University of Victoria

GRADING TOOLKIT

Division of Learning and Teaching Support and Innovation

August 2020

Contents

TOOLKIT OF RESOURCES REGARDING GRADING PATTERNS

The following toolkit[1] was developed to assist University of Victoria Chairs and Directors to carry out responsibilities in relation to grading, and to facilitate discussion in their respective units.

This Toolkit includes:

1. Literature review about assessment and related topics
2. Links to key documents (such as UVic grad descriptors and SAS grades portal)
3. Questions that can be used for discussion within units so that there is a shared understanding of grading practices and their relation to students' learning
4. Recommended practices for multi-section courses to achieve equitable assessment and grading practices
5. Best practices toolkit for instructors
6. Series of professional development sessions to assist instructors with assessment practices

Terminology used in this document:

Assessment – any mechanism instituted in a course to provide feedback and measure student learning, divided into two sub-categories: formative and summative (see Appendix 2 for a list of both of these forms).

Formative assessment – is *assessment for learning* that is primarily used to inform a student (through feedback by verbal, written or other means) about their learning and progression towards the requirements to successfully complete the course. Typically, formative assessment is not assigned a grade or, if it is, it has a low value.

Forms of formative assessment:

a. *Classroom or embedded assessment* – incorporated into instructional activities, such as using personal response systems, such as iClickers or classroom assessment techniques. Within a lesson plan for a class, pre-assessment and post-assessment are used. Pre-assessment is a diagnostic used at the beginning of class to determine what students know about concepts for that lesson. Post-assessment determines if students reached the intended learning outcomes set for that lesson.
b. *Peer-assessment* – students assess other students work with no value. There is much research about peer-assessment and most agree on two things: do not assign grades to peer-assessment, and when used as formative assessment, peer-assessment is very effective and contributes to students' ability to self-assess.
c. *Self-assessment* – students are required to assess their work against certain standards, such as a rubric, provided by the instructor. This provides an opportunity for students to reflect on their work to see how they can improve.

Summative assessment – is *assessment of learning* that is graded. Many forms of summative assessment exist but increasingly *authentic assessment* methods are being encouraged because they more accurately represent how student learning will be used in practice (Wiggins, 1998). For example,

[1] The Toolkit was designed to be useful to all units at UVic, however, some information contained in this document may not be applicable to your department. For specific policies, please refer to your Faculty's documents about procedures for issues related to grading.

according to Wiggins (1998), authentic assessment has the student "do" the subject by simulating real-world contexts so that students can actually apply the subject matter.

Literature review about assessment and related topics

Where does assessment take place?

Assessment can take place in the following areas:

- in the classroom with embedded, peer, self-, or classroom assessment techniques
- through formative assessment
- by benchmark or summative assessment

Who performs classroom, formative and summative assessment?

Despite the instructor of record being responsible for the final course grades, assessment can be performed by the following:

- instructor on record
- teaching assistants
- peers (only formative)
- students self-assess (only formative)

What are common approaches to assessment?

There are three main approaches to assessment:

1. Analytically (criterion-based assessment);
2. Holistically (impressionistic); and
3. Normative-based assessment.

Assessing analytically, developed over the past 50 years in higher education, requires a certain amount of pre-set criteria that is used to evaluate students' work (Sadler, 2009). The pre-set criteria can be determined by the instructor or together with students. Alternatively, in holistic or global grading, the instructor looks at the work as a whole and determines what grade should be assigned (Sadler, 2009). Both approaches have been challenged for their scorer reliability (consistency of grades between markers), but there may be other issues, such as the skills of the markers and if each approach can completely represent the whole assignment (Sadler, 2009). Normative-based assessment utilizes comparison to separate students, typically called 'grading on the curve' or 'bell curve', or norm-reference measurement (Aviles, 2001). At the University of Victoria, the use of normative-based assessment is strictly prohibited. The UVic calendar states:

> A primary purpose of evaluation and grading is to further effective teaching and learning. Any practices which assign a predetermined percentage of students a specific grade, that is, a certain percentage get A, another percentage get B and so on, without regard to individual achievement are prohibited.

Despite issues with all approaches, an **analytical assessment approach, criterion-based**, is currently being used due to its capacity to be more objective, reduce favouritism, increase transparency, provide more targeted and relevant feedback, and improve accountability (Sadler, 2009). However, Sadler (2009) points out that using analytic assessment can result in anomalies that render the pre-set criteria as insufficient for assessing the work. When this occurs, the instructor has to decide whether to adjust

and let students know, or not. But one of the foundations of analytic assessment is that it is transparent, whereas in holistic grading, the reasons why and how grades were assigned can be hidden. Therefore, Sadler (2009) argues for a combined approach, which he terms 'developing expertise,' that informs students of the criteria used but allows for informed academic judgement about the quality of the work (discussed below). Jackel, Pearce, Radloff, and Edwards (2017) call informed judgement by students 'assessment literacy' because it takes student knowledge about assessment beyond rubrics to really understand how assessment works, making assessment more transparent. At the University of Victoria, criterion-based assessment is the preferred approach to assessment.

What are the common problems with assessment design?

Within the decisions that can be made, the following are some of the common issues associated with assessment:

- inappropriate methods of assessment selected (timing, type, or not aligned with learning outcomes) (Pusateri, 2009)
- minimal assessment measures used (Pusateri, 2009)
- students don't have a chance to act on feedback (Pusateri, 2009)
- omission of formative assessment to support summative assessment (Pusateri, 2009)
- assessment methods used are not evaluated (Pusateri, 2009)
- instructors use assessment as a reward system by inflating grades (for those who put in more effort than others or showed marked improvement over the course) rather than assessing the quality of the task assigned (Sadler, 2009)

What is the purpose of summative assessment?

In general, summative assessment is viewed as a mechanism for student improvement. The term 'backwash' (Williams, 2014) has been used to describe the negative impact summative assessment can have on students, meaning that it may negatively affect and influence students rather than be beneficial and encourage further learning (Jackel, Pearce, Radloff & Edwards, 2017). Literature about the purpose[2] of summative assessment practices in higher education falls into four main categories: feedback, motivation, comparing students, and measurement of learning.

The first purpose states that summative assessment provides **feedback** to students about their learning and performance. Previously, feedback was seen as justification for a grade, whereas currently, feedback is seen as a mechanism to feed-forward. **Feed-forward** is feedback that is timely and constructive, with the aim to explicitly improve the next assignment in the course (Wimshurst & Manning, 2013). This 'new feedback' model is nicely summarized by Orsmond et al. (2013, p. 244):

Model of 'New Feedback'	
New Feedback Delivery	*Standard feedback delivery*
Encourages dialogue between giver and receiver of feedback	Monologue, often tutor-directed, one-way feedback
Involves peers	Does not involve peers
Explicitly encourages self-assessment/regulation	Does not explicitly encourage self-assessment/ regulation
Feedback on assignment process	Feedback on assignment product

[2] See Appendix 1 for a history of the purpose of assessment.

