A CONCEPTUAL FRAMEWORK FOR FORMATIVE ASSESSMENT IN LARGE-ENROLLMENT INTRODUCTORY STATISTICS

KIMBERLEIGH FELIX HADFIELD
Utah State University
k.hadfield@usu.edu

ABSTRACT

Implementing formative assessment in large-enrollment statistics courses is recommended by mathematics and statistics education communities. Yet research maintains that large-enrollment courses employ few, if any, formative assessments, exacerbating negative attitudes toward statistics and low student achievement. This conceptual essay applies an andragogical approach to the theories of self-efficacy, self-regulation, and formative assessment. A literature review explicates the associations between formative assessment with feedback and reassessment with student attitudes and achievement. Resulting from the review, a conceptual framework is proposed that illustrates the relations between the elements of formative assessment cycles and student attitudes and achievement. The implications of this conceptual framework suggest a comprehensive transformation of assessment practices to provide pathways for student success in statistics courses.

Keywords: Statistics education research; Formative assessment; Large-enrollment; Feedback; Attitudes toward statistics; Student achievement; Reassessment

1. INTRODUCTION

… given the present strength of the evidence for the effectiveness of formative assessment, or assessment for learning, it is somehow surprising that the implementation of better classroom practises has not been more evident.

(Hopfenbeck, 2018, p. 548, paraphrasing D. Wiliam, 2018)

Assessment practices in undergraduate statistics education have been the subject of international concern for several decades. Organizations such as the National Centre for Research Methods (NCRM), the Economic and Social Research Council (ESRC), and the American Statistical Association (ASA) call for pedagogical and curriculum strategies that create opportunities and pathways for students to successfully complete their introductory statistics courses. Several publications have resulted from these efforts (see for example, Abell et al., 2018; ASA Revision Committee, 2016; Bidgood et al., 2010, Nind & Lewthwaite, 2015; Peck, 2019; Pfannkuch, 2018). Specifically, the research recommends formative assessments, especially for students who take statistical methods courses as part of a service course for other departments (Abell et al., 2018; Snelgar & Maguire, 2010). *Formative assessments* are those that inform the teacher about their teaching and instruction and inform the student of their understanding using feedback from the assessment process (Black & Wiliam, 1998; Cowie & Bell, 1999; Ghaicha, 2016; Harlen, 2012; Shute, 2008). In contrast, evidence obtained from *summative assessments* provides judgments about student achievement with no cycle of feeding back (Harlen, 2012).

Additionally, the ASA and the Mathematical Association of America (MAA) delineate several recommendations for improving assessments in undergraduate introductory statistics courses. One recommendation by the ASA in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report* (2016) states that assessments should be used both formatively to improve learning and to summatively evaluate learning in a continuous process. Furthermore, the *GAISE* report stresses the need for assessments to provide feedback to students regarding their learning through utilizing frequent low or no-stakes formative assessments. Likewise, the *MAA Instructional Practices*

Guide (Abell et al., 2018) discusses how instructors and course designers can use formative assessments to improve curriculum and instruction in undergraduate quantitative courses by providing vignettes with examples of different formative assessment strategies. The guide emphasizes the need for evidence-based assessment practices in large-enrollment courses to enhance various cognitive and performance-based student outcomes. Additionally, several principles for assessments outlined in the *MAA Instructional Practices Guide* echo the *GAISE* report. Both reports recommend courses integrate assessments, not as stand-alone events, but rather as a "continuous cycle" of assessment throughout the course (Abell et al., 2018; ASA Revision Committee, 2016).

Despite these calls for improved assessment practices, large-enrollment undergraduate introductory statistics courses have been slow to implement formative assessment in due to the volume of students and the time-intensive nature of the assessment process (Garfield et al., 2002; Garfield et al., 2011). Cash et al. (2017) conducted a large study of university courses and found that students noticed a stark difference in the types of assessments used in large- and small-enrollment classes. Unsurprisingly, the findings revealed that assessments used in large-enrollment courses were less frequent and did not often vary in their design. In fact, summative, high-stakes assessments accounted for more than 95% of the assessments in large-enrollment classes (Cash et al., 2017). In addition, students reported only one to two exams and one final exam in these large courses, a glaring departure from the recommendations to utilize a continuous cycle of formative assessment with ongoing student and teacher feedback (Abell et al., 2018; ASA Revision Committee, 2016). Therefore, this conceptual essay focuses on examining the formative assessment literature to consider the impact and implementation of formative assessment as an embedded, continuous cycle with feedback and reassessment in large-enrollment introductory statistics curriculum.

1.1. LARGE-ENROLLMENT STATISTICS COURSES

Introductory statistics is quickly becoming the course that satisfies the quantitative literacy requirements in higher education, recently replacing college algebra (Hoang et al., 2017). The growth of undergraduate student enrollment in introductory statistics courses brings a diversity of student interests and mathematical background knowledge (ASA Revision Committee, 2016; Blair et al., 2018). These students often do not major in science, technology, engineering, or mathematics (STEM) and are experiencing the course as the only statistics course in their program of study. Fong et al. (2015) found that students enrolled in introductory statistics lack mathematical knowledge, potentially due to this growth in and diversity of students. This deficiency in mathematical knowledge can result in a financial burden when students require multiple attempts at their introductory statistics class to pass successfully, creating a "bottleneck" by slowing the path to their further studies (Complete College America, 2012; Fong et al., 2015). To further exacerbate these problems, introductory statistics courses tend to be large-enrollment courses in large universities, which do not allow for individualized academic attention (Blair et al., 2018; Cash et al., 2017).

Introductory statistics is arguably one of the most critical quantitative courses in a non-STEM student's university experience, due to the "societal need for statistically literate citizens and a more statistically trained workforce" (Zieffler, et al, 2018, p. 50). The non-STEM students only experience with statistical thinking and application is through their introductory course (Aliaga et al., 2005; Garfield et al., 2010).The 21st century requires students to navigate statistical information critically in a data-driven world (Rumsey, 2002; Tishkovskaya & Lancaster, 2012). As universities address both the increased enrollments and increased demand for introductory statistics courses, large-enrollment courses become the answer. Thus, introductory statistics courses will continue to amass many students with varied interests, goals, and mathematical preparedness (ASA Revision Committee, 2016; Garfield et al., 2002; Hoang et al., 2017; Peck, 2019). Therefore, the need to increase formative assessment, import the assessment recommendations, and support student success in large-enrollment statistics courses becomes even more paramount.

1.2. TWO IMPORTANT STUDENT OUTCOMES: ATTITUDES AND ACHIEVEMENT

The challenges faced by students to navigate large-enrollment statistics sections impact their experiences, affect whether they value statistics, and influence their attitudes toward statistics throughout their adult lives (Ramirez et al., 2012; Tichkovskaya & Lancaster, 2012). Evidence suggests that students who were nonmajors experienced greater statistics anxieties and tended to avoid courses in statistics, which negatively affected their attitudes toward statistics and their achievement (Chew & Dillon, 2014; Chiesi & Primi, 2010; Lavidas et al., 2020; Onwuegbuzie & Wilson, 2003; Williams, 2015). These negative attitudes toward statistics "stuck" with students long after they experienced their introductory course, affecting their motivation and appreciation for statistical literacy (Ramirez et al., 2012; Schau, 2003). Additionally, students' negative attitudes were associated with decreased achievement (Chiesi & Primi, 2010; Emmioğlu & Capa-Aydin, 2012). These anxieties and negative attitudes, linked to testing and assessment performance in students' introductory statistics courses, were due in part to their lack of mathematical understanding and preparation in prior math courses (Chiesi & Primi, 2010; Malik, 2015; Onwuegbuzie & Wilson, 2003). Although large-enrollment classes answer the need to support increased enrollment by many non-STEM students, they aggravate anxieties and negative attitudes, as large-enrollment courses tend to use high-stakes examinations with minimal if any, feedback (Cash et al., 2017).

Conversely, teachers who used innovative assessments influenced student attitudes, persistence, and achievement (Abell et al., 2018; ASA Revision Committee, 2016). Specifically, students' positive attitudes toward statistics were associated with greater motivation, higher achievement, and improved outcomes in statistics (Chiesi & Primi, 2010; Ramirez et al., 2012). Furthermore, when classrooms used formative assessment with feedback focused on learning rather than performance, students experienced a change in their mindset and relationship with their learning (Boaler & Confer, 2017). Finally, recent research has asserted that a "comprehensive approach to designing a successful statistics pathway" is needed to provide support structures in introductory statistics for underprepared students to "effectively complete their college-level statistics course" (Peck, 2019, p. 35). Thus, this conceptual essay posits a comprehensive reform of assessment practices to align with these efforts in large-enrollment courses.

