, norm-referenced assessment—have expanded the concept of assessment greatly, suggesting that the traditional model is “no longer adequate” [1]. Some even suggest that assessment, not instruction, is the major influence on students’ learning, influencing not only what is learned but how learners approach the learning process [2], [3].

Formative assessments are an essential part of this paradigm shift. They do more than just measure the amount of learning that took place; they support learning by yielding information that can be used to improve instruction and enhance the learning experience. They provide feedback that enables learners to identify their individual strengths and weaknesses in a non-threatening way. At their best, they also offer the learner a constructive way to address any deficits.

While summative assessments also support learning when they show which questions have been answered incorrectly, they offer less robust information. They rarely offer a means for learners to reflect on their own perform-
ance, and they rarely provide feedback to guide a learner to improve performance, make better judgments, or attain higher levels of skill.

In some instructional situations, data from summative assessments is also less valid. Summative assessments can inspire a test anxiety that suppress performance, thus leading to an inaccurate measure of a learner’s abilities.

By contrast, formative assessments are non-threatening. Learners may be given the opportunity to try again when responses are incorrect, such that all learners can ultimately get all answers correct. The assessment may be optional. Formative assessment data may lend itself less easily to being marked in percentage points or alphanumeric grading (e.g., a reflective essay, or an analysis of a case study. In some circumstances, assessment data may not be captured, or it may be captured solely for the purpose of improving future versions of the courseware.

Formative assessments can offer amazing flexibility and a wide range of learning strategies, particularly in the context of asynchronous online learning. Feedback can be provided instantaneously, allowing immediate remediation and helpful guidance while the content is still fresh in the learner’s mind. Additionally, as Velan et al observe, this delivery method supports “equity and inclusiveness by allowing students to attempt each assessment anonymously on multiple occasions, at any time, and from virtually anywhere” [4]. This is a particular advantage to learners from “culturally and linguistically diverse backgrounds [who may] fear embarrassment if found to be in error; …[for them], [o]nline formative assessments provide a safe environment, where trial and error is permitted” [4].

III. THE ROLE OF COMPLEX SIMULATIONS

A. Simulations

High quality simulations allow learners to apply what they have learned, testing themselves on their ability to use the new KSAs in a situation that mirrors a situation they may encounter in the workplace. First responders might explore potential responses to a natural disaster or terrorist threat. Medical personnel might explore options in helping patients. 3D modeling, video gaming, and complex decision tree branching can be used to assess a learner’s mastery of content, allowing the learner to see the consequences of certain decisions.

As an example, Rival Interactive offers interactive training for medical personnel in a Neonatal Intensive Care Unit (NICU). After exploring the instructional screens, the NICU training presents a scenario in which a newborn baby has stopped breathing. The learner must make and implement rapid decisions to save the infant’s life. Each decision can be virtually enacted; with each passing moment, the infant’s skin tone reflects lack of oxygen (“blue baby” syndrome) until the correct responses have been made. This is a very effective way in which to assess the impact of medical choices: student mistakes have no consequences to a live infant, but the consequences can be explored virtually. When the correct medical decisions are implemented, the infant returns to smiling health.

B. Game-Based Learning

Where assessments can be made to simulate games, learner motivation increases greatly. Within medical education, game formats have been modeled after Trivial Pursuit, Survivor, and Who Wants To Be A Millionaire [5]. In many instances, learners will replay learning games beyond what is instructionally necessary. Best-of-class include complex video games, such as those used by DOD to simulate situations encountered in war. These involve 3D modeling and complex decision-making threading.