Students encouraged to be proactive in working with feedback	Students encouraged to be reactive in working with feedback

Summative assessment informs students about how well they performed, but is feedback on summative assessment useful for students' learning?

Two forms of feedback to students on assignments exist: evaluative and descriptive. Evaluative involves judgement in the form of a grade or as written feedback that commends students for their work or admonishes them for their lack of work. Descriptive feedback, by contrast, gives students information about how they can improve their work in the future. A study by Butler and Nisan (1986) shows the impact of the two forms of feedback and of no feedback. Their research determined that evaluative feedback after completion of a task does not improve student performance in the future, but descriptive feedback significantly improved similar subsequent task performance. An important aspect of descriptive feedback is that it cannot be accompanied by a grade. If it is, the grade will overshadow the comments (Brookhart, 2008, as cited in Schinske & Tanner, 2014). The most common problem with descriptive feedback is that students do not know what the instructor means (Weaver, 2006 as cited in Schinske & Tanner, 2014), highlighting the need for clarity and for a dialogic approach (Nicol, 2010). Therefore, feedback coupled with summative assessment does not provide useful feedback to assist students' learning. Useful feedback that can help students improve their work as they are learning comes in the form of detailed descriptive feedback combined with dialogue with the instructor or teaching assistant (TA) and no grade. These points argue for the necessity of formative assessments for all summative assessments.

The next claims about purpose is put forward by Guskey (2011), who prefers to approach the topic by looking at the misconceptions that exist about assessment today in our universities. One of the misconceptions addresses **motivation**—that poor grades prompt students to try harder. Guskey (2011) states there is no evidence to show that this statement is true, whereas there is research showing that low grades prompt students to withdraw and feel helpless to improve (Selby & Murphy, 1992).

The next misconception addresses **comparing students**. The first point is: *Grades should provide the basis for differentiating students,* and Guskey asks, "Is this what teaching is about? Isn't teaching about developing knowledge in students instead of just identifying those who have demonstrated they know the topic?" If the purpose is to develop knowledge of the topic in all students in the class, then the end result should be that most students do well in the course.

The second point in this category is: *Grade distributions should resemble a normal bell-shaped curve.* Guskey (2011) states that teaching is an intentional act that aims to have all students do well in courses. If the result at the end of the course is normal distribution, then a significant number of students did not learn the course content.

The last two misconceptions address **measuring student learning**: *(a) Grades should be based on students' standing among classmates.* The problem with this approach is that it does not give an indication of how well students have actually learned. All students may have performed poorly. The other problem with this approach is that it sets up a competitive rather than a collaborative learning environment. *(b) Students should receive a single grade for each subject or course.* Instructors often group diverse assessments together, such as a final exam, participation in class, and lab assignments.

This is what Guskey (2002) terms a 'hodgepodge grade' - combining diverse sources distorts the meaning of the grade and does not indicate clearly what a student has learned.

In summary, if summative assessment is to provide some measurement of student learning, then we need to take a close look at how that can be done effectively, efficiently, and sustainably so that students and faculty can engage in meaningful practices to promote student learning. Before moving on, we need to address two issues that are closely linked to improving assessment practice — grade inequality and grade inflation.

What is grade inequality and inflation?

Grade inequality is when students receive different grades for work that is of similar quality across departments, courses or sections (Aufderheide et al., 2016).

Grade inflation is considered to occur when instructors give higher grades to student work when compared to comparable work in previous years (so no increase in achievement to warrant a higher grade) (Hunt, 2008). Kohn (2008) argues that it is impossible to prove such a claim since there is not enough data to support it. Rosovsky and Hartley (2002) analyzed grading patterns in the United States from the 1960s to 1990s and found evidence of grades increasing on average a 0.5 letter grade (from C level to B level), a slight increase that could have been due to multiple factors. For example, some historical reasons account for the increase, such as ensuring male students did not have to join the wartime military service, changes in curriculum, more reliance on student evaluations, increase in student enrolment, and more contingent faculty (Rojstaczer & Healy, 2012). In addition, Rosovsky and Hartley (2002) found increased disparity between disciplines (humanities higher than natural sciences), and a significant increase in grades issued by Ivy League schools.

Many factors have been identified to affect grade inflation, such as: grading practices; how students perceive grades; external interpretation and pressure by employers, parents, and other institutions; consistency of grading practices across departments and course sections; and the meaning and purpose of grades (Aufderheide et al., 2016). Hodges (2014) states that grade inflation occurs when instructors don't have clear expectations for students.

Different approaches have been suggested to ensure that grades are fair and not inflated: for example, contextualizing grades on transcripts that provide details about the difficulty of the course. Unfortunately, this approach requires significant effort by instructors and departments to determine the levels of difficulty (Aufderheide et al., 2016). Institutions that tried contextualizing grades found that this type of information actually led to grade inflation because students tried to take the more lenient courses (Aufderheide et al., 2016). Another approach is to cap the number of A's per course. This approach sets up a sense of competition for students and that certain quotas have to be met for instructors (Aufderheide et al., 2016). A more successful approach is to set grading standards and guidelines in departments or schools that aim to eliminate grade inequality and inflation (Aufderheide et al., 2016).

Despite evidence that grades have increased over the years due to multiple factors, the real problem is in grade inequality. The same factors mentioned above that affect grade inflation also affect grade inequality, and some more than others. For example, an increase in contingent (sessional) faculty teaching at higher education institutions (who sometimes are without significant background in teaching, knowledge about grading and support from the department, coupled with the desire to

maintain employment) may result in fewer lower grades being issued to students (the same may apply to other inexperienced or vulnerable faculty, such as tenure-track) (Kezim, Pariseau & Quinn, 2005; Nikolakakos, Reeves & Shuch, 2012). Hodges (2014) claims that rubrics with clear criteria eliminate the many factors that affect grade inflation and inequality. Rubrics indicate the standards that students need to attain to receive the assigned grade, with the result being that there is more consistency in grading and a decrease in instances of students challenging grades (Hodges, 2014).

In summary, to lower incidences of grade inflation and grade inequality, the following is recommended:

- Ensure that all instructors are sufficiently familiar with department grading policies
- Encourage instructors to give students rubrics when the assignments are given so that it is clear how they are going to be evaluated and helps ensure consistency and reliability when grading
- Suggest that both formative and summative assessment are incorporated into each course
- Check that all assessment is valid and aligned with intended learning outcomes of the course

How can assessment practices be improved?