1.3. RECOMMENDATIONS FOR ASSESSMENT PRACTICES

Recommendations for improving assessments are further detailed in reports from both the ASA and MAA, and books such as *Assessment Methods in Statistical Education: An International Perspective* (2010). The *GAISE College Report* (2016) explicitly addresses how instructors should employ assessments in introductory statistics courses. Specifically, to "[u]se assessments to improve and evaluate student learning" (p. 3), the *GAISE* authors stress that (a) students should receive timely feedback throughout the course, (b) assessments should align with learning outcomes, and (c) instructors should maximize the use of varying types of formative assessments in addition to summative examinations.

The *MAA Instructional Practices Guide* (2018) provides an assessment framework, gleaned from research explaining the benefits of evidence-based assessment practices, ASA recommendations (ASA Revision Committee, 2016), and the National Council of Teachers of Mathematics (NCTM) standards (NCTM, 2000). The assessment framework was based on the following six principles proposed by Steen (1999, p. 2–3):

1. Assessment is not a single event but a continuous cycle.
2. Assessment must be an open process.
3. Assessment must promote valid inferences.
4. Assessment that matters should always employ multiple measures of performance.
5. Assessment should measure what is worth learning, not just what is easy to measure.
6. Assessment should support every student's opportunity to learn important mathematics.

Several of the MAA's principles for assessments overlap with the *GAISE* recommendations, specifically that assessments are not singular events, but rather, when feedback cycles back to the teacher and student throughout the course, assessment becomes a cycle (Steen, 1999). Additionally, the

Instructional Practices Guide stresses assessment tasks must be linked to learning goals or objectives for the course (Abell et al., 2018). Furthermore, the *Instructional Practices Guide* devotes a section to assessment in large-enrollment courses, suggesting online response systems, online homework systems, and the use of technology to support instructors to provide timely feedback to students (Abell et al., 2018).

Davies and Marriott (2010) provided several recommendations for achieving the balance of formative and summative assessment, which the MAA recommendations (2018) punctuate. Specifically, to encourage deeper understanding, instructors must utilize formative assessments to assess those learning outcomes valued in the course. Learning outcomes that emphasize statistical thinking, statistical reasoning, and statistical literacy are recommended (Garfield et al., 2011; Schield, 2010). Finally, several authors encouraged large-enrollment courses to utilize individualized assessment practices through test banks and computer-based tasks to allow timely feedback and randomization of questions (Hunt, 2010; Simonite & Targett, 2010; Spencer, 2010; Stirling, 2010). Further research must address how instructors can import these recommendations more fully and use technology to create meaningful assessments in large-enrollment courses.

With the focus on contributing to the field of statistics education and heeding the call to create successful student pathways, the purpose of this conceptual essay is to examine formative assessment concepts and their relationships to student attitudes and achievement in the statistics education literature. The next section discusses the theoretical framework for formative assessment in higher education. Then, following the methods of the literature review, the relationships between the elements of formative assessment and student attitudes and achievement are explored. These results generate the conceptual framework for formative assessment in large-enrollment introductory statistics courses. The final sections discuss the constructs of the conceptual framework to provide foci for future research and implications for instructors, students, and curriculum creators. Finally, the paper concludes by posing questions for future research to analyze formative assessment cycles and their effects in practice.

2. DEVELOPING A THEORETICAL FRAMEWORK TO EXAMINE FORMATIVE ASSESSMENT IN LARGE-ENROLLMENT INTRODUCTORY STATISTICS

... learning is an increase, brought about by experience, in the capacities of an organism to react in valued ways in response to stimuli.

(Black & Wiliam, 2009, p. 10)

2.1. THEORIES OF FORMATIVE ASSESSMENT IN HIGHER EDUCATION

Andragogy. Andragogy is a perspective explaining how adult learners engage in their learning environments (Knowles, 1978). Viewing learning theories through the lens of andragogy provides unique insight and understanding as to how adult learners learn and process information. The collective research of Knowles (1978), Lindeman (1926), and Merriam (2001) posits five central tenets:

1. Adult learners are motivated and desire to learn.
2. Adult learners want to apply information to life situations directly.
3. Adults' life experiences provide a valuable resource to their learning.
4. Adult learners are self-directed.
5. As age increases, differences across individuals are vast and contextualized.

For nearly a century, these tenets have informed adult education worldwide (Merriam, 2001). As stated by Knowles (1978), "Adult education is an attempt to discover a new method and create a new incentive for learning" (p. 11). Additionally, andragogy is learner-centered, with educators given the charge to "involve learners in as many aspects of their education as possible" (Houle, 1996, p. 30), such that the educational climate fosters positive learning environments for adult learners.

The theoretical framework for this conceptual essay in Figure 1 also includes formative assessment theory (Black & Wiliam, 2009; Harlan, 2012). By taking up an andragogical lens, we can draw from these central tenets to leverage adult learners' engagement in their learning through assessment design. It is important to note that each of these theoretical discussions could span volumes of text (and

academic journals abound regarding these theories). For the purposes of this conceptual essay, a general discussion of these theories provides the foundation for the resulting conceptual framework.
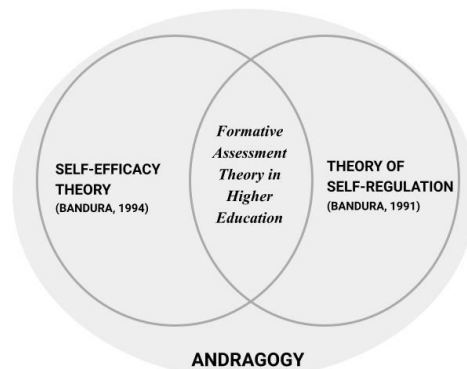

Figure 1. A theoretical framework for formative assessment in higher education

Andragogy and self-regulation theory. Schunk and Zimmerman (1997) considered self-regulation to be a skill acquired through several active processes. As a self-directed agent, the adult learner pursues tasks and goals with actions, such as planning, strategizing, and taking risks to pursue and solve problems (Martin, 2010). Martin suggested that classrooms provide "self-regulatory activities as embedded and entwined with their ongoing, active engagement in curricular tasks" (p. 143). Adult learners use their rich life experiences in the classroom; they actively engage in the problem-solving process as they adapt to the learning environments, acting as agents in their educational outcomes. Bandura defined agency as "the power to originate actions for given purposes" (Bandura, as cited in Martin, 2010, p. 138). One agentic mode considered within the lens of andragogy is that of *self* or *personal agency*: the self-regulative abilities of an adult learner are an active agent unto themselves—to self-motivate, self-guide, and self-correct (Bandura, 2001). These aspects of personal agency describe adults as self-directed agents.

Adult learners, as self-directed agents, set standards of performance regarding their learning goals and affect change through their environment. Furthermore, the adult learner creates paths of action to meet learning goals and must change, adapt, and act within these paths to accomplish those goals (Bandura, 2001). The self-regulation of motivation is bi-directionally associated with goal setting and proximity; goals in the near future engender greater motivation and self-regulatory action than goals in the distant future (Bandura, 2001).

Andragogy and self-efficacy theory. Self-efficacy is an essential factor in the self-regulation of an adult learner. *Self-efficacy* theory regards the beliefs which are "the foundation of human agency. Unless people believe they can produce the desired results and forestall detrimental ones by their actions, they have little incentive to act or to persevere in the face of difficulties" (Bandura, 2001, p. 10). Adult learners are internally motivated and more self-directed in their learning (Knowles, 1978; Merriam, 2001; Sosibo, 2019). In a sense, the adult learner is learning because they *want* to. The actions of the adult learner are rooted in their core beliefs that their efforts will produce the desired outcomes in their education (Bandura, 2001). Thus, if adult learners believe in their ability to achieve their goals, their actions stem from those beliefs (Mangels et al., 2006). Additionally, Good et al. (2008) suggested that students' learning beliefs are attributable to their performance in academic tasks.

2.2. FORMATIVE ASSESSMENT THEORY IN HIGHER EDUCATION

Formative assessment theory, in the intersection in Figure 1, creates the essential foundation for this research in tandem within the theories of self-efficacy and self-regulation, seen through an andrological lens. Assessment theory began as a study in elementary and secondary education (Black & Wiliam, 1998, 2009; Harlen, 2012). The application of formative assessment in tertiary education,

however, has brought undergraduate assessment research, regarding adult learners' needs, to light (Abell et al., 2018; ASA Revision Committee, 2016; Steen, 2006).

Situating formative assessment. "… assessment is operationally defined as part of the educational process where [faculty] instructors appraise student achievements by collecting, measuring, analyzing, synthesizing, and interpreting relevant information … under controlled conditions in relation to curricula objectives set for their levels" (Ghaicha, 2016, p. 212). This definition succinctly circumscribes all assessments. Thus, formative assessments are those assessments through which students can make progress towards the course's learning goals by receiving feedback on their learning and, therefore, take ownership of their education (Black & Wiliam, 2009).