As an example, Incident Commander, which was built in Mosbe, received critical acclaim for training first responders. As described in Business Week,

“Incident Commander…has already proved itself in the field. One of Katrina’s first responders, Joseph Barlow, used what he had learned from a test version to help set up an 800-bed field hospital in Baton Rouge. A paramedic from rural Illinois who had never supervised anyone before, Barlow drew on Incident Commander's step-by-step approach to disaster management to run logistics for the hospital, which treated 6,000 patients. It was "an excellent example of training becoming reality," says Glenn Schmitt, acting director of the Justice Dept.’s National Institute of Justice.” [6]

Developed by Breakaway Ltd., Mosbe offers develop- ers the ability to rapidly and cost-effectively build digital environments from real-world data, script scenarios and free-form exercises, and draw from a library of 3D models. It is used by business and the military.

C. Complex Writing Assignments

To assess higher levels of learning, learners can be asked to develop responses to case studies, formulate action plans, or perform other assessment tasks that can reflect their ability to apply concepts. Computational linguistics technology to assess longer forms of writing, such as essays, does exist, though to date it has been more primarily used in academic contexts. Existing studies indicate a high degree of accuracy in these settings (91-97%) [7], [8]. While these might be used in adult training—for example, to allow learners to design solutions for case studies—the existing software does not seem precise enough for most job-based applications, with the exception of those that assess quality of writing as well as content.

The e-rater, developed by the Educational Testing Service, has an exceptionally high accuracy rate: 97%, as compared with ratings given by human graders [8]. The e-rater scans for more than 50 features within an essay; many of these concern writing quality rather than content. It can be used for any type of content, but hundreds of samples must first be graded and processed by the system. It has been used successfully in national job training programs, such as Job Corps.

IV. COST-EFFECTIVE AND TIME-EFFECTIVE SOLUTIONS

The high tech options described above can involve considerable costs: in labor hours, in software costs, and/or in media costs. But effective formative assessments need not employ all the bells, whistles, and expenses to be effective.
COST-EFFECTIVE STRATEGIES FOR DEVELOPING FORMATIVE ASSESSMENTS IN ONLINE WORKPLACE TRAINING

A. Simple Quizzes

The simplest type of formative assessment is a quiz (multiple-choice, matching, true/false, etc). Periodic quizzes that do not directly affect a learner’s grade can serve as effective checks that learning has occurred, assisting the learner in determining which portions of content have not been absorbed.

While some instructional designers are wary of frequent quizzes in online learning—and it is definitely possible to quiz too often—employing quizzes throughout a course is instructionally effective, particularly where each section of a course builds upon its predecessor. By helping a learner know where essential content has not been mastered, learners do not risk building on a faulty foundation.

Because quizzes assess smaller portions of content than summative tests, a learner has higher motivation to review content: the content will be more recent, more compact, and easier to find. The learner may be concerned that the material will be on the final test, increasing motivation. After a summative evaluation, motivation is typically less—particularly if the final grade cannot be changed.

The literature indicates clearly that frequent testing with feedback improves learning (e.g., [9], [10], [11], [12], [13], [3], [14]). Frequent testing with feedback increases retention by helping learners recognize, understand, and remember the most essential elements of instruction. Test anxiety decreases, since learners have a sense of what they will be asked at on summative assessments, and performance improves in summative assessments. Learners who take optional formative online quizzes performed better—usually significantly better—on summative examinations than those who did not.

B. Quizzes to Assess Higher Level Thinking

Whereas simple quizzes are best at measuring knowledge and comprehension, more complex questions are typically needed to assess the higher levels of Bloom’s taxonomy for the cognitive domain: application, analysis, synthesis, and evaluation [15]. Even without expensive and complex essay grading software, it is possible to employ formative assessments that assist learners in applying, analyzing, synthesizing, and evaluating material.

In courseware developed for the EPA, we used low-tech formative assessments to help learners assess whether they could determine the best ways to reach to a chemical release scenario. In one assessment, learners were asked to evaluate the adequacy of an initial evacuation zone. After reading instructional screens, documents in a Resources section, and performing any additional research desired, they were asked to formulate an Action Plan. Learners could assess the correctness of their response in two ways: by comparing their work to an example of an effective plan, and by answering specific questions to evaluate their own Action Plan.