Kohn (2008) reminds us that learning is rarely mentioned in discussions about grading. Overall, literature about how to improve assessment practices focuses on the following areas:

- criterion-based assessment to be utilized (Sadler, 2009), which requires instructing all involved with grading in the course about how to satisfactorily use rubrics based on clear criteria
- clear, intended learning outcomes at the program, course, and module level that clarify good performance are developed
- rubrics are provided for all assignments to ensure grade consistency (Aufderheide et al., 2016; Hodges, 2014)
- proactive student role is encouraged rather than a reactive student role in assessment (Nicol & MacFarlane-Dick, 2006; Sadler, 2009)
- relevant examples of how to recognize the quality of work expected (students have to practice evaluating work themselves), so that students are aware of instructor's expectations (Sadler, 2009)
- exemplars are provided to help students peer-assess (Hendry, Armstrong & Bromberger, 2012)
- feedback from instructor and peers is scaffolded so that students can increase their meta-cognitive skills and eventually self-monitor the quality of their work (Sadler, 2009)
- multiple measures and sources that are implemented continuously to ensure fair representation of student learning (Pusateri, 2009)
- explicit instructions are provided for all assessment tasks (Pusateri, 2009)
- formative and low-stakes assessment that lead to an increase in self-regulated learning (Nicol & MacFarlane-Dick, 2006)
- self- and peer-assessment is used (Boud & Falchikov, 2006; Boud & Soler, 2016—term they use is sustainable assessment[3])
- assessment of authentic tasks (related to real world case studies, practice, portfolios)
- evidence-based approaches are taken to assessment (Joughin 2010)

[3] Sustainable assessment is "every act of assessment needs in some identifiable way to build students' capacity to manage and judge their own learning and thus equip themselves for the more challenging environments they will confront post-graduation" (Boud & Soler, 2016, p. 410).

- assessment practices are self- or peer-evaluated (Pusateri, 2009)
- the singular final grade is representative of what the student learned in the course (Guskey, 2002)

Keeping learning as the goal, Guskey and Baliey (2010) distinguished three distinct approaches that instructors can take in setting criteria for assessment: product, process, and progress, illustrating a more holistic approach to assessment (James 2006, 2012).

1. *Product criteria* – is used when instructors focus strictly on what students know as demonstrated in final products, such as examinations, reports and projects or other forms of culminating forms of learning.
2. *Process criteria* – is used by instructors who believe that product criteria do not provide the whole picture of learning that students undertook. How students learned the course material is an important contributor to the final grade. This means that the instructor values the learning process that students undertake and take this process into consideration. Process criteria therefore include formative assessments, homework, punctuality on assignments, attendance, class participation, and quizzes.
3. *Progress criteria* – focuses on how much students gain through their learning experience, which makes this approach very individualized. In this approach, a learning continuum is what a student's progress is measured against in order to see how much they have progressed in their learning (mastery approach).

Guskey and Bailey (2010) suggest that instructors using product, process, and progress approaches together, establish explicit criteria for each to provide a more accurate picture of what a student has accomplished and their academic achievement.

Schinske and Tanner (2014) concur with Guskey and Bailey (2010) regarding emphasizing and rewarding student processes in learning. In large classes, this can be done through iClicker participation marks, classroom assessment techniques (CATs) that have students complete one-minute papers or other CATs, homework that has students complete meaningful assignments and contribute to summative assignments, opportunities to practice some aspect and reflect on it, and more. These assignments can be peer-graded, marked for being turned in, or that they wrote the right amount of words or other criteria. These types of assignments can incentivize students, promote practice, and support student learning through various activities.

To assist with making the descriptive feedback process more efficient for the instructor, and develop self-regulated learning and metacognitive skills in students, self- and peer-assessment can be used by providing opportunities for students to self-reflect and to encourage dialogue with colleagues about their work (Nicol and Macfarlane-Dick, 2006). To ensure high quality of self- and peer-assessment, the following mechanisms can be employed: exposure to high quality work; comparison of their work to high quality work; and learning how to adjust their work to increase quality (Sadler, 2009).

These ideas resonate in the work of McDowell, Sambell and Davison (2009) who argue for a more integrated approach to assessment that includes an environment rich in feedback, the active participation of students leading to student autonomy, a reduction in summative assessments, and authentic tasks and assessment.

Links to key documents at UVic

UVic undergraduate grading scale

UVic graduate grading scale

SAS statistical software

For information related to Brightspace and online assessment, please see Teach Anywhere.

Questions for discussion within units for shared understanding of grading practices and their relation to students' learning

- Have all instructors received UVic's grading policies and procedures?
- Are all instructors aware of the department's approach to grading and the purposes of grading?
- What issues presently exist within or across courses?
- What anomalies exist and is the rationale clear as to why?
- Do all instructors use formative and summative assessments? Rubrics that are shared with students? Clear assignment instructions?
- Are assessment methods aligned with the intended learning outcomes?
- Is the weighting of the different methods appropriate to the importance of the outcomes and the time that students will spend on learning?
- Are the assessment methods distributed over the course in order to reduce stress and provide feedback to learners on progress towards achieving the intended learning outcomes?
- Will the assessment methods chosen reflect what the student has learned in the course?
- Are a range of methods used to adhere to Universal Instructional Design principles?

Recommended practices for multi-section courses to achieve equitable assessment and grading practices

Multi-section courses are unique in that they require extensive organization, as they involve several instructors and teaching assistants (TAs) across many sections to assure successful implementation of positive student learning. Therefore, recommendations for consistency, consideration, and collaboration are discussed to help ensure equitable assessment and grading practices for all students within multi-section courses. Not all suggestions will be relevant for every course/program, but thorough evaluation of each, including identifying potential limitations, will assist in employing the best recommendations for successful practice. See Appendix 4 for the full paper. For an annotated bibliography related to this topic, see this link: https://researchsotl.wordpress.com/2018/06/28/multi-section-grading-practices/

Best practice toolkit for instructors

Provide the following information (also provided as a separate document) to instructors in your department to help guide them to implement best practices in grading.

General

- Include definitions, types and examples of all forms of assessment in the syllabus: pre-assessment, classroom, formative, post-assessment, and summative.
- Show examples of alignment between all assessment forms, and alignment between intended learning outcomes, instructional strategies, and assessment.

Summative

- Use Track Changes when assessing papers and only correct an error once. You can highlight the same error if it occurs again and direct the student to the previous comment (Smith & Palenque, 2015).
- Create a comment bank arranged by topics, such as grammar, content, organization, so that you can quickly access when assessing (Smith & Palenque, 2015).
- Resist using red pen. Red signifies 'danger' 'stop' 'wrong'. If assessing on paper or computer, use blue or green ink for comments (Smith & Palenque, 2015).
- Avoid giving too much feedback. Receiving back a paper or assignment riddled with comments is overwhelming for students and detrimental to his or her learning. Instead, focus only on the most important points that the student needs to take into consideration to make his or her paper better (Smith & Palenque, 2015).
- Do not assess on the curve. Let students know that they are not in competition with the rest of the class and that they can achieve a good grade by doing the work required. If a large number of students perform poorly on certain questions in an exam, then give students another opportunity to resubmit those questions for credit (Schinske & Tanner, 2014).
- Do not repeat feedback, if students don't apply feedback provided on a draft in the final paper or assignment. Simply point the student to the comments provided for the draft (Smith & Palenque, 2015).