Black et al. (2004) defined *Assessment for Learning* as "any assessment for which the first priority in its design and practice is to serve the purpose of promoting students' learning" (p. 10). The Assessment Reform Group (2002) defines *Assessment of Learning* as summative assessments whose only purpose is for grading and reporting. Additionally, Harlen (2012) described a gradient of assessments ranging from formative to summative assessments and broadly as Assessment *for* Learning to Assessment *of* Learning, respectively. These definitions situate formative assessment theory in Assessment *for* Learning.

Cultivating self-regulation through formative assessment. Adult learners are agents for their learning, they decide actions to take to achieve learning goals. Formative assessment allows adult learners to recognize the learning goals or outcomes expected in the course curriculum (Yorke, 2003). Choosing the tasks for formative assessments must be "justified in terms of the learning aims that they serve" (Black & Wiliam, 1998, p. 143). Thus, learning outcomes must be clearly defined and communicated for formative feedback to be utilized (Stiggins, 2002).

The learning environment poses significant importance as to whether the adult learner understands the learning goals. That learning environment is contingent upon the educational community, such as the educator, curriculum developer, or course designer, to generate (Ben-Zvi et al., 2018). Stiggins (2002) posited several strategies for educators to transform their assessment processes into Assessments *for* Learning. First, educators must appraise the adult learners of the course's learning goals, specifically to understand what the intended purposes or outcomes are. Thus, Stiggins suggested that instructors must know and communicate those learning outcomes in the syllabus before students engage in course material over those objectives. Additionally, Black and Wiliam (2009) described implementing Assessment *for* Learning with three critical processes for educators to implement in the course design: (1) clarify learning outcomes, (2) create tasks consistent with those outcomes to provide evidence of student learning, and (3) provide feedback (p. 8). Moreover, the tasks and formative assessments provide evidence of the course outcomes and allow the adult learner to self-regulate and act upon the feedback.

Self-assessment and feedback. Black and Wiliam (2009) drew on a partnership between teachers and students in their formative assessment framework: that both students and teachers are equally responsible for student learning. Teachers must provide feedback to the students and students must be resources for themselves by taking ownership of their learning by utilizing the feedback provided by the teachers. Black and Wiliam (1998) cautioned that formative assessment may not be productive without students being able and willing to self-assess to further their learning goals. Students' self-assessment is the "essential component of formative assessment" (p. 143). Black and Wiliam (1998) made a case for formative assessment in classroom practice through a meta-analysis of over nine years of empirical research. Among the outcomes of this analysis was the importance of feedback in the cycle of formative assessment. Feedback that aligned with learning outcomes helped to identify gaps in student knowledge and assisted students to see where they should improve. Feedback directly assists adult learners, as these students willingly engage in self-assessment and are more motivated and able to self-regulate their learning (Ghaicha, 2016).

Stiggins (2002) also underscored critical features of formative assessment by explaining that when students employ continuous self-assessment with consistent learning outcomes, they can better reflect on their knowledge developed over time. Through incorporating regular self-assessment as part of the assessment cycle, students watch themselves grow as learners (Black & Wiliam, 2009; Stiggins, 2002;

Wride, 2017). "Indeed, unless linked to an effective process of reflection, assessment can easily become what many faculty fear: a waste of time and effort" (Steen, 1999, p. 2). Steen (1999) maintained that assessment must be an open process using learning goals to inform both the teacher and the student of the student's progress. Therefore, feedback must provide vital information to guide students in their learning. Moreover, the onus is on the adult learner to use that feedback to improve their understanding (Sosibo, 2019).

Students in large-enrollment courses reported that communication with the professor lacks one-on-one, personable interaction (Cash et al., 2017), making the feedback students receive from these classes even more critical. Bayerlein (2014) investigated undergraduate students' perceptions of feedback: both the *timeliness* of feedback and *constructiveness*. Constructiveness concerned the use of automatically generated feedback versus handwritten feedback. Interestingly, undergraduate students found the automatically generated feedback to be substantially more constructive than the manually-written feedback. Simple, automatically generated feedback with non-judgmental wording aligned with learning outcomes is all that is needed to create constructive feedback (Stiggins, 2002).

Complex feedback is unnecessary for students to gain information about correctness and learning goals; instead, productive, concise feedback improves learning and student outcomes (Abell et al., 2018; Shute, 2008). Simply stated, "provide feedback that moves learning forward" (Black & Wiliam, 2009, p. 8). Two meta-analyses conducted on formative feedback and achievement found that providing feedback had a positive effect on student achievement (Hattie & Timperley, 2007; Wisniewski et al., 2020). Specifically, the effect of computer-generated feedback and corrective feedback benefited student learning ($d = 0.79$; Hattie & Timperley, 2007). Wisniewski et al. (2020) found that feedback is most effective when it includes several factors: timeliness, strategies, and self-regulation ($d = 0.48$). When feedback maintained the qualities of timeliness, concise elaboration on the learning outcomes, and computer delivery, it was associated with increased performance and learning (Shute, 2008). As the adult learner is self-directed, these meta-analyses provide important evidence that concise, automated feedback is sufficient to elicit self-regulation and affect student achievement.

Self-efficacy and reassessment. Reassessment allows multiple attempts at learning tasks to progress through a curriculum. Having utilized feedback by learning from mistakes, reassessment allows the adult student an additional opportunity to show evidence of learning. Because the focus of formative assessment is on the process of learning rather than performance, formative assessment fosters self-motivation, goal-orientation, and positive motivational beliefs of persistence and confidence (Duckworth et al., 2007; Dweck, 2008; Yin et al., 2008). In fact, Grant and Dweck (2003) found these beliefs gave students the ability to persevere despite an initial poor performance. Additionally, one of the outcomes of Assessment *for* Learning, which Boaler and Confer (2017) found in their work, is that students changed their perceptions of who they are as learners, thinkers, and problem-solvers. Thus, Assessment *for* Learning can change the landscape of mathematics education, equipping quantitative learners with self-efficacy—hope and belief in their educational pursuits (2017).

Steen (1999) asserted, "Assessment that matters should always employ multiple measures of performance" (p. 3). One way of creating multiple measures of performance is by setting up assessments with opportunities for numerous attempts, retakes, or reassessments, giving students additional opportunities to demonstrate learning from formative feedback (Abell et al., 2018). Cognitive psychology research has consistently found that errors on tests can spark significant learning and retention, but only if the feedback is immediate, not delayed (Brame & Biel, 2015; Hays et al., 2013). Based on the cognitive psychology literature on testing, Brame and Biel (2015) recommended that low- and no-stakes testing environments offer the most benefit of "test-enhanced learning" (p. 9). Test-enhanced learning is when testing becomes a learning opportunity for students, and the vehicle that provides this tremendous learning value is using formative assessments with reassessment. Thus, the authors suggested that incorporating multiple testing opportunities such as reassessments or retakes increases student learning in undergraduate science courses. The increased frequency of formative feedback allowed adult students to view formative assessments as "learning events" and students began to evaluate their errors based on learning outcomes for the course, creating the potential for greater recall on the reassessment (p. 10).

The body of assessment research has provided evidence for formative assessments to be viewed not simply as evaluative but also transformative to student learning when accompanied by immediate

feedback and the opportunity to reassess. The formative assessment process changed the students' beliefs about testing situations and their abilities by directing the learner's goals from being oriented toward performance to learning (Grant & Dweck, 2003). Feedback accomplished this shift in goal orientation by helping students view learning as a skill advanced by practice, effort, and mistakes (Shute, 2008). Building confidence, students shifted responsibility for their learning to themselves, which can lead to life-long learning and even success in future classes (Ghaicha, 2016; Hassi & Laursen, 2015; Shute, 2008; Wride, 2017).

Thus, a literature review is warranted to connect the formative assessment cycle with the important student outcomes of attitudes toward statistics and improved achievement. The purpose of the literature review is to uncover research that studied elements of the formative assessment cycle in large-enrollment, undergraduate introductory statistics courses, and its effects on student outcomes. From the search, gaps in the research will also be identified.

3. METHODS

To explore empirical research connecting the curriculum intervention of formative assessment in large-enrollment, undergraduate quantitative courses with the outcomes of student attitudes and achievement, the author conducted an informal search of the literature from the years 2007-2020. This search was conducted from electronic search engines: ERIC, Google Scholar, and EBSCOhost using combinations of the search terms shown in Table 1. The selection of articles included empirical studies or meta-analyses from peer-reviewed journals. Research on formative assessments utilizing technology, and large-enrollment quantitative introductory courses measuring the outcomes of attitudes and/or achievement are targets for the review.

