Adding Rigor to Classroom Assessment Techniques for Non-Traditional Adult Programs: A Lifecycle Improvement Approach

Jason E. Thomas | Philip E. Hornsey

Formative Classroom Assessment Techniques (CAT) have been well-established instructional tools in higher education since their exposition in the late 1980s (Angelo & Cross, 1993). A large body of literature exists surrounding the strengths and weaknesses of formative CATs. Simpson-Beck (2011) suggested insufficient quantitative evidence exists on the utility of CATs for increasing the quality of student learning. Two quantitative studies by Simpson-Beck (2011) and Cottell and Harwood (1998) indicated no such correlation. We suggest that these deficiencies as applied to adult non-traditional programs may be due to a lack of rigor in the construction of many formative CATs, as well as a failure to properly match assessments to real learning objectives. In this article, we propose a nine-step framework to facilitate proper selection of formative CATs with appropriate rigor and implementation in the classroom.

Instructors need to provide meaningful feedback to promote effective learning and monitor the effectiveness of instructional techniques. However, students in higher education often find instructor feedback to be missing or of low quality (Ferguson, 2011). Additionally, in classes where knowledge builds upon itself, such as in the teaching of finance or accounting, failure to learn a core principle could hinder a student's progress in the rest of the course.

A primary challenge in higher education is assessment of the progress of students. Assessment is the practice of determining definitive and measurable goals, compiling and interpreting evidence that reveals the extent to which the goals have been achieved, and determining and effecting changes to enhance attainment of the desired goals (Suskie, 2004). This challenge is exacerbated when periods of time are compressed and instructor contact is limited as often happens in accelerated non-traditional adult learning programs.

This situation is further complicated by the growing number of non-traditional adult learning programs. According to Ross-Gordon (2011), adult non-traditional students account for a large portion of undergraduate enrollments. The National Center for Education Statistics (2009) indicated that in 2007 38% of the more than 18 million college students were over the age of 25. The agency also predicted enrollment would grow or remain stable during the period of 2007-2018 (Hussar & Bailey, 2009).

Whereas a college semester can range from 14-20 weeks with two to three classes per week in traditional programs with instructor contact ranging from 28-60 instances of student-instructor contact, non-traditional adult programs are trending toward compressed time frames with three hour classes being delivered in as little as five weeks with only five instances of student-instructor contact (Wlodkowski, 2003). The combination of less frequent and extended classroom sessions makes it imperative for an instructor to quickly ascertain if his or her students are struggling; otherwise, a large part of the class may elapse before a change in instructional tactics is implemented. Therefore, it is critical that instructors be able to accurately assess student learning quickly and make adjustments as needed.

Limited instructor contact in non-traditional adult programs provides fewer opportunities for in-

structor-student feedback and interaction. Accordingly, feedback must be timely and meaningful. The most important tool in an instructor's arsenal is the provision of timely, detailed, and individualized feedback (Jarzebowski, Palermo, & van de Berg, 2012). Since the late 1980s, one of the most influential feedback tools has been Classroom Assessment Techniques (CATs) (Carduner, 2002; Choinski & Emanuel, 2006; Eisenbach, Golich, & Curry, 1998; Goldstein, 2007; Simpson-Beck, 2011; Suskie, 2004; Walker, 2012). It is our belief that the lack of correlation–focused research studies may be due to a failure to select the correct formative CAT to support the required learning objectives and lack of sufficient rigor embedded in such exercises for compressed classroom and student-instructor interactions.

CLASSROOM ASSESSMENT TECHNIQUES DEFINED

Classroom Assessment Techniques include a broad array of tactics that can be grouped into two categories: summative and formative (Simpson-Beck, 2011). Summative CATs, such as testing or student-based ratings of instruction (e.g. surveys), are often used at the end of an instructional period to determine the effectiveness of the learning experience and how much information has been retained (Simpson-Beck, 2011). Formative assessment can be defined as quick exercises integrated into instructional contact time, which allow for an on-the-spot assessment of a student's understanding of a learning objective (Simpson-Beck, 2011). Theoretically, such a tool allows an instructor to determine which topic students have failed to properly assimilate, as well as offer immediate feedback to those students who might not have realized that they grasped a concept inadequately (Angelo & Cross, 1993; Simpson-Beck, 2011).

The division between formative and summative CATs is an important one. A formative CAT is a tool to guide and manage instruction (Simpson-Beck, 2011). A summative CAT is a final evaluation of the student's understanding of the topic when instruction is complete (Adams, 2004). Summative CATs may be indicators of instructional aptitude, student industry, student preparation (or lack thereof), basic student intellectual ability, or a combination of all four (Adams, 2004). When the term "CAT" is used generically and without the modifier "summative" or "formative," the meaning is ambiguous. Indeed, the nearly endless debate in primary and secondary education with regard to student testing often includes the term "CATs" and virtually always refers to some form of summative testing (Adams, 2004). The authors of this article are interested in exploring formative assessment. Our goal in doing so is to suggest a framework for the analysis and construction of formative assessments, which might better deliver on the promise of these techniques. Further, we suggest that for non-traditional adult programs CATs should provide immediate feedback in the class session due to the limited number of classroom interactions.