Formative

- Have students submit a draft so that you can provide feedback that the student can apply. Think strategically, though. If the student needs to make major revisions due to issues regarding the topic, then don't provide comments on grammar and organization since the paper will change significantly (Smith & Palenque, 2015).
- Ask students, before handing in a final paper or assignment, to write to you, the instructor, a letter on the back that begins with Dear Dr. (Name Here), and then describes the main point in their paper and how they think they did on it (Svinicki & McKeachie, 2011). When this method is used, most students self-declare if they did well or not and why. This provides an opportunity for students to enter into a conversation with you about their work and to share responsibility in assigning the grade.
- Ensure that your comments make sense and are understandable by the student. Avoid making vague statements or posing questions that the student may not be able to interpret what you mean. Remember that the purpose of feedback is to help students learn (Smith & Palenque, 2015).

- Give students rubrics (see Appendix 4) when the assignments are given so that it is clear how they are going to be evaluated (Smith & Palenque, 2015).
- Use Classroom Assessment Techniques (CATs) so that you can observe learning taking place in your classroom and notice when learning is not taking place (Angelo & Cross, 1993).
- Develop student expertise, called 'assessment literacy,' so that students really understand how assessment works (Jackel, Pearce, Radloff, & Edwards, 2017). Sadler's (2009) model is as follows:
    1. Expose students to a variety of works within the genre in which they will be working, which can include previous students' work.
    2. Students need to see a spectrum of poor to excellent quality work.
    3. Students need to see responses from an instructor to assignments.

  The learning process is to have students develop criteria and expand on it as they become more adept at judging the work. Especially important is to have students get to the stage where they can manifest latent criteria. Students often begin with mechanics, such as grammar, punctuation, referencing style, and organization. Sadler (2009) provides an example of how to do this: students need to submit formative tasks, such as "extrapolating, making structural comparisons, identifying underlying assumptions, mounting counter-arguments or integrating elements" (p. 12). In tutorials, students appraise peers' work, provide feedback and discuss the process. Sadler (2009) terms this "produce and appraise" rather than "study and learn" (p. 12).

See Appendix 2 for more examples of summative and formative assessment.

Classroom Assessment Plan (to ensure quality assessment)

The classroom assessment plan (CAP) helps an instructor scaffold assessment so that students are first taught what is expected and how to recognize quality work; then instructor/TA modeling of feedback; then peer feedback so that they can practice doing it followed by feedback on feedback from instructor/TA; and finally, self-monitoring.

Who, where and how, in order of occurrence	**In the classroom**	**Formative assessment**	**Summative assessment**
Instructor on record	CATs, such as the one-minute paper, muddiest point, etc.	Feedback on low-stakes formative assessment pieces that students submit (like those suggested above by Sadler, 2009)	Final assignments that had parts submitted for formative assessment but now all components are together in the final assignment
Teaching assistants	CATs in tutorials and labs	Practice opportunities with ample feedback	
Once instructor and TA have modeled feedback, move onto:			
Peer-assessment	Of upcoming assignments, ongoing practice activities	Feedback provided on peers' assessment	
After feedback received on peer-assessment, move onto:			Students write a reflective piece about their own progress in the course
Self-assessment	About progress and work in the course; reflective activities	Feedback from TA or instructor about progress	

Assessment framework – see Appendix 3 for examples of learning outcomes and assessment alignment.

Series of professional development sessions to assist instructors with assessment practices (see LTSI workshop listings on **learning central** for dates and times)

There are many issues related to the practice of assessment. This series will address the different facets of grading, such as formative and summative assessment, grade inflation, and aligning assessment with learning outcomes. If instructors attend the whole series (all four sessions), they will be awarded a certificate of completion – a great addition to their dossier indicating their commitment to professional development and to enhancing their performance as an instructor. The Division of Learning and Teaching Support and Innovation at UVic typically offers the following topics each fall and winter term.

Rubrics: What they are, why they are useful, and how to create them for your course assignments

Rubrics are a useful framework to ensure consistency when grading assignments. Ideally, rubrics are designed at the same time as assignments are but can also be developed from the criteria stated for an assignment. In this session, we will discuss the various forms that rubrics can take, the many advantages of using rubrics, and how to create well-designed rubrics that will provide consistency in grading. We will also demonstrate how we use rubrics to provide clarity, both for ourselves and our students, in terms of staging learning throughout the course, and offering clear expectations of learning goals in specific situations. You will have a chance to design a rubric for an assignment in your course and receive feedback.

High-stakes assessment of students' performance: The challenges of summative assessment

How can one make a final decision about a student's achievement in a course upon its completion? Find a high-stakes mechanism by which a student's final performance can be measured with respect to that of other students or against some standard or benchmark? In this workshop, we will consider the challenges associated with summative assessment, including the alignment of summative grades with the course intended learning outcomes and the design of assignments and tests - in particular, multiple-choice tests.

Making assessment enjoyable: Is that possible?

Formative assessment can alleviate some of the challenges associated with summative assessment. Formative assessments are the 'practice' aspects of assessment that prepare students for the exams and final papers where they receive grades. A common concern when considering implementing formative assessment into a course is that it will take too much of an instructor's time. In this session, we will discuss the significance of formative assessment, how it aligns with summative assessment, and how the time invested in formative reduces the time spent on summative. We will also work on how to implement formative assessment, such as forms of low-stakes assignments and scaffolding assignments that might actually make assessment enjoyable!

Fostering accessible learning for all students: Universal Instructional Design

Universal Instructional Design (UID) principles help create inclusive learning environments that benefit all students and instructors. UID is a key strategy for designing accessible and vibrant courses. It also reduces the need for course modifications to address accommodation requirements. This interactive workshop will explore how post-secondary classrooms can respond to a broad range of learning needs. We will introduce some principles of universal instructional design, explore how these principles can be applied in learning environments, and discuss how to address challenges that may arise

References

Angelo, T. A. & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*, (2nd Ed.), San Francisco: Jossey-Bass.

Aufderheide, P., Brannon, K., Connaughton, V., Dulaney, C., Figley, P., Freeman, D., Krishnan, G., McCurdy, H., Miller-Idriss, C., Oakes, S., Porzecanski, A., & Sonenshine, R. (2016). Final report of the ad hoc committee on grade inflation. American University. Retrieved September 12, 2018: https://www.american.edu/facultysenate/upload/Senate-report-on-Grade-Inflation-Final-10-20-16.pdf

Aviles, C. B. (2001). Grading with norm-reference or criterion-referenced measurements: To curve or not to curve, that is the question. *Social Work Education, 20*(5), 603-608. DOI: 10.1080/02615470120072869

Boud, D. & Falchikov, N. (2007). Introduction: Assessment for the longer term. In Boud, D. & Falchikov, N. (Eds.) *Rethinking Assessment for Higher Education: Learning for the Longer Term*, 3-13. London: Routledge.