Table 1. Search criteria categories

Criteria 1	Criteria 2	Criteria 3	Criteria 4
Introductory Statistics Statistics Education	Undergraduate Large Enrollment Tertiary	Formative Assessment Feedback Reassessment	Attitudes Achievement

Some articles found through the literature search were excluded for the following reasons: small sample sizes, practitioner papers, and those whose interventions were not formative assessment elements. Other exclusion criteria included studies that did not measure cognitive attributes (e.g., motivation, attitudes, persistence) or achievement scores. Books were also not included. Searching the ERIC database provided 215 articles, yielding 5 articles applicable to this essay. EBSCOhost produced 268 articles, of which six were appropriate to the search terms. The *Google Scholar* search of the search terms gave three empirical articles out of a total of 46 articles. Thus, the author turned to the *MAA Instructional Practices Guide* (2018), which provided five articles and a meta-analysis that demonstrated the use of formative assessment in quantitative, large enrollment courses (beyond introductory statistics) to impact student attitudes and achievement. From those three articles, three more studies related to the search terms. See the Appendix for a list of the empirical studies and the country in which the studies were conducted.

A total of 22 empirical studies were found for this conceptual essay of which 15 articles studied introductory statistics courses and courses that had a statistics component (e.g., psychology or business statistics). Few empirical studies were found to address large-enrollment classrooms using formative assessment to improve the important student outcomes of attitudes toward statistics or achievement. Given the research by Cash et al. (2017), it is not surprising that the research lacks large-enrollment courses importing the recommendations for formative assessments.

Although the author does not purport this search to be exhaustive, this search illustrates the large gap in the important research of statistics education and formative assessment in large-enrollment undergraduate settings. More research in introductory statistics is needed where formative assessment is incorporated into the curriculum of a large-enrollment introductory statistics course to improve student attitudes and student achievement. The findings are explained in the next section.

4. THE CONCEPTUAL FRAMEWORK FOR FORMATIVE ASSESSMENT CYCLES IN LARGE-ENROLLMENT INTRODUCTORY STATISTICS

As explicated in theory, formative assessment is a cycle of assessment utilizing three elements: lower or no-stakes assessments, feedback, and reassessment opportunities. These three elements, embedded into a large-enrollment course, create continuous, open assessments (Steen, 1999). First, by employing frequent, low-stakes assessments rather than few high-stakes assessments, students are given more assessment opportunities versus a single midterm for a large portion of a student's grade. The MAA describes assessment as more than a few high-stakes tests, but rather, as a "wider set of measures," in which varied assessments measure students' progress on learning outcomes (Abell et al., 2018, p. 50). Moreover, having more frequent low- to medium-stakes assessments embeds continuous assessments in a course (Steen, 1999).

Second, instructors must connect timely feedback from the formative assessments to learning goals. Students see their level of understanding, self-assess their learning, and utilize self-regulation to create actions for improving their knowledge from the learning goals. Finally, by allowing reassessment of the learning goals not yet mastered, students engage in self-efficacy as they see their learning and understanding improve. As suggested for large-enrollment courses, technology allows these large class sizes to be assessed ongoingly with feedback and reassessments to achieve course objectives (Abell et al., 2018; Simonite & Targett, 2010; Spencer, 2010; Stirling, 2010). Furthermore, implementing formative assessments with feedback and reassessment opportunities creates a continuous cycle rather than isolated assessment events.

Research has reported that student attitudes and achievement are important outcomes in introductory statistics (Ramirez et al., 2012; Xu & Schau, 2019). When improved, these two student outcomes could be key in providing pathways for students to successfully complete their quantitative requirements for their undergraduate studies (Peck, 2019). The literature review revealed studies regarding at least one of the formative assessment cycle elements and measuring the outcomes of student attitudes and achievement. Thus, the conceptual framework implements the entire cycle as the base of the introductory statistics curriculum to improve both attitudes and achievement (see Figure 2). The dotted lines in Figure 2 suggest a possible association between the formative assessment cycle and student attitudes toward statistics or student achievement.
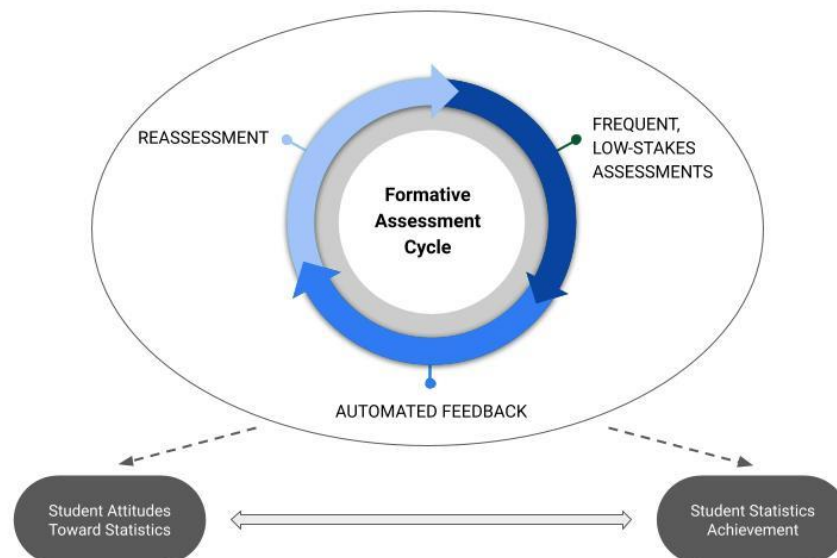

Figure 2. The conceptual framework for embedding formative assessment cycles in large-enrollment introductory statistics to improve student outcomes of attitudes and achievement.

4.1. COMPUTER-ASSISTED ASSESSMENTS

Using technology in mathematics and statistics courses has been studied for over 35 years. An overwhelming body of evidence supports its use in introductory statistics courses (Abell et al., 2018; ASA Revision Committee, 2016; Simonite & Targett, 2010; Spencer, 2010; Stirling, 2010). Additionally, using technology for online homework and quizzes results in positive responses from students and instructors. For example, the American Mathematical Society (2009) surveyed over 1,200 mathematics and statistics departments in universities in the United States and found three main advantages to online homework and learning systems: immediate feedback, multiple attempts allowed for problems, and less grading. The studies described in the next section from the literature review employed computer-based elements of formative assessment cycles which appeal to large-enrollment classes to minimize grading time for the instructor.

The results presented next synthesized this research to illustrate the connections with the three elements of formative assessment and student attitudes and achievement in large-enrollment quantitative literacy courses.

4.2. FORMATIVE ASSESSMENT AND ATTITUDES AND ACHIEVEMENT

Formative assessments comprise low to no-stakes assessments with feedback. Simple yet effective feedback is a necessary condition for an assessment to be formative. Formative feedback consistently is the most critical element of formative assessment to assist learning (Hattie and Timperley, 2007). Recall that feedback should connect to learning goals. The adult learner then uses self-assessment to learn from the feedback. Thus, feedback is an inseparable component of the definition of formative assessment. Therefore, formative assessment as no- to low-stakes assessment with feedback is the focus of this section, relating to student attitudes and achievement.

Formative assessments can be employed both in-and out-of-class. For example, in-class assessment strategies can utilize an online student response system (OSRS) or polling system for large-enrollment courses. Several studies suggest these in-class, no-stakes assessments effectively elicit student motivation and activate learning (Freeman et al. 2014; Gundlach et al., 2015, Muir et al., 2020). What is apparent is that through using in-class low and no-stakes formative assessments, the curriculum embeds active-learning approaches, which positively affects many aspects of the educational climate in large-enrollment courses: student achievement, attitudes, motivation, engagement, and perceived achievement (Freeman et al., 2014; Gundlach et al., 2015, Muir et al., 2020). For example, in a meta-analysis of 225 studies in STEM courses, online student response systems promoted active learning in large-enrollment course lectures. In these activated lectures, exam scores increased by 6% over traditional lectures (Freeman et al., 2014).

Gundlach et al. (2015) used both an OSRS and online homework in a large face-to-face section of introductory statistics. They measured student attitudes using the *Survey of Attitudes Toward Statistics* (SATS-36) instrument (Schau, 2003), finding that student attitudes improved in both the *affect* and *cognitive competence* subscales. In addition, these students' summative exam scores were higher than the flipped and online introductory statistics sections (Gundlach et al., 2015). In a large-enrollment introductory statistics study for non-statistics majors by Hodgson and Pang (2012), online formative activities (OFAs) improved self-regulation. More than 60% of students reported increased motivation, and over 70% stated that the OFAs helped their learning and understanding of statistics.