PROMISE OF FORMATIVE ASSESSMENT

Formative assessments are used to demonstrate how successfully knowledge has been transferred to a student and to provide an opportunity to adjust and correct the instructional process prior to summative assessments (Adams, 2004). This is a mutual opportunity for the student to request assistance or the instructor to provide assistance that the student might not know is required. Based on this, one might reasonably suggest that, assuming they work, implementing CATs in the classroom would consist of using a series of progressive formative assessments to maximize learning. This would then be measured by student performance on summative assessments, which are generally recognized as the principal indicator of success in the coursework.

Cross (1987), Angelo and Cross (1993), and Mosteller (1989) are responsible for the early promotion of formative assessment CATs in the United States. In particular, Angelo and Cross (1993) created a movement around formative assessments in the late 1980s, which has been embraced by both researchers and instructors. As such, even those who might question the efficacy of formative assessment consider these works canonical.

BACKGROUND

The utility of formative CATs has been an important topic since Angelo and Cross's work in the late 1980s. The publication of Classroom Assessment Techniques: A Handbook for College Teachers in 1993 cemented the concept within academia. The discussion in the literature rapidly moved from the importance of these techniques to an examination of those with the greatest utility. While Gold-

stein (2007) might have asserted that formative evaluations are less common in higher education, he proceeded to outline several and then demonstrated statistically that formative evaluations are of the highest utility in the teaching of statistics itself. Nine years before, Cottell and Harwood (1998) expounded in the same manner, selecting specific formative CATs and demonstrating their usefulness for those studying accounting. In the same year, Melland and Volden (1998) provided similar advice for the instruction of nursing students. In fact, the literature is replete with examples of this type of paper, wherein an academic of a given discipline contends that formative assessments are used in their discipline too infrequently, seeks to prove their utility for their area of study, and then suggests the techniques best suited to promote that erudition. Penetration is so complete that Choinski and Emanuel (2006) examined Angelo and Cross's (1993) work for its utility in the assessment of the effectiveness of one-hour library classes.

In recent publications, Carduner (2002) examined the efficacy of formative assessments in Spanish language instruction. Steward, Mickelson, and Brumm (2004) evaluated both summative and formative assessments and concluded that both had reinforcing effects, helping to provide insights into the instruction-learning process. In each case, techniques were selected from the growing list of recommended formative assessments, and, in most cases, some demonstration was made of their overall utility. It is far rarer to see any treatment as to how the assessments may be successfully developed or what cross-disciplinary traits distinguish a useful formative assessment from one of limited utility. To that end, it is important to consider what might happen if an instructor used a CAT that seemed effective at the time, but it did not produce the desired results. Might he or she end up giving up on CATs altogether, labeling them a mere fad?

Doubts about formative classroom assessment techniques do exist. Simpson-Beck (2011) decried the anecdotal nature of qualitative studies of CATs that have been published since Angelo and Cross (1993) brought them to the fore. Mosteller (1989) suggests that the qualitative assessment, the "Muddiest Point," one of the first and most commonly cited techniques, may not contribute to learning. This then begs the question if the technique is being used properly. In fact, Angelo and Cross (1993) have been clear from the beginning that CATs are ineffective if they are not matched to the course goals; however, Simpson-Beck's (2011) study was designed to control for this, and still raises questions at to Classroom Assessment Techniques' utility.

With the literature thus divided between the firmly convinced, who write principally to aid in the selection of the proper formative CAT, and those who are concerned that only anecdotal evidence exists as to their utility, academics are left with questions. The first is, "Are formative CATs effective?" It seems intuitive that using techniques to glean student comprehension is better than not doing so at all. Perhaps the fact that the formative CATs often seem more like a simple formalization of the phrase "do you understand what I just said?" has facilitated their quick adoption.

Perhaps the question better put is, "Is there a reason that an intuitively useful formative CAT such as the Muddiest Point would fail?" One might therefore look at a means to reduce a formative CAT to its constituent parts to find such flaws. If this provides an answer as to why Simpson-Beck's (2011) test failed to find utility in the Muddiest Point, then perhaps this reductionist framework might be used to provide a methodology to both evaluate and construct useful assessments.

Supporting both the need for and potential of such a framework, Eisenbach et al. (1998) authored a paper on the cross-disciplinary effectiveness of CATs. As a result of this study and such venerable works as Dewey (1933), who suggested that all learning is predicated on cross-disciplinary principles, it is reasonable to suggest that all useful formative CATs must conform to certain principles and be subject to consistent methods of analysis. Tying this together is the perceived need to provide frequent and meaningful feedback to all students regardless of discipline (Jarzebowski et al., 2012).