Boud, D. & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education, 41(3),* 400-413, DOI: 10.1080/02602938.2015.1018133

Guskey, T. R. (2011). Five obstacles to grading reform. *Educational Leadership, 69(3*), 16.21.

Hendry, G. D., Armstrong, S. & Bromberger, N. (2012). Implementing standards-based assessment effectively: Incorporating discussion of exemplars into classroom teaching. *Assessment and Evaluation in Higher Education, 37(2),* 149-161.

Hodges, L. (2014). Demystifying learning expectations to address grade inflation. *College Teaching, 62*(2), 45-46.

Hunt, L. H. (Ed.) (2008). *Grade inflation: Academic standards in higher education*. New York: State University of New York Press.

Jackel, B., Pearce, J., Radloff, A. & Edwards, D. (2017). *Assessment and feedback in higher education*. York, UK: The Higher Education Academy.

James, M. & Lewis, J. (2012). Assessment in harmony with our understanding of learning: Problems and possibilities. In J. Gardner (Ed.), *Assessment and Learning*, (2nd Ed.), 187-205. London: Sage.

Joughin, G. (2010). The hidden curriculum revisited: A critical review of research into the influence of summative assessment on learning. *Assessment & Evaluation in Higher Education, 35*(3), 335–345.

Kezim, B., Pariseau, S. E. & Quinn, F. (2005). Is grade inflation related to faculty focus? *Journal of Education for Business, 80*(6), 358-364.

Kohn, A. (2008). The dangerous myth of grade inflation. In L. H. Hunt (Ed.), *Grade Inflation: Academic Standards in Higher Education*, pp. 1-11. New York: State University of New York Press.

Knight, P. (2001). *A briefing on key concepts: Formative and summative, criterion and norm-referenced assessment*. York: Learning and Teaching Support Network Generic Centre.

McDowell, L., Sambell, K. & Davison, G. (2009). Assessment for learning: A brief history and review of terminology. In *Improving Student Learning Through the Curriculum*, 56-64. Oxford: Oxford Centre for Staff and Learning Development.

Nicol, D. J. & Macfarlane-Dick, D. (2007). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199-218, DOI: 10.1080/03075070600572090

Nikolakakos, E., Reeves, J. L. & Shuch, S. (2012). An examination of the causes of grade inflation in a teacher education program and implications for practice. *College and University, 87*(3), 2-13.

Orsmond, P., Maw, S. J., Park, J. R., Gomez, S. & Crook, A. C. (2013). Moving feedback forward: Theory to practice. *Assessment and Evaluation in Higher Education, 38(2),* 240-52.

Pusateri, T. (2009). The assessment cyberguide for learning goals and outcomes. *American Psychological Association Education Directorate.* Accessed August, 2018: https://www.apa.org/ed/governance/bea/assessment-cyberguide-v2.pdf

Rojstaczer, S., & Healy, C. (2012). Where A is ordinary: The evolution of American college and university Grading 1940-2009. *Teachers College Record, 114*(7).

Rosovsky, H., & Hartley, M. (2002). *Evaluation and the Academy: Are we doing the right thing?* Cambridge, MA: American Academy of Arts and Sciences.

Sadler, D. R. (2009). Chapter 4: Transforming holistic assessment and grading into a vehicle for complex learning. In G. Joughin (Ed.), *Assessment, Learning and Judgement in Higher Education*, 1-15. DOI: 10.1007/978-1-4020-8905-3_4.

Schinske, J. & Tanner, K. (2014). Teaching more by grading less (or differently). *CBE Life Sciences Education, 13*(2), 159-166, doi: 10.1187/cbe.CBE-14-03-0054.

Smith, V. & Palenque, S. M. (2015). Ten tips for more efficient and effective grading. *Faculty Focus:*

Higher Ed Teaching Strategies from Magna Publications.

Svinicki, M. & McKeachie, W. J. (2011). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers,* (13th ed.).Belmont, CA: Wadsworth.

Wiggins, G. (1998). Ensuring authentic performance. In G. Wiggins (Ed.) Educative Assessment: Designing Assessments to Inform and Improve Student Performance, (pp. 21 – 42). San Francisco: Jossey-Bass.

Wimshurst, L. & Manning, M. (2013). Feed-forward assessment, exemplars and peer marking: Evidence of efficacy. Assessment and Evaluation in Higher Education, 38(4), 451-65. http://doi.org/10.1080/02602938.2011.646236.

Appendix 1: History of assessment in higher education in North America

Beginning in 1646, Harvard required exit exams before a degree could be awarded (Smallwood, 1935, as cited in Schinske & Tanner, 2014). Yale in 1785 had the first official record that grades were recorded for seniors in four categories: Optimi, second Optimi, Inferiores, and Perjores. (Stiles, 1901, cited by Smallwood, 1935, in Schinske & Tanner, 2014). That increased to recording grades (not accessible to students, though) for each course in 1837. At this same time, some professors at Harvard were using a 100-point system, but most institutions kept no records at all.

With the increase in schools in the late 19th century, the need arose to be able to coordinate between schools. Some schools had adopted an A to F recording system that included E, while others adopted the 100-points percentage system. Wide variability existed within these systems, and in the early 1900s, there was an effort to understand and find an assessment system that could be uniform between institutions. They found that the 100-point system was highly unreliable and that the five letter category was more reliable. By the 1930s, E was dropped and F was kept to represent Fail (Schinske & Tanner, 2014).

At this same time, research about intellectual ability influenced the idea of a normal curve. It was posited that if a curve existed for intellectual ability, then the same must occur in the classroom, which neatly addressed the issue of subjectivity associated with assessment (Schinske & Tanner, 2014).

The pressure to find uniformity in assessment systems led to the dominance of the A-F assessment systems in the 1940s, but other systems still prevailed at some institutions. Assessment systems and whether to grade or not are still highly contested issues. Presently, some institutions have eliminated numerical and categorical assessment completely and, instead, enter into contracts with students or require students to reflect on their learning. As the history above shows, assessment was developed primarily for communication and for organizational purposes within and between institutions and had no pedagogical value (Schneider & Hutt, 2013, as cited in Schinske & Tanner, 2014).