C. Effective Feedback

Feedback is, in many ways, the most central part of formative assessment. Effective feedback does more than identify strengths and weaknesses. It assists learners in learning more effectively, providing insights into good learning processes. And feedback can be effective without being costly to develop.

Even in a basic multiple-choice quiz, feedback can be contextualized. In addition to the standard congratulatory messages for correct answers, feedback for incorrect responses may include additional information or links to supplementary resources (e.g., articles, web pages). Links can also go to the pages within the lesson itself, so learners can review the screens in which the concept was first presented. Most low-budget technologies permit some type of contextual feedback.

Where providing feedback for each potential answer choice is prohibitively time-consuming, general help can be provided for all answers missed in a quiz: general suggestions can help learners find the correct answers for themselves.

When learners are offered the opportunity to determine the correct answers themselves, feedback has the greatest influence. [9] Though simply explaining the correct answer is useful, there will be a greater influence when users have an opportunity to reflect further and make a correct choice. To prevent learners from simply clicking any answer to access the correct answers, many assessment creation tools and learning environments offer question banking and answer shuffling.

D. Simulations

Since adult learners are most highly motivated and retain information best when the instruction mirror situations they may encounter in their workplace, simulations can be highly effective. For many instructional purposes, simulations need not be costly or complex. Firefly if an effective technology for building simulations for software training, since it offers easy to use, robust features that offer the user an immersive learning experience that duplicates the experience of using the actual software—users often think they are using the software itself rather than a simulation of it. Captivate, by Adobe, offers a less robust but easier to use simulation development package.

In Firefly, a learner first progresses through instructional screens that explain functionalities, applications, and any appropriate standard operating procedures, then moves to a screen with a link to a simulation in which the learner “uses” the software, following captioned instructions. If a learner makes an incorrect response, the system offers to demonstrate the correct procedure—the user can accept, watch the procedure being demonstrated, then try it, or just click “no” and make another try. If desired, a course can be programmed such that a learner must complete a simulation before progressing further into the course.

E. Game-Based Learning

Game-based learning need not be expensive to be fun or effective. As Hudson and Bristown show, the natural competitiveness of the learner population can be effectively harnessed by this method [5]. With a basic understanding of Flash programming—or some inexpensive or free software—it is possible to create Jeopardy-style games and other games that make assessments less intimidating. At their best, such games encourage learners to replay situations and learn more. There are many free and low-cost game creation software tools on the web.

As an example, Information Experts created a Flash-based formative assessment game for the EPA, entitled the RadTown SpeedQuiz. The learner has 30 seconds to answer as many questions as possible. Correct answers are worth 10 points; wrong answers subtract 5 points, and 3
right answers end the game. When the game ends, a "fun fact" appears for each question answered. Learners can play the game multiple times, with new questions appearing from the question bank.

F. Skill Meters

Learners sometimes want or need to compare themselves with others—with an average learner, a top learner, or even an expert. Struggling learners may be encouraged to know that their errors are no greater than those of the average learner. High-performing learners may enjoy comparing their expertise with other top learners or with an expert in the field. A skill meter can allow these comparisons with others and/or with course expectations, as desired.

Submission of results can be optional or required. Making submission optional reduces test anxiety, but it also reduces the validity of the results, since less successful learners may be less likely to submit. Some type of incentive to participate may be helpful. Bull, Quigley and Mabbott describe an Open Learner Model (OLM) in which the only way learners can receive feedback was to access an "OLMlet" [16]. Learners accessed it often, and their feedback indicated that they found it helpful. Of the 223 users in Bull and Mabbob's experiment, only one commented that it was "demotivating" to discover that his/her mastery was less than that of other learners [16].

In addition to showing learners where they need to review of material already covered, a skill meter can facilitate the setting of realistic goals—e.g., if few learners achieve more than 80% on a given skill meter, their mastery may be sufficient to move ahead. Note that skill meters require that a course be connected to a database; for coursework that resides in an LMS, developers would need access.