Few quantitative studies exist to validate the effectiveness of formative CATs (Simpson-Beck, 2011). Variability in the quality of instruction, the capacity of students to learn, the duration of classes, and frequency of contact make quantitative studies on the effectiveness of CATs challenging. With this said, perhaps the greatest ally that the mass adoption of formative CATs has had is the desire among instructors that they should work. Non-traditional learning environments increasingly feature smaller numbers of class meetings with more limited

contact between students and instructors (Ross-Gordon, 2011; Wlodkowski, 2003). Consequently, a clear need exists to assess quickly if students are properly assimilating information. It is for this reason that so many colleges and universities focused on adult learners distribute Angelo and Cross's (1993) text and will undoubtedly continue to do so.

ROLE OF CLASSROOM ASSESSMENT TECHNIQUES

As previously stated, the goal of formative assessment is to evaluate a student's understanding of course objectives at a particular stage of the learning process. This can help to determine if the instructor needs to adjust or elaborate and can help to inform the student of their actual level of accomplishment in the class to that point as opposed to their perception of this. Ideally, a formative assessment would test both knowledge and, in the author's opinions, an ability to analyze and synthesize information conveyed in the coursework. Angelo and Cross (1993) stated that formative assessment must be implemented early and often, providing students with ungraded feedback so that both student and instructor are able to assess the quality of learning. This then allows the student to provide feedback for improving instruction. Angelo and Cross (1993) provided three specific guidelines for implementing CATs: (a) planning, (b) implementing, and (c) responding.

The planning step requires the instructor to reflect on what might be gained by implementing formative CATs in the classroom (Angelo & Cross, 1993). How such activities will improve class interaction and student learning is an important question to consider. Selected CATs must align with the goals of the course or the learning objectives. Moreover, the selected CATs should provide sufficient feedback to facilitate improvements in learning and be easily implemented in class. Finally, we suggest the CAT must work well with the instructor's individual teaching style.

The implementing step involves explaining to the students the tasks that are used in the assessment with a clear explanation of the activity as well as its purpose. This ensures proper execution of the assessment and makes it more likely that the students will engage in the tasks. Generally, formative CATs ought not to be viewed in an adversarial light, as a summative assessment might be. It is important that the assessment be presented as a means of communication rather than a punitive or final evaluation. If course credit is offered, it is generally for participation credit rather than a large number of points. The assessment is typically completed immediately, then collected and analyzed within the class period (Angelo & Cross, 1993).

The third step recommended by Angelo and Cross (1993), which is responding, relates to providing students with feedback on the assessment. Feedback should illuminate what the instructor has learned from administering the CAT and what they intend to do with that information. For example, if an assessment indicates a high level of learning on a task, the instructor might express approval and continue down the established path at an enhanced pace. If the assessment indicates a lack of understanding of course concepts, the instructor might reteach those concepts from a different perspective. Finally, if the assessment indicates a failure in prerequisite preparation, the instructor might suggest a student seek aid, such as supplementary instruction in basic mathematics.

Angelo and Cross (1993) offered a list of sample CATs that range in complexity, time to implement, preparation, and time needed to analyze results. Examples include:

- **The Muddiest Point:** a writing assignment where students are asked to note the point or part of the lesson they are the least clear on or are having significant trouble understanding.
- **The Minute Paper:** a 60-second writing assignment where students are asked to write about the most important topic discussed in class that day.

These CATs are among the most common in the literature; however, the first, the Muddiest Point, has been identified as problematic in some studies (Simpson-Beck, 2011). Therefore, it is the authors' intent to suggest a framework for the analysis and construction of formative assessments, which might more effectively deliver on the promise of these techniques.

CLASSROOM ASSESSMENT TECHNIQUES EXAMINED

We believe that traditional formative CATs may suffer from a lack of emphasis on analysis and synthesis. This may take the form of not testing for the synthesis of information or testing for it in an inappropriate manner. For example, Angelo and Cross-'s (1993) Muddiest Point exercise does not evaluate students' understanding of the data but

merely their perception of their own understanding of the course material. The one major quantitative study on the effectiveness of using the Muddiest Point indicated no significant difference between groups who were taught using this CAT and the control groups who were not exposed to it (Simpson-Beck, 2011).

The study conducted by Simpson-Beck (2011) is persuasive in its assertion that there were no accrued benefits for the learners; however, there are some concerns with the study. The description of the methodology used in this study does not describe the subject matter under investigation and this precludes any analysis of issues surrounding the selection of the Muddiest Point as a formative CAT appropriate to the learning objectives. Additionally, we suspect that much of the weakness of the Muddiest Point is in its reliance on student perception of learning. For example, it is possible for a student to feel as though they understood a concept and later be surprised when they fail a final exam. In this case, the Muddiest Point may not have been the most useful CAT to select. We propose that if the formative CAT is not properly aligned with the learning objectives, then there will be no demonstrable benefit from their application.