Computer-assisted assessments such as online homework have increased course performance in quantitative classes. For example, in a large study of both freshmen mathematics and statistics students, computer-based formative assessments helped identify under-performing students and improved final exam scores in learning mathematics and statistics (Tempelaar et al., 2014). Another large-scale study evaluated the use of web-based homework for a calculus course and found that freshman students' grades improved on average by two letters of those who utilized the web-based homework (Hirsch & Weibel, 2003). Finally, several studies found web-based homework significantly improved test scores over pen and paper homework in large cohorts of introductory students (Balta & Güvercin, 2016; Chua-Chow et al., 2011; Jonsdottir et al., 2017). The authors of these studies all noted that the improvement of exam scores could be attributed to the immediate feedback the online homework provided.

Frequent, low stakes assessments via online homework or tests have been shown to improve both student motivation and achievement. Broadbent et al. (2018) suggested breaking up larger summative assessments into smaller, lower-stakes assessments to assist large-enrollment courses with employing formative assessments. Used in a large undergraduate psychology class, students reported increased motivation, improved ability to self-assess, and higher achievement in the course.

In large-enrollment quantitative literacy courses, the Just in Time Teaching (JiTT) model was developed in 1996 to aid undergraduate science and mathematics students to use out-of-class time more effectively (Novak et al., 1999). With web-enhanced learning, the use of JiTT quizzes before class to inform teachers of student understanding relative to learning outcomes has been improved and made simpler (Abell et al., 2018). The implementation of JiTT has also been successful in introductory statistics courses at major universities. Testing JiTT's effectiveness by analyzing pre- and post-test scores, the average post-test scores were higher than semesters where JiTT was not implemented (McGee et al., 2016). Implementing these formative assessments before class time can improve student participation and ownership of class material (McGee et al., 2016; Natarajan & Bennett, 2014).

4.3. FORMATIVE FEEDBACK AND ITS EFFECT ON QUANTITATIVE LEARNING IN LARGE-ENROLLMENT COURSES

Feedback is often studied apart from the formative assessment as students report motivation and competence with self-assessment and find that feedback assists them in improving their performance on assessments (Abell et al., 2018; ASA Revision Committee, 2016; Shute, 2008). Automatic feedback given through computer assessment is the focus of the conceptual framework for large-enrollment quantitative courses. When feedback maintains the qualities of timeliness, concise elaboration on the learning outcomes, and computer delivery, it is associated with increased performance and learning (Shute, 2008).

Feedback can also improve self-efficacy in adult learners. Several studies reported that the automated feedback in the formative assessments contributed to students' enjoyment of their learning experience (Balta & Güvercin, 2016; Beemer et al., 2018; Chua-Chow et al., 2011; Hodgson & Pang, 2012; Jonsdottir et al., 2017; Krause et al., 2009; Posner, 2011). For example, Krause et al. (2009) found that e-learning provided university students with perceived competence in statistics. Additionally, the feedback helped students' achievement with little prior statistics knowledge or experience, which is particularly notable as students experience anxiety toward statistics and lack confidence in their quantitative abilities in introductory statistics courses (Onwuegbuzie & Wilson, 2003; Williams, 2015). Broadbent et al., (2018) studied a large-enrollment undergraduate course ($N \sim 1,500$) that used formative assessments and found over 83% of the students agreed that the online formative feedback on assessments motivated them to learn, improved their understanding, and increased their learning. Moreover, because the instructors improved upon the feedback for the following semesters, students' average achievement increased over 10% in the subsequent semesters as the feedback improved (Broadbent et al., 2018). Thus, feedback can provide essential information to the adult learner in large-enrollment courses which results in improved attitudes and achievement.

In addition, large-enrollment introductory statistics courses employing formative examinations with feedback reported improved learning outcomes. Massing et al. (2018) studied the use of computer-assisted assessments with automated feedback in a large-enrollment statistics course. They found three significant increases in academic outcomes linked to the computer assessments using automatic feedback: student effort, student success in achieving learning outcomes, and final grades (Massing et al., 2018). Together, these studies support formative assessment with feedback as instructional interventions to improve student outcomes in large-enrollment introductory statistics. However, to make the cycle complete, giving adult learners the opportunity to retest their knowledge and learn from their mistakes is needed.

4.4. REASSESSMENT AND ATTITUDES AND ACHIEVEMENT

An often-overlooked aspect of the formative assessment cycle is reassessment. With formative feedback as a necessary condition of formative assessment, allowing students to learn from their mistakes is a natural next step. Unfortunately, there is less literature studying reassessment opportunities

in large-section introductory quantitative courses. However, as andragogy posits that adult learners are motivated and desire to learn through self-assessment, reassessment is a valuable element in the formative assessment cycle for creating successful student pathways in introductory statistics. Grant & Dweck (2003) studied pre-med majors in a large chemistry class. They found that the key to making retakes work was the ability for students to recover from a poor initial attempt by using achievement goals to provide coping, motivation, and higher achievement on exams.

Students in mathematics and statistics courses have benefitted from reassessments. A large study of undergraduate mathematics and statistics students by Collins et al. (2019) investigated reassessment opportunities. The assessments allowed multiple attempts with credit only for mastery. Well over 80% of the students felt the assessments and reassessment opportunities helped them understand the material, prepared them for problem-solving, and reflected their knowledge. Allowing reassessments led to reduced pressure on students during examinations and improved attitudes because of added opportunities for student success. In addition, the reassessments also showed that students felt the motivation to utilize feedback, which stated simply whether the student "mastered," was "progressing," or was "insufficient" in that concept for the next assessment attempt. Because the learning outcomes were directly related to assessment material, the students revisited course material to develop further understanding (Collins et al., 2019). In another study of undergraduate mathematics students using web-based homework with the opportunity for students to revise and resubmit answers, Hirsch and Weibel (2003) found a high correlation ($r = .944$) between attempts made at homework problems and the percentage of problems solved in a calculus course. This study suggested that students displayed persistence with the web-based homework system: they kept working on a problem until they achieved the correct solution, despite students' prior mathematical ability for the course. Thus, students' persistence was more a function of "effort rather than ability" (p. 14). Lenz (2010) found that mathematics undergraduates' homework scores improved due to web-based homework assessments that utilized both feedback and the opportunity to resubmit homework multiple times.

Several studies specifically evaluated reassessment in introductory statistics. Posner (2011) found that students who chose to resubmit work increased their proficiency in introductory statistics concepts, compared to those who achieved proficiency with only one submission. Hodson and Pang (2012) noted that computer-assisted assessments created favorable attitudes in quantitative classes when online formative activities used feedback and multiple attempts. They found that students were highly satisfied with the online approach and noted that their abilities to self-assess increased as they actively worked on finding solutions. A large study conducted in Italy allowed multiple attempts at exams: students who failed the exam could utilize feedback to self-assess and learn from their mistakes (Chiesi and Primi, 2010). Using a mathematical preparedness test score taken prior to the semester, Chiesi and Primi divided the subjects into low and high mathematics preparedness groups to determine the change in attitudes among these groups. Both groups' attitudes increased. The students with lower mathematical preparation, however, experienced lower feelings of competence, which impacted students' self-efficacy in statistical and mathematical concepts. Reassessment provided these students additional chances to succeed, experiencing higher self-confidence and greater worth in statistics at the end of the course. The authors noted that creating pedagogical interventions to increase students' feelings of competence could impact student achievement especially in students with lower mathematical preparedness. These studies punctuate the importance of reassessment for student attitudes and achievement.

As the literature highlights and the conceptual framework posits, student attitudes and achievement can be impacted effectively when all three elements of the formative assessment cycle are implemented in large-enrollment quantitative courses.

5. IMPLEMENTING THE FORMATIVE ASSESSMENT CYCLE IN INTRODUCTORY STATISTICS

Several practitioner papers and book chapters provided examples of large-enrollment introductory statistics courses implementing elements of the formative assessment cycle using technology. These papers are excellent resources for ways to apply computer-based formative assessments in the introductory statistics curriculum. For example, using online student response systems (OSRS) such as clickers or polling software import formative assessment with automatic feedback and the opportunity

to reassess a students' understanding. Full lesson plans, including OSRS questions and how the instructors used these lesson plans in their classes for exam reviews, data exploration, and lectures are available (Bruff, 2011; Gunderson & McGowan, 2011; Murphy et al., 2011). In addition, Peck (2011) created a website with links to OSRS questions and other online resources (http://mathquest.carroll.edu/resources.html). An informal experiment by Peck (2011) found that students who used an OSRS had higher average final exam scores than those students who did not. An important confounder to this informal study is that students in the OSRS sections of introductory statistics had higher attendance.