G. Computer-Adaptive Tests

Computer-adaptive tests (CAT) can be very useful in formative assessments. In a CAT, the difficulty of questions is dynamically matched to the proficiency level of each learner. The first questions are of medium difficulty; correct responses trigger more difficult questions, while incorrect responses trigger less difficult questions. Thus, assessments are customized to learner needs with relatively simple programming. In the model described by Lilley and Barker, if there is no historical data on questions, the ratings are updated after each assessment session based on learner performance [13].

Feedback can also be individualized. This is particularly important for learners who are having difficulty with the content, since as Lilley and Barker observe, "Feedback provided at a level too high for a student is less than useful if they do not understand basic concepts. Equally there is no point in providing feedback on questions that a student already understands and can answer. With a CAT, students are tested at the boundary between what they understand and what they do not know. This is an important boundary, as at this level students have good motivation, neither being discouraged by questions that are too hard, or demotivated by questions that are too easy" [13]. In this way, the CAT provides scaffolding for the next level of KSAs.

The CAT can also include a personalized revision plan to address learner needs. In one model, performance on the assessment triggers "a set of revision tasks that match [learners'] current level of ability within the subject domain, [ensuring] that less able students are not provided with revision tasks that are too hard and therefore bewildering or frustrating ...[and that] more able students are not presented with revision tasks that are unchallenging and therefore de-motivating" [13].

H. Decision Trees

Another way to adapt a formative assessment to a learner is to employ decision trees. At their most complex, they are time- and cost-intensive, as in video games, but simpler models are also effective. At their simplest, they are still effective.

As an example, a learner’s answer to a question in a scenario can move that learner to a question or game scenario that builds upon that choice. As long as the branching always returns to the same place (e.g., all learners progress to the beginning of the next module), development time is kept within reasonable bounds. Information Experts is currently building training for Federal Retirement Thrift Investment Board that offers multiple learning paths. Learners’ goals as well as success with the content (e.g., understanding of investment fundamentals) determine which games and other formative assessments will be presented to them in a given module.

V. Avoiding the Pitfalls of High Cost and Ineffectiveness

The first step to putting those strategies into effect is to recognize and avoid the barriers to creating effective formative assessments in asynchronous online learning. These include limited cost, time, and programming expertise. They also include avoiding common pitfalls in formative assessment creation.

A. Control Cost and Time

Cost and time are relatively easy to control. Graphics can contribute heavily to costs, but an instructional designer or developer need not sacrifice visual quality: there are many low-cost graphics online. If time is not an issue, capturing photos with a digital camera not only reduces costs; it is an effective way to tailor an assessment the population being trained.

Where information from the formative assessments does not need to be captured, less programming may be required, minimizing costs. Using inexpensive software packages and add-ons that automate assessment production can reduce both costs and development time.

Instructional design time can be reduced by creating formative assessments that resemble summative assessments. Many researchers believe formative assessments are most effective when they are similar in format to the summative assessments (e.g., if the summative exam involves multiple-choice and matching questions, then the formative assessments should include similar questions [11], [14]).

B. Ensure that Learners Take the Assessments

For formative assessments to work, learners must take advantage of them. They may not be used if no incentives were present [17], [10]. If learner motivation may be an issue, requiring assessment completion may be advisable. As long as learners understand that their results are not
being captured, there should be no text anxiety or other negative effects.

Where formative assessments are optional, participation does not necessarily depend upon ability. A study of 1300 undergraduates at the University of Sydney showed that more advanced students used formative assessment less frequently than struggling students [18]. In Olson and McDonald’s study of 98 dental students who were offered formative online assessments, 45 chose to take the assessments, with “no statistical differences between the accumulative GPAs of either group or the accumulative GPAs of either group minus science courses that utilized the formative exam questions” [3].

C. Avoid Inappropriate Feedback

Not all feedback is good feedback. The classic study by Kluger and DeNisi found that while feedback on average does increase achievement (60%), in 40% of the studies they analyzed, feedback had the opposite effect of decreasing performance [19].