A student may not be in a position to know what they do not yet know. Angelo and Cross (1993) suggest that the Muddiest Point be used for initial evaluation of new material. However, this is still a perception issue. For instance, a student whose attention has slipped in a history course may not understand that they do not recall a critical event that is necessary for the understanding of later events. An example might be the fact that the invention of the cotton gin by Eli Whitney substantially increased the monetary value of slave labor on cotton plantations. The student might dwell on other facts when asked after a lecture about the Muddiest Point, but when required to write an essay exam on why slavery might have become more important to cotton-producing states after 1794, they would undoubtedly fail the summative assessment. It simply never occurred to them that they had missed something important.

In summary, the authors of this article propose that effective formative CATs must include methods to test if learning objectives have been synthesized, applied, and ultimately absorbed by students. The following example illustrates this point a primary school teacher believes the day has been spent teaching the concept of addition but wishes to test this. The teacher discusses the process of addition and uses the examples of 1+1=2 and 2+2=4. The teacher then conducts a quick verbal CAT asking students what they have learned so far. The students reply, "We have learned addition!" Then when asked to elaborate, they reply, "We learned that 1+1=2 and 2+2=4." Have the students learned this concept or have they simply parroted back the two data points given in class?

While this is an exceptionally simple example, it illustrates two separate concepts, one very intuitive and one perhaps more obscure. From the example, it is certain that the students have learned something. At the very least, they have memorized the data that 1+1=2 and that 2+2=4. This is useful data, but it is far more useful if they have also absorbed the concept of addition and are able to successfully apply it.

In order to assess students' understanding of the concepts of addition rather than simply memorization, the instructor might utilize a simple CAT and ask another question such as, "If you have learned addition, what does 3+3 equal?" This question follows the same simple progression that students have been exposed to, and the instructor might be heartened to hear the answer "six"; however, this may still not be the best assessment. It is important to be mindful that sometimes students are able to provide the correct answers for the wrong reasons. In this case, it is possible that we have inadvertently taught the concept of doubling, which is in fact multiplication by two. In all of our examples, a whole number is added to itself. It is not too much of a stretch to imagine a precocious first grader internalizing this concept rather than the principles of addition that we are attempting to teach. Thus, a CAT related to the addition of 3+3 may not be robust enough to provide sufficient assessment.

To assess with more certainty whether the CAT has achieved our learning objectives, we might vary the question by asking, "If you learned addition, then what does 1+2 equal?" We might also increase the complexity by asking, "What does 7+5 equal?" This new CAT requires students to apply the learning rather than remember an answer already given in class, thus testing application of the proper concept. This will provide an immediate and accurate assessment as to whether or not the principle of addi-

tion has been properly communicated. Additionally, the three questions, 3+3, 1+2, and 7+5, may be scaffolded into a more rigorous CAT exercise. Used together, these steps will assess not only if the correct principle has been taught but may also uncover if another undesired concept has been taught instead. Finally, by creating a ladder of increased complexity, it is possible to gauge the level of understanding of the students and ascertain at what point they begin to feel intimidated by the assessment.

When adding rigor to the assessment, two key guidelines should be observed:

1. The instructor must attempt to use the appropriate amount of rigor based on topic progression but not so much as to intimidate the student or prevent them from properly performing the assessment.
2. Instructors should be mindful of incorporating rigor that may require yet to be obtained or outside knowledge.

When asking students to demonstrate applied knowledge, it is important to consider the appropriate level of rigor and ramp up complexity accordingly. For example, we might hesitate to give our imagined first graders the problem 574+2. While quite simple for most, requiring the students to simply take 574 and count off two more numbers, they might look at the problem and shut down due to the emotional impact of seeing this kind of question, which on the surface looks like a far more difficult problem. This is made more problematic if the assessment includes work that requires prerequisite knowledge that may not have been learned properly or which is perceived as an inherently difficult concept by students.

A simple example of this would be a test of an early primary school student's ability to perform subtraction after the student learned the concept of addition the week before. In this case, we assume that the student has learned to add three two-digit numbers together. Given these prerequisite skills, we give them the problem 52+27-2. This is probably not a good assessment of their capacity to subtract. It adds a great deal of complexity and if the child did not properly learn addition, they may fail the assessment for that reason before ever getting to the subtraction portion of the assessment.

Our example here is basic, but it demonstrates the concept of using simple CATs in the form of questions of increasing complexity to gauge learning and student confidence. While this may seem rudimentary, the process carries through to higher learning where the same principles apply. As an example, the automated finance lab assessments of one major university tests knowledge of the concept of weighted average cost of capital (WACC). In essence, this is an equation that says if a company's cost of money raised by stocks is 20% per dollar, and the cost to raise the same amount with bonds is 10% per dollar, if the company raises 50% of the money by stock and 50% by bonds, the weighted average cost of capital is 15% overall. The concept is simple and easily accessible to any student who can imagine going to the store to buy ten pairs of socks and being forced to buy five pairs of bargain socks and five pairs of premium socks because there is not enough of one or the other on the shelf.