Appendix 2: Potential forms of formative and summative assessment (not an exhaustive list)

1. Artefacts/products, especially in fashion, design, engineering, etc.
2. Bullet point summaries
3. Assessment banks. Students have access to a question & answer bank. They learn how to answer all of them but are assessed on a sample
4. Assessment of work-based learning (in a variety of ways, many times, by a variety of people, for different purposes)
5. Book, website or program reviews
6. Classroom assessment techniques. They are *brief* tasks that tell the teacher something about the *class's* grasp of the material. (See Angelo and Cross, 1993)
7. Completing structured summaries of readings, debates, etc.
8. Computer-based self-assessment
9. Concept maps. Excellent way of seeing how students understand complex content and relationships
10. Contribution to threaded electronic discussions
11. Defence of lab records
12. Design and build (similar to 2, above)
13. Dissertations and theses
14. Electronic monitoring of web searches, program use and communications
15. Essay writing broken down into and submitted in chunks until complete – topic, annotated literature review, thesis, etc. until paper is complete
16. Small writing pieces
17. Exhibitions of work, posters, products. History students have curated museum exhibitions in lieu of doing a dissertation
18. Field work and lab work assessment
19. Formative assessment of logs/journals/portfolios (when the purposes are formative, students identify areas for discussion. If summative, sampling within the logs, etc., is recommended, especially if students know in advance the areas that are likely to get closest attention.20. Games and simulations
21. 'General' assessments, drawing together learning in several modules
22. Making annotated bibliographies for next year's students
23. Making models (literally, in some subjects, conceptual models in others)
24. Multiple choice questions (they do not have to be only tests of information, although it is a lot quicker to write MCQs like that. See also 3, above)
25. New tests in which learners use old software/programmes/notes
26. Objective Structured Clinical Examination (OSCE)
27. Open-book, end of course exams
28. Orals
29. Peer assessment
30. Performances. Vital in the assessment of competence. Note massive problems assessing complex performances fairly and reliably. Simulations sometimes possible.
31. Personal response assessments. Usually done in classes where each student has an electronic response pad. Teachers ask questions and they press a key to show their answer. Can be used for classroom assessment or test purposes.
32. Posters
33. Production of structured logs of project/dissertation progress and reflection on it

34. Projects
35. ‘Real’ problem working, which involves defining ‘fuzzy’ situations, bringing some order to ill-defined issues, analysing the problem and suggesting solutions
36. Replication of published inquiries
37. Role-playing
38. Self-assessment
39. Seminar presentations (in or out of role; with or without use of video, PowerPoint, etc.)
40. Short answer questions. (MCQs *plus* some explanation of the thinking: limit to 100 words per response?)
41. [Short] appraisals of target papers
42. Small-scale research or enquiry
43. Statements of relevance, which are short pieces of writing (1000 words, perhaps), making claims about the relevance of a workshop, article, field observation, etc., to another task or activity
44. Submission of claims to achievement, with reference to portfolio (if this is to be summative, I suggest grading on the claim alone, provided that sufficient evidence supports the claims)
45. Takeaway papers/questions/tests
46. Terminal, unseen examinations and other individual, time-constrained assignments
47. Two-part assessments. Elements of a task are formatively assessed, but the final product is summatively assessed
48. Web page creation
49. Writing exams/tests/assessments to tutor specification
50. Writing memoranda or journalistic summaries

Knight (2001)

Appendix 3: Sample learning outcomes and table to align assessment

Extract from a social research methods course handbook

Intended Learning Outcomes

By the end of the course, you will:

1. Demonstrate knowledge of mainstream educational and social research methods
2. Critically engage with issues concerning the relationships between research and knowledge, and the fitness of different research methods for different purposes
3. Skillfully read and evaluate research reports
4. Design a feasible, small-scale research inquiry
5. Organize and manage your own learning
6. Work effectively with others, both to their benefit and yours
7. Present your conclusions orally to an audience

This table links these outcomes to the main pieces of work to be done.
(S) stands for summative and (F) stands for formative assessment.

Assessment tasks	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Coursework 1 – submit a topic (F)	x				x		
Coursework 1 – submit example of annotated literature review (F)	x	x	x		x		
Coursework 1 – peer review of annotated literature review (F)	**x**	**x**	**x**		**x**	**x**	
Coursework 1 – write a 200 word literature review (S)	**x**	**x**	**x**		**x**		
Coursework 2 – submit draft research inquiry for peer review				**x**	**x**	**x**	
Coursework 2 – design a research inquiry (S)				**x**	**x**		
Coursework 3 – write an evaluation of a published research paper (S)	**x**	**x**	**x**		**x**		
Submit bullet point lists related to each week's set readings (F)					**x**		
Coursework 4 – submit reflective piece about your organization, management, and participation in course (S)					**x**	**x**	
Oral presentation of work (S)	**x**	**x**			**x**	**x**	**x**

Adapted from Knight (2001)

Appendix 4: Recommendations for Equitable Assessment and Grading Practices in Multi-Section Courses

By Alisha Duwyn, Brock University

** This document was prepared and written by Alisha Duwyn as part of the course EDUC 5P51, Department of Graduate and Undergraduate Studies in Education, Faculty of Education, Brock University. Ms. Cynthia Korpan, LTSI Professional Development Programs, University of Victoria, asked Ms. Duwyn to write this paper as part of her course. Ms. Duwyn gave her permission to include the paper in this document.*

Assessment is a key aspect of instruction as there must be a method to evaluate students and produce data to identify their level of achievement. Assessment tasks are often determined by instructors but can be mandated by the faculty or university who have specific policies to be followed (Lingnan University, 2017). Assessment and grading can become complex when a course has a large class size that must be taught by several instructors who teach different sections and/or several teaching assistants (TAs) who often facilitate seminars with smaller groups of students (Glazer, 2014). These classes are referred to as multi-section courses since all students are not in one class together being assessed by the same person (Glazer, 2014). As a result, disparities can exist in assessment and grading practices when several instructors/TAs are involved with the same course. This brings forth the issue of how one ensures equitable treatment for all students in multi-section courses.

According to CUPFA (2015) “equitable treatment includes similar course content, similar evaluative exercises, similar grading schemes” (p. 1). This does not necessarily mean that each student is treated the same but that comparable practices are utilized no matter the individual that is being assessed. As a result, it is important to review the literature to provide recommendations towards achieving equitable assessment and grading practices for multi-section courses. I will further outline potential limitations and implications for these recommendations including areas for further research within this topic.

Background

Assessment can be identified based on the format such as qualitative assessment which includes written papers and essays and quantitative assessment usually in the form of math problems and lab reports (Knight, Allen, & Mitchell, 2012). Assessment can also be distinguished by when it is practiced such as formative assessment, which occurs throughout a course and is meant to provide students with feedback, versus summative assessment, which is focused on end of course learning outcomes (Lingnan University, 2017). Both are often included in a student's final grade however formative might consist of lab reports and presentations and summative might be a final written exam or essay (Glazer, 2014).