There are also many options for online and web-based homework systems for introductory statistics: from textbook publishers to free and open-source homework systems mentioned in Lunsford and Pendergrass (2016). The authors also discussed the importance of affording students' multiple attempts on each homework question, allowing students to take advantage of the automated feedback available in the online environment. Additionally, Lunsford and Pendergrass provided suggestions for instruction. They suggested using 10 minutes of class time to discuss the online homework, using student homework notebooks to take notes of their problem-solving processes, and assigning frequent, even daily, online homework. Implementation of web-based homework benefited students when options to resubmit homework are provided.

Lastly, but importantly, implementing the formative assessment cycle with frequent, low-stakes examinations in large-enrollment introductory statistics is made possible with computer-generated testing and grading. Creating individualized testing can be time consuming but can provide easy grading for the instructor of large-enrollment courses, and immediate, individualized feedback making the time-commitment worth the effort to import formative assessments for all (Spencer, 2010; Stirling, 2010). The effort of creating test banks, randomizing questions, and providing individualized assessments is an investment of instructor and departmental resources, but one that affords students the opportunity to experience automatic feedback, self-assessment, and retakes or reassessments in a large-enrollment learning environment (Simonite & Targett, 2010).

6. CONCLUSIONS

But all can be judged by the same criteria: an open process, beginning with goals, that measures and enhances students' mathematical performance; that draws valid inferences from multiple instruments; and that is used to improve instruction for all students. (Steen, 1999, p. 5)

There is a large body of research examining formative assessment, bringing awareness to the educational community of the need for students to experience more courses with formative assessment: lower-stakes assessments, automatic feedback aligned with learning outcomes, and the opportunity to be reassessed. In addition, with the recent advent of learning management systems, the internet, and online resources, large-enrollment classes can provide formative assessment cycles by incorporating embedded, computerized formative assessments with feedback and multiple attempts in the curriculum. Due to high-stakes summative examination practices, introductory statistics courses often evoke anxiety and negative attitudes (ASA Revision Committee, 2016; Chew & Dillon, 2014; Malik, 2015; Williams, 2015). The use of computer-based online assignments that utilize the elements of assessment theory for importing high-quality, continuous assessments with automatically generated feedback and multiple attempts can transform large, introductory statistics courses into those where formative assessment is the norm.

A conceptual framework of a three-pronged approach to formative assessment resulted from the review of empirical studies in large-enrollment, quantitative courses. Additionally, the research suggested using web-based or online assessments to manage the large enrollments. Finally, the literature indicated that students reported improved attitudes using web-based formative assessment and these assessments prepared them for summative exams. Thus, the formative assessment cycle could help answer the call for a comprehensive transformation of introductory statistics courses, creating pathways for students to complete their quantitative literacy requirements successfully (Pfannkuch, 2018; Peck, 2019). Providing a learning environment with formative assessment cycles embedded in the curriculum could foster self-efficacy and self-regulation in adult learners. Although this review suggests that pieces of the formative assessment cycle have been studied to some extent, empirical research lacks in

implementing the entirety of the formative assessment cycle in large-enrollment undergraduate statistics courses. Thus, future studies are needed in large-enrollment introductory statistics to examine formative assessment cycles' impact on students' self-efficacy and self-regulation through andragogical tenets to improve student and course outcomes, providing a gateway rather than a bottleneck in furthering students' educational pursuits.

As student attitudes towards statistics are an outcome in introductory statistics courses the statistics education community has deemed valuable and aspires to improve (Xu & Schau, 2019), there is a need for studies of formative assessment cycles as a curricular intervention that potentially improves student attitudes toward statistics, thereby fostering an appreciation for statistics and statistical thinking that extends beyond the classroom (Ramirez et al., 2017) especially for students with such diverse and varying mathematical backgrounds found in introductory statistics courses. Where there is a bidirectional association between student attitudes and achievement, the impact of formative assessment cycles could be the key to successfully completing the introductory statistics course.

The effects of formative assessment cycles have far-reaching implications for instructors, students, statistics departments, and curriculum designers. Currently, instructors tend to be their own barrier to implement a change in assessment in a classroom of several hundred students (Cash et al., 2017; Garfield et al., 2002, 2011). Instructors require support for applying evidence-based pedagogical transformations to their courses, stating the "lack of time to plan, practice, use and reflect" being of great concern regarding making those changes (Ghaicha, 2016, p. 222). Instructors must also respond to student feedback constructively and without judgment, which is difficult with large-enrollment courses without computer-assisted testing (Shute, 2008). To address these concerns, mathematics and statistics departments can allocate resources for instructors to create and implement formative assessment cycles in large-enrollment courses. Moreover, with future research, introductory statistics curricula and course designers will more readily design course curriculum employing formative assessment cycles, creating a learning environment which empowers all adult learners.

The purpose of this essay was to provide a conceptual framework for formative assessment cycles (FACs) in the introductory statistics curriculum. As evidenced by the literature, the individual elements of FACs can improve the student outcomes of attitudes toward statistics and achievement. Thus, embedding FACs can potentially renovate the introductory statistics curriculum, providing all students successful pathways to achieve their quantitative literacy requirement in higher education.

The following research questions are offered for future study:

- How does the implementation of FACs in large-enrollment introductory statistics courses support achievement?
    - Do pass/fail/withdrawal rates improve?
    - Do assessment (homework, quiz, exam) scores improve when offered reassessment?
    - Do FACs improve summative scores?
    - Do FACs improve statistical thinking and statistical understanding?

- What is the impact on student attitudes toward statistics, statistics anxiety, self-efficacy, self-regulation, and other cognitive attributes after experiencing a large-enrollment course with FACs?

- Do FACs provide an equitable student experience among students of diverse mathematical backgrounds, majors, genders, ethnicities, and interests?

With the current focus on improving statistics achievement and creating successful experiences for students, assessment must be integrated as a continuous cycle—the foundation of the introductory statistics experience (ASA Revision Committee, 2016; Steen, 1999). Formative assessment cycles can impact students during the course as they complete their education, and potentially improve their attitudes toward statistics throughout their lives. As students experience increased opportunities for success through self-assessment from formative feedback and repeating assessment opportunities, formative assessment cycles could reduce statistics anxiety, improve attitudes towards statistics, and increase statistics achievement, preparing a new generation of statistically literate citizens for a data-driven world.

ACKNOWLEDGEMENTS

REFERENCES

Abell, M., Braddy, L., Ensley, D., Ludwig, L., & Soto, H. (2018). *MAA instructional practices guide*. Mathematical Association of America. https://www.maa.org/programs-and-communities/curriculum%20resources/instructional-practices-guide

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., Moore, T., Rossman, A., Stephenson, B., Utts, J., Velleman, P., & Witmer, J. (2005). *Guidelines for Assessment and Instruction in Statistics Education College Report*. https://www.amstat.org/asa/files/pdfs/GAISE/2005GaiseCollege_Full.pdf

American Mathematical Society. (AMS). (2009). *AMS homework software survey*. http://www.ams.org/profession/leaders/WATSreport.pdf

American Statistical Association Revision Committee. (2016). *Guidelines for assessment and instruction in statistics education (GAISE) college report*. American Statistical Association. http://www.amstat.org/education/gaise

Assessment Reform Group. (2002). *Assessment for learning:10 principles*. https://dspace.stir.ac.uk/bitstream/1893/32458/1/Assessment%20for%20learning%2010%20principles%202002.pdf

Balta, N., & Güvercin, S. (2016). Increasing undergraduate students' exam performances in statistics courses using software Socrative. *Turkish Online Journal of Educational Technology: Special Issue for IETC, ITEC, IDEC, ITICAM 2016* (pp. 314–321). http://www.tojet.net/special/2016_7_1.pdf

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bayerlein, L. (2014). Students' feedback preferences: How do students react to timely and automatically generated assessment feedback? *Assessment and Evaluation in Higher Education*, *39*(8), 916–931. https://doi.org/10.1080/02602938.2013.870531

Beemer, J., Spoon, K., Fan, J., Stronach, J., Frazee, J. P., Bohonak, A. J., & Levine, R. A. (2018). Assessing instructional modalities: Individualized treatment effects for personalized learning. *Journal of Statistics Education*, *26*(1), 31–39. https://doi.org/10.1080/10691898.2018.1426400

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. In D Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 473–502). Springer. https://link.springer.com/chapter/10.1007/978-3-319-66195-7_16

Bidgood, P., Hunt, N., & Jolliffe, F. (Eds.). (2010). *Assessment methods in statistical education: An international perspective*. Wiley.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, *86*(1), 8–21. https://doi.org/10.1177/003172170408600105

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *The Phi Delta Kappan*, *80*(2), 139–148. https://doi.org/10.1177/003172171009200119