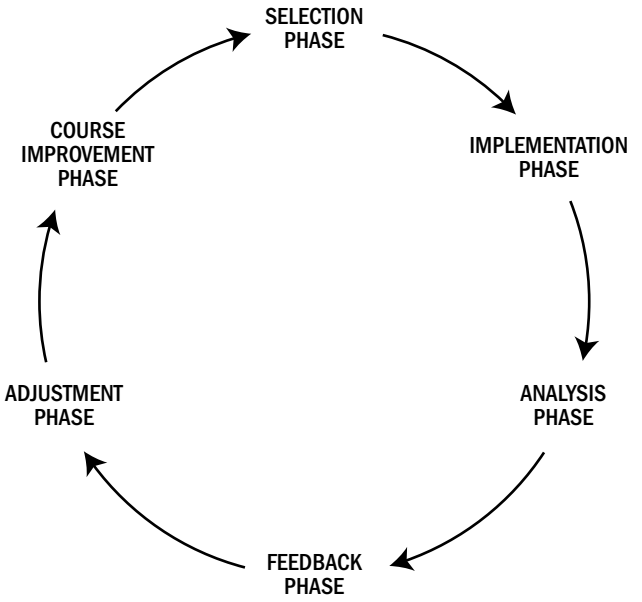
Unfortunately, this formative CAT assessment, in an attempt to compel the student to demonstrate information from a prerequisite class, requires the students to first calculate the rate of the bond, applying discounts and taking into consideration the tax deductibility of interest. This is a far more complex mathematical calculation than the basic WACC formula and is far more intimidating. It also assumes that students know how to calculate bond value for a discounted or premium bond on the primary market. As a formative assessment, this question would fail our test unless preceded by at least two other questions on the same topic with building complexity. An example of this might be one in which the cost of bonds is simply given, followed by one in which the cost of bonds is offered, along with a question about the tax deductibility of bonds and given tax rate. In this way, the assessment would deliver more complete data on student learning and it would indicate whether remedial education is necessary.

COMPONENTS OF AN EFFECTIVE FORMATIVE CLASSROOM ASSESSMENT TECHNIQUE

Angelo and Cross (1993) recommended three simple steps to implementing CATs; we would respectfully suggest additional guidelines to help appropriately apply CATS:

Diagrammatically, the proposed cycle created is as follows:

Figure 1: Cycle of CATs.



The cycle created may proceed quickly, particularly in a non-traditional setting. If in a traditional university there are effectively two classroom generations per year or three if including summer school, then an instructor teaching a given course once a semester will teach it perhaps one to three times a year. In an adult learning or non-traditional university, one might teach a popular or required course six times a year or more (Ross-Gordon, 2011). This reduced generation time allows for the improvement of a course offering either formally through the revision of curriculum or informally through independent modification by an individual instructor. It is our belief that attempts to formalize this selective revision process may be worthy of further study.

Table 1 (below) provides more detail on each phase and its sub-components.

Table 1: Components, key features, and importance of CATs.

SELECTION PHASE

STEP 1: EVALUATE LEARNING OBJECTIVES
Allows us to ensure that a CAT will be aligned with learning objectives.

This provides focus for the activity and ensures that it aligns goals of the course. Often activities may be exciting and there is a temptation to do too much or create a classroom "effect" outside of the goal of assessment. Noting what is being assessed first will help an instructor keep focus.

STEP 2: SELECT ASSESSMENT (OR CONSTRUCT A NEW CAT)
Determines the appropriate method of interaction.

Assessments may be of many different types, delivered in many ways such as verbal, written, role-play, and so forth. Determining the appropriate method that both performs the assessment and fits the instructor's style, and the student's disposition is essential for an effective formative CAT.

STEP 3: EMBED SYNTHESIS
Ensure an accurate formative assessment of the true learning objectives.

Exercises should be rigorous enough to determine the degree of synthesis of the information. Have the rudimentary elements of the data been absorbed? Is the student merely parroting back what they have been told? Can the information be applied dynamically to multiple scenarios? Complex topics must be assessed in layers to determine what has been learned and what must be re-communicated. This is the heart of the design process.

STEP 4: VALIDATE ASSESSMENT FOR TIME AND EASE OF USE
Prevents a poorly planned CAT from overrunning class time and impacting instructor credibility.

Activities must be smoothly integrated into the classroom experience; they should be a smooth part of class and most importantly should be concise and not overrun the rest of the day's learning activities.

The instructor might mark time, give the students instructions on how to perform the assessment to a mirror, perform the assessment themselves, considering if the students will have necessary supplies with them or if they would need to be provided. This will yield the time necessary to deliver the assessment.