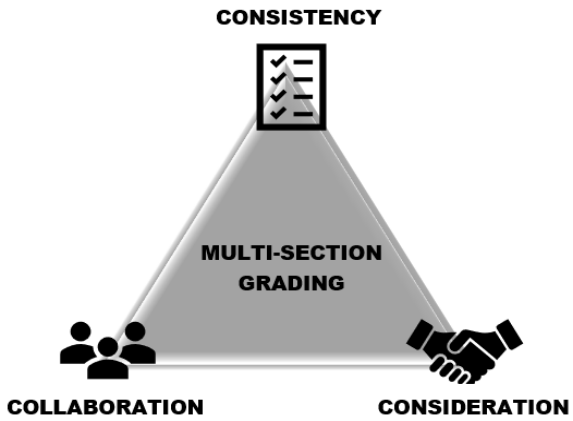
As with all courses, there are several factors that must be strictly followed based on university and/or faculty policies. These would be found in the course syllabus, the faculty's webpage, and the university's academic calendar and would include plagiarism, grade conversions, late assignments (which can be up to the instructor), special needs, and accommodations (i.e., religious, medical). On the other hand, factors that are relevant to multi-section courses (and even all courses) and would often be at the discretion of the instructor(s), consist of the course outline, learning outcomes/objectives, marking criteria, textbook, exam format (if included), and any assignments (include labs, seminars, discussions etcetera) (Adams, n.d.; CUPFA, 2015).

Recommendations

Research on assessment and grading within multi-section courses primarily focuses on large first-year undergraduate classes. After reviewing the literature and identifying recommendations within each study, I have divided the themes into three main categories of consistency (Glazer, 2014; Lingnan University, 2017; O'Neill, Birol, & Pollock, 2010, Preston & Holloway, 2006), consideration (Gladding, 2007; Lingnan University, 2017; Roth & Sandford Bernhardt, 2016), and collaboration (CUPFA, 2015; Knight et al., 2012; Perrin, Rusnak, Zha, Lewis, & Srinivasan, 2009) that are essential in ensuring that equitable assessment and grading practices are developed within multi-section courses (see Figure 1). This does not mean that every idea within each category must be fulfilled, as this will depend on the course and

program, but it is important to evaluate and reflect on each. Consistency involves ensuring that procedures and practices are similar across all sections and are being effectively implemented by all stakeholders (supervisors, instructors, TAs etcetera). Consideration is important when discussing fairness for both the instructors/TAs who are administering and assessing the course material as well as for students who are being evaluated. Lastly, collaboration is an integral element for multi-section courses as all stakeholders must cooperate and work as a team to assure equitable assessment and grading practices.

Figure 1. The 3 Cs for equitable assessment and grading practices in multi-section courses



Consistency

Review of the literature reveals that consistency is key for assessment and grading practices across multi-section courses. Consistency involves ensuring that all aspects of a course are similar so that students are being assessed in an unbiased way. This can be accomplished by developing consistent learning outcomes across all sections (Lingnan University, 2017). These learning outcomes need to be communicated to all students at the beginning of a course and should be aligned with assessment tasks (Lingnan University, 2017; O'Neill et al., 2010). As well, all sections should be developed under one common course syllabus with standardized assessment tasks and course objectives (Perrin et al., 2009; Preston & Holloway, 2006). Glazer (2014) found that the inclusion of both formative and summative assessment tasks for a first-

year chemistry class improved grading consistency across sections which advocates for the importance of both forms of assessment.

When evaluating assessment tasks, it is important to ensure consistency among instructors/TAs. This will allow students to feel confident that no matter who is grading their work, equitable practices are followed. Standardization of exam questions across all sections of a course are important for consistency as well as to ensure that exam questions reflect course objectives and material learned (O'Neill et al., 2010). For qualitative assessment tasks such as written essays, marking criteria should be standardized, provided to students well in advance, and aligned with learning outcomes (Lingnan University, 2017). Several studies have looked at the benefits of using marking rubrics with multi-section courses and found that they helped to reduce variability in grades (Knight et al., 2012), improve consistency in feedback (Glazer, 2014), and reduce marking time (Roth & Sanford-Bernhardt, 2016). Marking rubrics identify a specific set of criteria for evaluation and can be utilized across all subjects including essays; however, students must be made aware of the grading criteria prior to completion of the task.

Consideration

While consistency is essential, consideration involving fair practices for both instructors/TAs and students is also significant when discussing equitable assessment and grading. Gladding (2007) states that lectures and class material should be the same across all sections of a course. Although uniformity is important for student learning in multi-section courses, autonomy and choice for instructors/TAs is also noteworthy as they have their own areas of interest and specialisms that guide their learning and should be considered when developing assessment tasks. To promote teacher's academic freedom, O'Neill et al. (2010) suggest that a few learning outcomes can be tailored to the individual instructors, but these need to be clearly identified to students. Due to this, some questions on assessments could be personalised by the instructor but this opportunity would need to be available to all instructors of a course to ensure fairness.

Other recommendations centered around fairness are more specific to grading and the importance of consulting with all instructors/TAs before evaluating student work. One way is

for individuals to each mark a set of samples separately and then compare to assess grading consistencies or lack thereof (Roth & Sandford Bernhardt, 2016). This would be relevant for courses with qualitative written responses such as English and social science courses. Another suggestion is to have one individual mark a specific set of questions on all exams/assignments to help ensure accurate grading (Lingnan University, 2017). Rubrics, as previously mentioned, are also a good way to ensure fairness in grading across multi-sections, but it is important to consider instructor/TA effectiveness which would require discussions to assess accurate grading practices.

Collaboration

Collaboration among all stakeholders is pivotal in the effective organization and running of multi-section courses. Collaboration cannot be possible without effective communication and cooperation between all individuals and must be maintained throughout the entire course from start to finish (Perrin et al., 2009). First, this must involve routine meetings with all instructors/TAs (Preston & Holloway, 2006; Sundberg, 2002; Williams, 2017) as well as with directors/supervisors who oversee the course (Knight et al., 2012). These meetings should occur throughout the year and allow for input on aspects of course design including developing the course syllabus and learning outcomes as well as monitoring for consistent grading and evaluation of assessment tasks. Another important aspect of collaboration within assessment is that all instructors need to contribute to the creation and marking of assignments (exams, essays, projects, reports etc.) and need to be given sufficient time and opportunity to provide feedback and revision (CUPFA, 2015).

Other Recommendations

There are some recommendations within the literature that indirectly relate to assessment and grading but benefit the overall organization of multi-section courses. Perrin et al. (2009) suggest developing an administrative structure that involves appointing a director who would allocate course supervisors responsible for the monitoring of instructors/TAs for each section. This type of hierarchal structure can be very effective for large universities with several multi-section courses. Even if a director is not necessary, having a course supervisor or a head instructor can assist in effective communication.

O'Neill et al. (2010) recommend that a resource area should be created for all instructors/TAs which includes "instructors guide to learning outcomes, specific examples, active learning activities, clicker and exam question banks" (p. 19). By cooperating with all individuals responsible for each course and creating a common resource area, equitability will be fostered as well as the sharing of good practice. These resources would be available to instructors/TAs in all sections and would also help to ensure consistency amongst course material.

