

Exploring Secondary Teachers' Perceptions of Fairness in Assessment

Student Researcher: Samantha Cullen

Supervisor: Dr. Christina van Barneveld

Committee Member: Dr. Alex Lawson

Internal Reviewer: Dr. Tanya Kaefer

External Reviewer: Dr. Ruth Childs

Faculty of Education, Lakehead University

Abstract

This mixed-methods study explored secondary school teachers' perceptions of fairness in assessment practices in Ontario classrooms. This research explored how educators' practices aligned with fairness principles in the *Standards for Educational and Psychological Testing* and Ontario Ministry of Education policy documents, particularly *Growing Success*.

An explanatory sequential mixed-methods design was used. In Phase 1, 33 secondary school teachers from northern Ontario completed an anonymous online questionnaire involving hypothetical assessment scenarios. In Phase 2, semi-structured Zoom interviews were conducted with 8 educators to further examine and contextualize the survey findings.

Results revealed inconsistencies between teacher assessment practices and established fairness guidelines. Teachers demonstrated substantial variation in assigning final grades to identical student performance profiles, with differences of up to 35 percentage points. Findings also indicated that assessment fairness may be influenced by interpersonal dynamics, as compliant students or those perceived as similar to teachers were more likely to receive favourable consideration.

The study concludes that fair assessment practices cannot rely solely on positive pedagogical intentions. Teacher decision-making is shaped by personal beliefs and social contexts in ways that may conflict with fairness principles. The findings highlight the need for clearer, evidence-based guidance from school boards regarding grading practices, and targeted professional development to mitigate construct irrelevant variance as well as implementing culturally responsive assessment approaches.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Christina van Barneveld, for her guidance, encouragement, and invaluable feedback throughout this project. Her patience, experience and wisdom helped me to have one of the best educational experiences, ever! I also thank the members of my thesis committee for sharing their outstanding expertise, thoughtful suggestions and support. I am grateful to the Department of Education at Lakehead University for providing resources and support that made this research possible. Special thanks are extended to those who approved this project and to my valued colleagues for sharing their time and valuable insights. Through their participation in various parts of this project, they demonstrated a passion for the profession and level of care for students that was inspirational.

Finally, I would like to thank my family and friends for their patience, encouragement, and unwavering support during the completion of this thesis.

Table of Contents

<i>Exploring Secondary Teachers' Perceptions of Fairness in Assessment</i>	1
<i>Abstract</i>	2
<i>Acknowledgements</i>	3
<i>Table of Contents</i>	4
<i>List of Figures</i>	6
<i>List of Tables</i>	7
Chapter 1: Introduction	8
Purpose Statement	8
Research Question	8
Chapter 2: Literature Review	9
Overview of Chapter Structure	9
Section 1: Defining Fairness	9
Growing Success	10
Standards for Educational and Psychological Testing	11
Section 2: Opportunity to Learn	12
Student Engagement: Student Absences	13
Student Engagement: Marginalized Groups	13
Student Anxiety	14
Culture	15
Socio-Economic Factors	15
Disability	16
Formative Assessment	16
Teacher Practice	17
Section 3: Construct Irrelevant Variance	20
Grading Practices: Mathematical Processes	20
Variance from Teacher to Teacher	24
Section 4: Teacher Worldview	26
Centralizing the teachers' worldview	28
Teacher Professional Judgment	30
Social Context	31
Synthesis of the Four Sections of the Literature into a Conceptual Visual Model	33
Contribution of this Research to the Literature	35
Chapter 3: Method	36
Participants	38
Instruments for Data Collection	38
Online Questionnaire	38
Semi-Structured Interview	39

Data Analysis	40
Quantitative Data Analysis	40
Qualitative Data Analysis	40
Integration of Quantitative and Qualitative Data	40
Design Rationale and Limitations	41
Ethical Considerations	41
Chapter 4: Results	43
Phase One: Quantitative Results	43
Phase Two: Qualitative Results	62
Opportunity to Learn	63
Student Engagement: Student Absences.....	63
Student Engagement: Marginalized Groups.....	64
Student Anxiety.....	65
Teacher Practice.....	65
Construct Irrelevant Variance	68
Mathematical Processes.....	68
Variance from Teacher to Teacher.....	70
Cheating or Student Use of AI.....	72
Teacher Worldview	73
Teacher Professional Judgment.....	74
Social Context.....	77
Teacher-Student Relationships.....	78
Chapter 5: Discussion	80
Summary of Key Findings	80
Inequities in Opportunities to Learn	80
Mathematical Processes as Sources of Construct Irrelevant Variance	83
Teacher Worldview Introduces Bias	85
The Updated Model	88
Why Policy Alone Is Insufficient	88
The Centrality of Teacher Worldview	88
Systemic Pressures as Catalysts for Divergence	89
Explaining Divergence from Fairness Principles	89
Unexpected Findings and Alternative Explanations	91
Limitations of the Study	92
Implications for Educational Practice and Policy	93
1. Targeted Professional Development in Assessment	93
2. Embedding Equity and Culturally Responsive Practices in Assessment	94
3. Review and Reform Assessment Policies	94
Recommendations for Future Research	95
References	96
Appendix A	105
Appendix B	122
Appendix C	124