Game-based learning sometimes offers feedback that references the learner rather than the task (e.g., “Clueless!” rather than “Review this article for more information”). Research indicates that feedback is most effective when it focuses on the content, not the person answering the question. The focus should be on why an answer is incorrect—not the person who gave the answer. What may be amusing to one learner can be humiliating to another.

Where assessments provide only information about what was done incorrectly, or not achieved, without providing the confidence and means to address these issues in a non-ego-threatening way, actual pain can result—and pain is not conducive to learning. A 2003 study at UCLA used magnetic resonance imagery (MRI) to look at the effects of negative feedback from a computer game on the brain. In this study, Eisenberger determined that the psychological stress of negative feedback was translated into brain activity in the part of the brain associated with physical pain—in other words, subjects suffered pain. Pain was markedly greater where the student was rejected as an individual (where the students were rejected by what they thought were other students, as opposed to being knocked out of the game by what seemed to be a technical problem) [20].

VI. THE CONTEXT OF HYBRID COURSES AND EVOLVING ASYNCHRONOUS ONLINE COURSES

Asynchronous online assessment tools are sometimes used within hybrid courses. Here, formative assessments can play an important role in improving the quality of the instruction as well as the quality of the learning. Formative assessments typically provide knowledge at the topic or lesson level, as opposed to the unit or course level, and this can be essential to those designing the education. A live trainer can put additional emphasis on points not fully understood by many learners.

Where a hybrid course includes synchronous content, additional elements can be incorporated into formative evaluations. With audio or videoconferences, online chat, webinars, computer-equipped classrooms, or learners working independently but synchronously, peer projects and multi-player learning games can be incorporated.

Similar advantages exist in online courses that are updated or re-released. Informed by formative assessment data, instructional designers can add text, graphics, supplementary information, etc., to expand upon content areas that learners have more difficulty absorbing.

VII. SUMMARY

Technological advances have opened a rich array of cost-effective possibilities in developing formative assessments. Simulations and game-based learning can allow a learner to apply new KSAs to (virtual) real-life situations. Quizzes, including those that involve higher order thinking, can offer a learner timely and non-threatening insights into content areas that may require review. Well-written feedback can offer guidance on the most appropriate next steps for an individual learner. Skill meters, CATs, and Decision Trees further individualize the learning experience.

As online learning tools for the workplace, formative assessments can facilitate deep learning as well as assess basic comprehension. They ensure a solid foundation for content mastery, and they provide feedback that facilitates a rich and productive learning experience.

As a means of increasing and sustaining learner motivation, formative assessments are also highly effective. They boost the learner’s confidence that the material can be learned, and they offer a non-threatening path to continual learning. At their best, they recognize each learner’s unique combination of existing knowledge, motivation, learning styles, and need to apply the new learning. In this way, formative assessments offer a way to customize assessments to learner needs, increasing instructional effectiveness as well as the learning experience itself.

REFERENCES

COST-EFFECTIVE STRATEGIES FOR DEVELOPING FORMATIVE ASSESSMENTS IN ONLINE WORKPLACE TRAINING


AUTHORS

Theresa M Mackey is Director of Learning and Training at Information Experts, Reston VA 20190, USA (e-mail: tmackey@informationexperts.com). She is also a faculty member at Northern Virginia Community College, Annandale VA.

David R. Derr is Vice President of Education Services Information Experts, Reston VA 20190, USA (e-mail: dderr@informationexperts.com). He is also a doctoral student in Instructional Technology at the University of Texas, Austin TX.

Ellie O’Connor is a User Experience Analyst with Information Experts, Reston VA 20190, USA (e-mail: eoconnor@informationexperts.com).

This article was modified from a presentation at the ICELW 2009 conference in New York, NY, USA, June 2009. Submitted 3 July 2009. Published as resubmitted by the authors on 10 October 2009.