As well, many universities utilize an online learning management system (LMS) to organize and disseminate course material but it is usually up to the instructor(s) to decide on the level of usage for their course. Yalcin, Gardner, Anderson, Kirby-Straker, Wolvin and Bederson (2015) found that using an LMS was beneficial for multi-section courses as online learning communities promoted instructor/TA interaction and each section could more easily be monitored by the course supervisor/head instructor.

Limitations

Developing equitable assessment and grading practices for multi-section courses also presents challenges including budget limitations, time management, staff turnover, training requirements, and technological advancements which must be considered when implementing any changes within a course (Glazer, 2014; Perrin et al., 2009). Challenges can arise with having a large class size and multiple instructors for each course regarding the time needed for effective training (Sundberg, 2002). This makes it difficult to implement similar grading practices across sections and if staff turnover is high, then new individuals must be trained frequently which requires resources and money (Lignan University, 2017).

There is also a degree of bias that is present with subjective grading of essays and qualitative written responses with courses like English (Glazer, 2014). However, with implementing best practices such as by using rubrics and group marking practice, this subjectivity can be diminished as much as possible.

Grades are very important to students as they often dictate scholarships and admittance into further studies such as master's programs (Knight et al., 2012). Students may become frustrated if they feel that grades are being unevenly disseminated without any explanation for the differences (Knight et al., 2012). Therefore, equitability must be assessed across grade distributions and years to prevent preferential choice of one course/instructor over another (Glazer, 2014; Knight et al., 2012). Courses must be frequently assessed usually overseen by a supervisor to assure grade alignment is unbiased.

Implications

When considering implications regarding equitable assessment and grading practise in multi-section courses it is important to look at the impact of the recommendations described above in three areas consisting of practice, theory and future research**.**

Practice

Equitable assessment and grading practices promote positive student learning and are important for multi-section courses where several stakeholders are involved in the evaluation of student work. Equitable practices across all sections are vital to ensure impartiality and transparency (Lingnan University, 2017). Applying good practices involving multi-section grading will involve implementing recommendations based on the needs of the institution. This may result in modifications to individual courses/sections or possibly a university wide evaluation. Ultimately, the degree to changes that would be made to courses would be mandated by policies that must be followed (CUPFA, 2015).

Theory

Recommendations regarding multi-section grading have been centered on the 3Cs of consistency, consideration and collaboration (see Figure 1). These three areas have been created based on the desire to ensure equitability across all sections of a course concerning the assessment of tasks which will determine a student's overall grade. It is important to remember that recommendations need to be evaluated for their individual applicability. For example, an institution may assess that there is excellent collaboration between instructors/TAs but discover that there is no consistency or consideration between their assignments. As well, if it is determined that gaps exist within each category, suggestions may be implemented across all 3C's to improve equitable practices.

Future Research

Research on assessment and grading in multi-section courses has focused on standardization across all sections; however, there are several areas requiring further study. Glazer (2014) noted that effective feedback was important for student learning but stated that many instructors struggle with providing feedback and this area needs to be explored. Instructors/TAs must balance using similar assessment methods with all students while still providing personalized feedback. Ultimately, assessment is only effective and relevant if it is used by the student and instructor to inform further practice (Glazer, 2014). Analyzing learning outcomes and identifying gaps/discrepancies between outcomes and grades is important for student success and universities need to develop procedures to assess their own courses (Jefferson College, 2009; O'Neill et al., 2010). Lastly, as technology continues to advance it is important to consider the impact on multi-section courses and the influence this may have on assessment and grading practices.

Conclusion

Multi-section courses are unique in that they require extensive organization as they involve several instructors/TAs across many sections to assure successful implementation for positive student learning. Therefore, recommendations for consistency, consideration, and collaboration will help ensure equitable assessment and grading practices for all students within multi-section courses. Not all suggestions will be relevant for every course/program but thorough evaluation of each including identifying potential limitations will assist in employing the best recommendations for successful practice.

See also, the following annotated bibliography:
https://researchsotl.wordpress.com/2018/06/28/multi-section-grading-practices/

References

Adams, D. (n.d.) [Review of the book *Effective grading: A tool for learning and assessment,* by B.E. Walvoord & V. Johnson Anderson]. Retrieved from

http://cla.auburn.edu/cla/assets/docs/assessment/Summary%20of%20Walvoord%20and%20Anderson%20Effective%20Grading.pdf

CUPFA. (2015, April 25). Best practices - multi section courses. *Concordia University Part-time Faculty Association.* Retrieved from http://www.cupfa.org/best-practices-for-multi-section-courses/

Gladding, G. (2007). Reforming large courses at a research university: A case study from physics at Illinois. Retrieved from http://www.cwsei.ubc.ca/Files/Gladding_talk.ppt

Glazer, N. (2014). Formative plus summative assessment in large undergraduate courses: why both? *International Journal of Teaching and Learning in Higher Education, 26*(2), 276-286.

Jefferson College (2009, February). Assessment of multi-section courses. *Assessment Update.* Retrieved from http://vega.jeffco.edu/szak/handouts/CorrectedAssessmentNewsletterFeb09.pdf.

Knight, J., Allen, S., & Mitchell, A. M. (2012). Establishing consistency measurements of grading for multiple section courses. *Journal of the Academy of Business Education, 13*, 28-47.

Lingnan University (2017, March). Assessment guidelines for Lingnan University. *Registry.* Retrieved from https://www.ln.edu.hk/reg/docs/assgud.pdf.

O'Neill, A., Birol, G. & Pollock, C. (2010). A report on the implementation of the blooming biology tool: aligning course learning outcomes with assessments and promoting consistency in a large multi-section first-year biology course. *The Canadian Journal for the Scholarship of Teaching and Learning, 1*(1), 1-24.

Perrin, K., Rusnak, L., Zha, S., Lewis, D. & Srinivasan, S. (2009). Using blended learning to ensure consistency and quality in multiple course sections. *Journal of the Research Centre for Educational Technology, 5,* 42-48.

Preston, M.M. & Holloway, R. (2006). Case study of a basic course: using assessment to legitimize innovation. *Basic Communication Course Annual, 18,* 1-20.

Roth, M., & Sanford Bernhardt, K.L. (2016). Using concept maps for assessment and improvement of a multi-section introduction to engineering course. *American Society for Engineering Education.* Retrieved from https://www.asee.org/public/conferences/64/papers/17112/view.

Sundberg, M.D. (2002). Assessing student learning. *Cell Biology Education, 1*, 11-15.

Williams, R.L. (2017). Retrospective integration of research conducted on a multi-section educational psychology course over a fifteen-year period. *Innovative Higher Education, 42*, 207-224.

Yalcin, M.A., Gardner, E., Anderson, L.B., Kirby-Straker, R., Wolvin, A.D. & Bederson, B.B. (2015). Analysis of consistency in large multi-section courses using exploration of linked visual data summaries. *PeerJ PrePrints, 3*, 2-15.