If the assessment is to be scored immediately in class, this should be done and the results multiplied by the expected number of students, with thought being given as to how scores will be recorded and interpreted. This is particularly critical as it is possible to "lose the class" if one spends too much time in such evaluation.

IMPLEMENTATION PHASE

STEP 5: IMPLEMENT ASSESSMENT
Applies the technique in the classroom.

Techniques must be executed appropriately in the classroom to be effective and provide the appropriate experience for students to meet goals.

ANALYSIS PHASE

STEP 6: ANALYZE RESULTS
Provides the data to assess performance of both students and instructor.

Responses must be analyzed to understand the degree of absorption of the material and the student's ability to use and apply the material. This is used to provide student feedback.

Responses must also be analyzed to understand any deficiencies in classroom presentation which may be present. Often this will involve comparing these assessments to ones previously performed by the instructor in class with other students.

FEEDBACK PHASE

STEP 7: PROVIDE FEEDBACK
Provides information to the student and instructor on performance.

This is crucial for the students to understand how they have absorbed the material. If the students have done well, this provides affirmation and the confidence in applying the material or skills presented. If not, it provides data about how much progress they have obtained. This is vital as it is possible for students to believe they understand a concept when they do not or fear they do not understand a concept when they do.

While formative CATs will generally be performed by the instructor themselves, if they are being performed by a third party for the instructors benefit, the feedback would be presented along with analysis.

ADJUSTMENT PHASE

STEP 8: ADJUST PRESENTATION / PROVIDE REMEDIATION

Based on the analysis and feedback the instructor can transition to a remediation stage if necessary, elaborating on the topic, providing additional information, or corrective action. This is a tactical adjustment meant to improve the learning and experience

COURSEWORK IMPROVEMENT PHASE

STEP 9: RECONSIDER COURSEWORK/CURRICULUM

Improves the overall quality of the education product for the next class of students.

When a formative assessment surfaces issues across several populations of students, particularly when the issues affect the proper communication of the course learning objectives, it calls into question the strength of the curriculum. It is therefore important for the instructor to use the assessment to identify weak-points in their course curriculum and/or their own presentation style and correct these where possible. This is a strategic adjustment to curriculum and or instructor skills meant to benefit future students

The final step of the above framework is one that is not always considered. As previously noted, the current trajectory of non-traditional adult education necessitates compressed class timelines with course delivery over a few weeks and perhaps as few as five instances of classroom contact (Ross-Gordon, 2011). This makes it even more critical to know when one is failing to effectively impart learning objectives, generating the need, in our opinion, for robust CATs. Consequently, the suggested framework not only allows for the proper construction and evaluation of formative CATs but also creates a process improvement cycle.

Steps one through four of the framework comprise the selection cycle of a formative assessment. These steps are used to evaluate a new or established assessment technique. Please refer to Appendix A for an example of how the framework may be used to evaluate the technique the "Muddiest Point or to revise the One Minute Paper assessment.

CONCLUSIONS AND RECOMMENDATIONS

There is no question the use of CATs, particularly of the formative variety, is embraced by many in academia (Carduner, 2002; Choinski & Emanuel, 2006; Cottell & Harwood, 1998; Eisenbach et al., 1998; Goldstein, 2007; Simpson-Beck, 2011; Suskie, 2004; Walker, 2012). It is our belief that the lack of correlation-demonstrated research studies may be due to a basic failure to select the correct formative CAT to support the learning objectives, along with lack of rigor embedded in exercises for compressed classrooms and student-instructor interactions. In an attempt to mitigate this, we have proposed a framework to assist in the evaluation of formative assessments. This framework works in conjunction with a lifecycle approach to improvement, which might be used to drive the improvement of either curriculum (formally) or the presentation of curriculum (informally) by instructors who adopt our concepts.

It is our hope that the framework presented may act as a schematic tool or checklist to increase the quality of formative CATs in classrooms. Simply evaluate student learning is not sufficient; one must turn the assessment inward to examine instructional style, presentation, and curriculum (Angelo & Cross, 1993). Our hope is that adding rigor to CATs will increase confidence in the utility of formative CATs, which could be used to improve students' experiences and reduce concerns in the classroom for both the teacher and the learner. This concept must be examined through further research. We invite researchers to test these assertions and the framework presented.

References

Adams, P. (2004). Classroom assessment and social welfare policy: Addressing challenges to teaching and learning. *Journal of Social Work Education, 40*(1), 121-142. Retrieved from http://www.jstor.org/stable/23043983

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers.* San Francisco, CA: Jossey-Bass.