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Blair, R. M., Kirkman, E. E., Maxwell, J. W., & National Science Foundation (U.S.). (2018). *Statistical abstract undergraduate programs in the mathematical sciences in the United States : Fall 2015 CBMS survey*. American Mathematical Society. https://www.ams.org/profession/data/cbms-survey/cbms2015-Report.pdfBoaler, J., & Confer, A. (2017). *Assessment for a growth mindset*. Stanford University. https://www.youcubed.org/wp-content/uploads/2017/03/1439422682-AssessmentPaper.pdf

Brame, C. J., & Biel, R. (2015). Test-enhanced learning: The potential for testing to promote greater learning in undergraduate science courses. *CBE Life Sciences Education*, *14*(2), 1–12. https://doi.org/10.1187/cbe.14-11-0208

Broadbent, J., Panadero, E., & Bouda, D. (2018). Implementing summative assessment with a formative flavour: A case study in a large class. *Assessment and Evaluation in Higher Education*, *43*(2), 307–322. https://doi.org/10.1080/02602938.2017.1343455

Bruff, D. (2011). Engaging statistics students with classroom response systems. In K. Cline & H. Zullo (Eds.), *Teaching mathematics with classroom voting with and without clickers* (pp. 61–70). The Mathematical Association of America. https://doi.org/10.1017/CBO9781614443018.010

Cash, C. B., Letargo, J., Graether, S. P., & Jacobs, S. R. (2017). An analysis of the perceptions and resources of large university classes. *CBE Life Sciences Education*, *16*(2), 1–12. https://doi.org/10.1187/cbe.16-01-0004

Chew, P., & Dillon, D. (2014). Statistics anxiety update: Refining the construct and recommendations for a new research agenda. *Perspectives on Psychological Science*, *9*(2), 196–208. https://doi.org/10.1177/1745691613518077

Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal*, *9*(1), 6–26. https://doi.org/10.52041/serj.v9i1.385

Chua-Chow, C., Chauncey, C. A., & McKessock, D. (2011). Enhancing the study of business statistics with an e-homework system. *Journal of Education and Vocational Research*, *1*(3), 96–105. https://doi.org/10.22610/jevr.v1i3.15

Collins, J. B., Harsy, A., Hart, J., Haymaker, K., Hoofnagle, A. M., Janssen, M. K., Kelly, J. S., Mohr, A. T., & OShaughnessy, J. (2019). Mastery-based testing in undergraduate mathematics courses. *PRIMUS*, *29*(5), 441–460. https://doi.org/10.1080/10511970.2018.1488317

Complete College America. (2012). *Remediation: Higher education's bridge to nowhere*. https://completecollege.org/wp-content/uploads/2017/11/CCA-Remediation-final.pdf

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *International Journal of Phytoremediation*, *21*(1), 101–116. https://doi.org/10.1080/09695949993026

Davies, N., & Marriott, J. (2010). Assessment and feedback in statistics. In P. Bidgood, N. Hunt, & F. Joliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 3–20). Wiley.

Duckworth, A. L., Peterson, C., Matthews, M. D., Kelly, D. R., Farah, M., Latham, G., Rozin, P., & Shell, R. (2007). Grit : Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Dweck, C. S. (2008). *Mindsets and math/science achievement*. Carnegie Corporation of New York. http://www.growthmindsetmaths.com/uploads/2/3/7/7/23776169/mindset_and_math_science_achievement_-_nov_2013.pdf

Emmioğlu, E., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal*, *11*(2), 95–102. https://doi.org/10.52041/serj.v11i2.332

Fong, K. E., Melguizo, T., & Prather, G. (2015). Increasing success rates in developmental math: The complementary role of individual and institutional characteristics. *Research in Higher Education*, *56*(7), 719–749. https://doi.org/10.1007/s11162-015-9368-9

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Jordt, H., Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Science of the United States of America*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Garfield, J., Hogg, B., Schau, C., & Whittinghill, D. (2002). First courses in statistical science: The status of educational reform efforts. *Journal of Statistics Education*, *10*(2), 1069–1898. https://doi.org/10.1080/10691898.2002.11910665

Garfield, J., Delmas, R., & Zieffler, A. (2010). Assessing important learning outcomes in introductory tertiary statistics courses. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 75–86). Wiley.

Garfield, J., Zieffler, A., Kaplan, D., Cobb, G. W., Chance, B. L., & Holcomb, J. P. (2011). Rethinking assessment of student learning in statistics courses. *American Statistician*, *65*(1), 1–10. https://doi.org/10.1198/tast.2011.08241

Ghaicha, A. (2016). Theoretical framework for educational assessment: A synoptic review. *Journal of Education and Practice*, *7*(24), 212–231. www.iiste.org

Gold, B., Keith, S., & Marion, W. A. (1999). *Assessment practices in undergraduate mathematics*. Mathematical Association of America. https://www.maa.org/sites/default/files/pdf/ebooks/pdf/NTE49.pdf

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, *29*(1), 17–28. https://doi.org/10.1016/j.appdev.2007.10.004

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, *85*(3), 541–553. https://doi.org/10.1037/0022-3514.85.3.541

Gunderson, B. K., & McGowan, H. M. (2011). Using clickers in a statistics classroom. In K. Cline & H. Zullo (Eds.), *Teaching mathematics with classroom voting with and without clickers* (pp. 53–60). The Mathematical Association of America. https://doi.org/10.1017/CBO9781614443018.009

Gundlach, E., Richards, K., Nelson, D., & Levesque-Bristol, C. (2015). A comparison of student attitudes, statistical reasoning, performance, and perceptions for web-augmented traditional, fully online, and flipped sections of a statistical literacy class. *Journal of Statistics Education*, *23*(1). https://doi.org/10.1080/10691898.2015.11889723

Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. In John Gardner (Ed.). In *Assessment and learning* (pp. 87–102). SAGE Publications. http://dx.doi.org/10.4135/9781446250808.n6

Hassi, M. L., & Laursen, S. L. (2015). Transformative learning: Personal empowerment in learning mathematics. *Journal of Transformative Education*, *13*(4), 316–340. https://doi.org/10.1177/1541344615587111

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hays, M. J., Kornell, N., & Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*(1), 290–296. https://doi.org/10.1037/a0028468

Hirsch, L., & Weibel, C. (2003). Statistical based evidence that web-based homework helps. *MAA Focus*, *23*(2), 14–15. https://www.maa.org/sites/default/files/pdf/pubs/feb03.pdf

Hoang, H., Huang, M., Sulcer, B., & Yesilyurt, S. (2017). *Carnegie math pathways 2015-2016 impact report: A five-year review*. www.carnegiefoundation.org/resources/publications/carnegie-math-pathways-2015-2016-impact-report-a-five-year-review/

Hodgson, P., & Pang, M. (2012). Effective formative e-assessment of student learning: A study on a statistics course. *Assessment and Evaluation in Higher Education*, *37*(2), 215–225. https://doi.org/10.1080/02602938.2010.523818

Hopfenbeck, T. (2018). Classroom assessment, pedagogy and learning–twenty years after Black and Wiliam 1998. *Assessment in Education: Principles, Policy and Practice*, *25*(6), 545–550. https://doi.org/10.1080/0969594X.2018.1553695

Houle, C. O. (1996). *The design of education* (2nd ed.). Jossey-Bass.

Hunt, N. (2010). Individualised assessment in statistics. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 203–210). Wiley.

Jonsdottir, A. H., Bjornsdottir, A., & Stefansson, G. (2017). Difference in learning among students doing pen-and-paper homework compared to web-based homework in an introductory statistics course. *Journal of Statistics Education*, *25*(1), 12–20. https://doi.org/10.1080/10691898.2017.1291289

Knowles, M. S. (1978). Andragogy: Adult learning theory in perspective. *Community College Review*, *5*(3), 9–20.

Krause, U. M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction*, *19*(2), 158–170. https://doi.org/10.1016/j.learninstruc.2008.03.003

Lavidas, K., Barkatsas, T., Manesis, D., & Gialamas, V. (2020). A structural equation model investigating the impact of tertiary students' attitudes toward statistics, perceived competence at mathematics, and engagement on statistics performance. *Statistics Education Research Journal*, *19*(2), 27–41. https://www.iase-web.org/ojs/SERJ/article/view/108 Lenz, L. (2010). The effect of a web-based homework system on student outcomes in a first-year mathematics course. *Journal of*

Computers in Mathematics and Science Teaching, *29*(3), 233–246. https://www.learntechlib.org/p/32318/

Lindeman, E. (1926). *Workers' education and the public libraries* (Vol. 7). Workers' Education Bureau of America.