Carduner, J. (2002). Using classroom assessment techniques to improve foreign language composition courses. *Foreign Language Annals, 35*(5), 543-53. doi: 10.1111/j.1944-9720.2002.tb02722.x

Choinski, E., & Emanuel, M. (2006). The one-minute paper and the one-hour class: Outcomes assessment for one-shot library instruction. *Reference Services Review, 34*(1), 148-155. doi: 10.1108/00907320610648824

Cottell, P. G., Jr., & Harwood, E. M. (1998). Using classroom assessment techniques to improve student learning in accounting classes. *Issues in Accounting Education, 13*(3), 551-564. Retrieved from http://aaahq.org/pubs/issues.htm

Cross, P. K. (1987). The adventures of education in wonderland: Implementing educational reform. *Phi Delta Kappan, 68*(7), 496-592. Retrieved from http://eric.ed.gov/?id=EJ349185

Dewey, J. (1933). *How we think.* Lexington, MA: Heath.

Eisenbach, R., Golich, V., & Curry, R. (1998). Classroom assessment across the disciplines. *New Directions for Teaching & Learning, 78*(59). doi: 10.1002/tl.7506

Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education, 36*(1), 51-62. doi: 10.1080/02602930903197883

Goldstein, G. S. (2007). Using classroom assessment techniques in an introductory statistics class. *College Teaching, 55*(2), 77-82. doi:10.3200/CTCH.55.2.77-82

Hussar, W., & Bailey, T. (2009). *Projections of education statistics to 2018. Table 3: Actual and middle alternative projected numbers for college enrollment rates, by sex, attendance status, and age: Fall 2007, 2013, and 2018.* Retrieved from http://nces.ed.gov/pubs2009/2009062.pdf

Jarzebowski, A., Palermo, J., & van de Berg, R. (2012). When feedback is not enough: The impact of regulatory fit on motivation after positive feedback. *International Coaching Psychology Review, 7*(1), 14-32.

Melland, H. I, & Volden, C. M. (1998). Classroom assessment: Linking teaching and learning. *Journal of Nursing Education, 37*(6), 275-277.

Mosteller, F. (1989). The 'muddiest point in the lecture' as a feedback device. *Journal of the Harvard Danforth Center on Teaching and Learning, 3*, 10-21.

National Center for Education Statistics. (2009). *Digest of educational statistics 2009. Table 192. Total fall enrollment in degree-granting institutions by control and type of institution, age, and attendance status of student: 2007.* Retrieved from http://nces.ed.gov/programs/digest/d09/tables/dt09_192.asp?referrer=list

Ross-Gordon, J. M. (2011). Research on adult learners: Supporting the needs of a student population that is no longer nontraditional. *Peer Review, 13*(1), 26-29

Simpson-Beck, V. (2011). Assessing classroom assessment techniques. *Active Learning in Higher Education, 12*(2), 125-132. doi: 10.1177/1469787411402482

Steward, B. L., Mickelson, S. K., & Brumm, T. J. (2004). Formative and summative assessment techniques for continuous agricultural technology classroom improvement. *NACTA Journal, 48*(2), 33-41. Retrieved from http://www.nactateachers.org/journal.html

Suskie, L. (2004). *Assessing student learning: A common sense guide.* San Francisco, CA: Anker.

Walker, D. (2012). Classroom assessment techniques: An assessment and student evaluation method. *Creative Education, 3*, 903-907. doi: 10.4236/ce.2012.326136

Wlodkowski, R. (2003). Accelerated learning in colleges and universities. *New Directions for Adult and Continuing Education, 97*, 5-15. doi: 10.1002/ace.84

Appendix A

Included in this appendix are examples that detail the use of the CAT Cycle Framework to evaluate and modify proposed CATs prior to adoption.

ONE MINUTE PAPER – REVISION OF AN ESTABLISHED FORMATIVE ASSESSMENT

In this example, students write for one minute about the most important thing they learned that day. We then take an assessment, which is not well suited to our purpose, and revise it to fit our learning objectives.

Step 1: Evaluate Learning Objectives – Assume that this is a third year history class with 15 history majors studying the evolution of chattel slavery in the United States from approximately 1619 to the end of the American Civil War in 1865. The goal is to assess if the students comprehend the economic factors associated with the introduction of the cotton gin from its patenting in 1794 and to determine if they can apply that knowledge to an essay question given in the summative assessment. This assessment includes a five question final exam wherein one question will be, "What technology introduced in the late 18th century had an important economic impact in the development of chattel slavery in the Southern United States? Please discuss this technology and how it affected the economics of the South leading up to the Civil War." As we are able to define the objective, we move to Step 2.

Step 2: Select Assessment (or Construct a New CAT) – We select the One Minute Paper as described by Angelo and Cross (1993) and move on to Step 3.

Step 3: Embed Synthesis – We note that there are weaknesses in our selected assessment technique. The first issue is that students select what they consider to be the most important topic of the day for their paper. We might reason that other topics are presented on that day. As such, a student may not select the chosen topic of the cotton gin, effectively reducing the value of our assessment to an acknowledgement that he or she did not understand the importance of this subject. Additionally, the topic is a rich one in which the instructor hopes to impart nuance and to connect to other topics already discussed; therefore, one minute is too little time to write a paper sufficient for an effective evaluation. Thus stymied in our attempt to use the One Minute Paper as originally described, we revise our assessment to become our own "Five Minute Topical Summary List." In this new assessment, the students are asked to write for five minutes summarizing the day's most important topic, which they are explicitly told is the economic implication of the introduction of the cotton gin. They are invited to list the most important factors that they remember regarding this topic but are told to write only one descriptive sentence for each point. This increases the time necessary for the assessment to 10-15 minutes. As this is acceptable, we move on to Step 4.

Step 4: Validate Assessment for Time and Ease of Use – At the end of class, each student is asked to write for five minutes on the selected topic, namely the economics of the cotton gin's introduction. We estimate that this will take 10-15 minutes. We then assess if we believe the students are likely to have writing instruments and paper and obtain extra of each to assist those who take notes exclusively electronically. We then practice the instructions and attempt to anticipate their questions until we are satisfied that this will not surpass the available time.

Given the time factors, we score the assessment between classes and provisionally accept The Five Minute Topical Summary List as an appropriate formative assessment, pending its successful use. Assuming that actionable data emerges, we add this assessment to our arsenal of formative CATs.

It should be noted that establishing a new formative CAT is essentially the same as the revision that we completed above. The instructor conceptualizes a classroom assessment technique and then implements the CAT.

THE MUDDIEST POINT – EVALUATION OF A "FAULTY" FORMATIVE ASSESSMENT

The Muddiest Point is a formative assessment in which students are asked to write the point or part of the lesson that is most unclear to them. To evaluate if this assessment is valid, one simply walks through the framework and performs an analysis at each level.

Step 1: Evaluate Learning Objectives – In this example, we assume that the learning objective is to test the comprehension of historical data

presented to a group of 150 freshmen undergraduates in a large recitation hall. The data we present is related, as it is from the same period in history, but as the summative test at the end of the course will be multiple choice rather than an essay, facts are more or less independent of each other. As we are able to identify the objective, we move to Step 2.

Step 2: Select Assessment (or Construct a New CAT) – In this case, we select the Muddiest Point and move to Step 3.

Step 3: Embed Synthesis – The Muddiest Point does not lend itself to synthesis in this case. The large number of independent facts makes it possible for each of the 150 students to ask about different parts of the week's lecture, necessitating either re-lecturing of the material the entire week or ignoring the data. Similarly, any student who fails to hear a key point of the lecture or fails to understand its implications would not necessarily know that they did not know this. It is therefore not possible to impart synthesis with this technique and another CAT is selected. It should be noted that our rejection of the Muddiest Point as an assessment does not constitute a rejection of the desire to understand our students' perceptions. In this case, it simply means that this technique does not align with our intended goals.

Using the framework's creation phase, this would normally mean we would end the process here; however, in the interest of clarity, we will complete the following steps as though the Muddiest Point had not been rejected.

Step 4: Validate Assessment for Time and Ease of Use – Each student is invited to note the part of the lesson that needs clarity. With 150 students, this will take approximately five to 15 minutes once a week. This is an appropriate expenditure of time, so we note the time required and examine how the CAT should be administered. Since the Muddiest Point requires the submission of a short written response, we see to it that sufficient paper and extra writing instruments are available to preclude a flurry of activity surrounding those students who have neither. We also practice the instructions to the class so that they can be disseminated in less than five minutes.

Finally, we evaluate how we score the assessment. In the above class of 150 students, analysis would take place prior to the next class session. If this is an acceptable timeframe, as it would likely be in a traditional university setting, we accept this. In an adult learning setting with a dozen students, we also accept this as it would be possible to both read the results and analyze them in a few minutes. With our planning done, we move to the implementation phase as described in the framework.

Author Biographies

Jason E. Thomas, MBA is the Chief Operating Officer (COO) for The Collective Group in Austin, TX, a Lead Faculty Area Chair for University of Phoenix, and an Adjunct Instructor for Grand Canyon University and Concordia University. His research interest focuses on best practices for Non-Traditional Accelerated Degree Programs and Consumer Behavior. Jason received his B.S. in General Studies from the University of the State of New York, a B.S. in Business Information Systems and a MBA from University of Phoenix, and has attained candidacy status as doctoral student pursuing a Ph.D. in Business with a concentration in Marketing from Northcentral University.

Philip E. Hornsey, MSA is a General Partner in the Kom Do Kwan Martial Arts Association, an Investor, and Adjunct Instructor for Grand Canyon University and the University of Phoenix in the area of Finance and Accounting. His research interests include Non-Traditional Education methodologies, Entrepreneurial Business Finance Strategies, and the relationship of Revenue Models to consumer buying decisions. Philip received his B.S. in Anthropology from the University of Missouri at Columbia, and an MSA from the University of Phoenix, and is pursuing a Ph.D. in Business with a concentration in Finance from Northcentral University.