Lunsford, M. L., & Pendergrass, M. (2016). Making online homework work. *PRIMUS*, *26*(6), 531–544. https://doi.org/10.1080/10511970.2015.1110219

Malik, S. (2015). Undergraduates' statistics anxiety: A phenomenological study. *The Qualitative Report*, *20*(2), 120–133. https://doi.org/10.46743/2160-3715/2015.2101

Mangels, J. A., Butterfield, B., Lamb, J., Good, C., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective Neuroscience*, *1*(2), 75–86. https://doi.org/10.1093/scan/nsl013

Martin, J. (2010). Self-regulated learning, social cognitive theory, and agency. *Educational Psychologist*, *39*(2), 135–145. https://doi.org/10.1207/s15326985ep3902

Massing, T., Schwinning, N., Striewe, M., Hanck, C., & Goedicke, M. (2018). E-assessment using variable-content exercises in mathematical statistics. *Journal of Statistics Education*, *26*(3), 174–189. https://doi.org/10.1080/10691898.2018.1518121

McGee, M., Stokes, L., & Nadolsky, P. (2016). Just-in-time teaching in statistics classrooms. *Journal of Statistics Education*, *24*(1), 16–26. https://doi.org/10.1080/10691898.2016.1158023

Merriam, S. B. (2001). Andragogy and self-directed learning: Pillars of adult learning theory. *New Directions for Adult and Continuing Education*, *2001*(89), 3–13. https://doi.org/10.1002/ace.3

Muir, S., Tirlea, L., Elphinstone, B., & Huynh, M. (2020). Promoting classroom engagement through the use of an online student response system: A mixed methods analysis. *Journal of Statistics Education*, *28*(1), 25–31. https://doi.org/10.1080/10691898.2020.1730733

Murphy, T. J., McKnight, C. C., Richman, M. B., & Terry, R. (2011). Clickers in introductory statistics courses. In K. Cline & H. Zullo (Eds.), *Teaching mathematics with classroom voting with and without clickers* (pp. 43–52). The Mathematical Association of America. https://doi.org/10.1017/CBO9781614443018.008

Natarajan, R., & Bennett, A. (2014). Improving student learning of calculus topics via modified just-in-time teaching methods. *PRIMUS*, *24*(2), 149–159. https://doi.org/10.1080/10511970.2013.854853

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.

Nind, M., & Lewthwaite, S. (2015). *NCRM quick start guide to: Principles for effective pedagogy*. National Center for Research Methods.

Novak, G., Patterson, E., Gavrin, A., & Christian, W. (1999). *Just-in-time-teaching: Blending active learning with web technology*. Prentice Hall.

Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics anxiety: Nature, etiology, antecedents, effects, and treatments--a comprehensive review of the literature. *Teaching in Higher Education*, *8*(2), 195–209. https://doi.org/10.1080/1356251032000052447

Peck, R. (2011). Incorporating clicker technology in the introductory statistics course. In K. Cline & H. Zullo (Eds.), *Teaching mathematics with classroom voting with and without clickers* (pp. 71–76). The Mathematical Association of America. https://doi.org/10.1017/CBO9781614443018.011

Peck, R. (2019). The evolving statistics pathway. In R. Hartzler & R. Blair (Eds.), *Emerging issues in mathematics pathways case studies, scans of the field, and recommendations* (pp. 33–40). Charles A. Dana Center at The University of Texas at Austin.

Pfannkuch, M. (2018). Reimagining curriculum approaches. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 387–414). Springer.

Posner, M. A. (2011). The impact of a proficiency-based assessment and reassessment of learning outcomes system on student achievement and attitudes. *Statistics Education Research Journal*, *10*(1), 3–14. https://doi.org/10.52041/serj.v10i1.352

Ramirez, C., Schau, C., & Emmioğlu, E. (2012). The importance of attitudes in statistics education. *Statistics Education Research Journal*, *11*(2), 57–71. https://doi.org/10.52041/serj.v11i2.329

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, *10*(3), 1–12. https://doi.org/10.1080/10691898.2002.11910678

Schau, C. (2003). Students' attitudes: The "other" important outcome in statistics education. *The 2003 Joint Statistical Meetings* (pp. 3673–3683). http://statlit.org/pdf/2003SchauASA.pdf

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist*, *32*(4), 195–208. https://doi.org/10.1207/s15326985ep3204_1

Schield, M. (2010). Assessing statistical literacy: Take CARE. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 133–152). Wiley.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Simonite, V., & Targett, R. (2010). Assignments made in heaven? Computer-marked, individualised coursework in an introductory level statistics course. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 235–246). Wiley.

Snelgar, R., & Maguire, M. (2010). Assessing for success: An evidence-based approach that promotes learning in diverse, non-specialist student groups. In P. Bidgood, N. Hunt, & F. Joliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 35–46). Wiley.

Sosibo, Z. C. (2019). Self-assessment: A learner-centered approach towards transforming traditional practices and building self-directed learners. *South African Journal of Higher Education*, *33*(5), 76–97. https://doi.org/10.20853/33-5-3586

Spencer, N. (2010). An adaptive, automated, individualised assessment system for introductory statistics. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 211–222). Wiley.

Steen, L. A. (1999). Assessing assessment. In B. Gold, S. Z. Keith, & W. A. Marion. (Eds.), *Assessment practices in undergraduate mathematics* (pp. 1–6). Mathematical Association of America. https://www.maa.org/sites/default/files/pdf/ebooks/pdf/NTE49.pdf

Steen, L. A. (2006). *Supporting assessment in undergraduate mathematics*. Mathematical Association of America. https://www.maa.org/sites/default/files/pdf/ebooks/pdf/SAUM.pdf

Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758–765. https://doi.org/10.1177/003172170208301010

Stirling, D. (2010). Random computer-based exercises for teaching statistical skills and concepts. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 223–234). Wiley.

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2014). Computer-assisted, formative assessment and dispositional learning analytics in learning mathematics and statistics. In M. Kalz & E. Ras (Eds.) *Research into E-Assessment. CAA 2014. International Computer Assisted Assessment Conference.* Communications in Computer and Information Science (vol. 439, pp. 67–78). Springer. https://doi.org/10.1007/978-3-319-08657-6_7

Tishkovskaya, S., & Lancaster, G. A. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, *20*(2), 1–55. https://doi.org/10.1080/10691898.2012.11889641

Watson, J. M. (1997). Assessing statistical thinking using the media. In I. Gal & J. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107–121). IOS Press and The International Statistical Institute.

Williams, A. (2015). Statistics anxiety and worry: The roles of worry beliefs, negative problem orientation, and cognitive avoidance. *Statistics Education Research Journal*, *14*(2), 53–75. https://doi.org/10.52041/serj.v14i2.261

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, *10*, 1–14. https://doi.org/10.3389/fpsyg.2019.03087

Wride, M. (2017). *Academic practice and elearning (CAPSL) resources: Guide to self-assessment.* https://www.tcd.ie/CAPSL/assets/pdf/Academic%20Practice%20Resources/Guide%20to%20Student%20Self%20Assessment.pdf

Xu, C., & Schau, C. (2019). Exploring method effects in the six-factor structure of the survey of attitude toward statistics (SATS-36). *Statistics Education Research Journal*, *18*(2), 39–53. https://doi.org/10.52041/serj.v18i2.139

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., Tomita, M. K., & Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, *21*(4), 335–359. https://doi.org/10.1080/08957340802347845

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, *45*, 477–501.

Zieffler, A., Garfield, J., & Fry, E. (2018). What is statistics education? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 37–70). Springer. https://doi.org/10.1007/978-3-319-66195-7_2

KIMBERLEIGH FELIX HADFIELD

Department Of Mathematics and Statistics

Utah State University

3900 University Hill

Logan, Utah 84322

APPENDIX

AUTHORS OF EMPIRICAL STUDIES FROM THE LITERATURE SEARCH

Article	Country
American Mathematical Society (2009).	United States
Balta & Güvercin (2016).	Turkey
Beemer et al. (2018).	United States
Broadbent et al. (2018).	Australia
Chiesi & Primi (2010).	Italy
Chua-Chow et al. (2011).	Canada
Collins et al. (2019).	United States
Freeman et al. (2014).	United States
Grant & Dweck (2003).	United States
Gundlach et al. (2015).	United States
Hirsch & Weibel (2003).	United States
Hodgson & Pang (2012).	Hong Kong
Jonsdottir et al. (2017).	Iceland
Krause et al. (2009).	Germany
Lenz (2010).	United States
Massing et al. (2018).	Germany
McGee et al. (2016).	United States
Muir et al. (2020).	Australia
Natarajan & Bennett (2014).	United States
Novak et al. (1999).	United States
Posner (2011).	United States
Tempelaar et al., (2014).	Netherlands