List of Figures

Figure 1.	
A Working Model of Relationships Between Factors Affecting Fairness in Assessment	34
Figure 2. Overview of the Dynamic Relationships Between Factors Affecting Fairness in Assessment	90

List of Tables

<i>Table 1. Methodological Overview of the Investigation into Ontario Secondary Teachers' Understanding of Fairness in Assessment</i>	37
<i>Table 2. Participants' Responses (n, %) to Scenarios Related to Student Absences</i>	44
<i>Table 3. Participants' Responses (n, %) to Scenarios Related to Tests and Students' Opportunity to Learn</i>	46
<i>Table 4. Participants' Responses (n, %) to Scenarios Related to Feedback on a Formative Assessment</i>	48
<i>Table 5. Participants' Responses (n, %) to Scenarios About Determining the Optimal Grade</i>	50
<i>Table 6a. Participants' Responses (n, %) to a Scenario Where Different Students Earn the Same Grade</i>	52
<i>Table 6b. Participants' Responses (n, %) to Scenarios Related to Professional Judgment and Relationships</i>	54
<i>Table 7. Participants' Responses (n, %) to Scenarios Related to Alternative Methods as Accommodations</i>	56
<i>Table 8a. Participants' Responses (n, %) Related to Assessment Practices Demonstrating Maximum Variance</i>	58
<i>Table 8b. Participants' Responses (n, %) Related to Assessment Practices Demonstrating Moderate Variance</i>	59
<i>Table 8c. Participants' Responses (n, %) Related to Assessment Practices Demonstrating Minimum Variance</i>	61

Chapter 1: Introduction

There is a lack of recent work that describes secondary school teachers' perceptions of fairness in assessment in Ontario. Some studies have addressed the issue of fairness, in general, but most were not conducted in secondary schools, were done decades ago, and often didn't include concepts like teacher worldview or student learning contexts. To address this gap, I explored secondary teachers' understanding of concepts that relate to fairness, with a particular interest in exploring ways that teachers perceive fairness in assessment for selected learning contexts. This study documents ways in which teachers' perceptions about fairness align and misalign with fairness principles in policy documents and existing research literature. Misalignments may point to barriers to fairness in student assessment and highlight areas for improvement in teacher practice. Dissemination of the findings could open the door for professional dialogue and possibly, changes in practice that enhance fairness in assessment of student learning.

Purpose Statement

The purpose of this study was to explore secondary teacher perceptions about fairness in assessment practices and explore the extent to which teacher perceptions aligned with current research literature and policy documents. Towards that end, my research question was the following.

Research Question

To what extent do secondary teachers' responses to hypothetical assessment scenarios align with the principles of fairness in assessment, as defined in the *Standards for Educational and Psychological Testing* and Ontario Ministry of Education policies and relevant research?

Chapter 2: Literature Review

Overview of Chapter Structure

This chapter is organized into four sections. Each section describes a key concept related to fair assessment. These four concepts were selected because they are key themes in the research literature on fair assessment. The four sections are entitled:

- Defining fairness,
- Opportunity to learn as a key component of fairness,
- Construct irrelevant variance and unfair assessment, and
- Teacher worldview.

The research literature review for each of the four key concepts is found below.

Following the literature review, I created a model that describes the relationships between these four concepts based on the literature. This model informed the research method and interpretation of results.

Section 1: Defining Fairness

In this section, I examine the key factors that define fairness in assessment and highlight its significance as outlined in two major policy documents. The first policy document is the government of Ontario's K-12 assessment policy document entitled *Growing Success: Assessment, Evaluation, and Reporting in Ontario Schools* (Ontario Ministry of Education, 2010), hereafter referred to as *Growing Success*. The second is the *Standards for Educational and Psychological Testing* (American Psychological Association [APA], American Educational Research Association [AERA], & National Council of Measurement in Education [NCME], 2014). These policy documents were selected because the former offers policy guidelines for

secondary schools in Ontario, which is the context for this research. The latter describes rigorous standards for fairness in assessment that are based on scholarly research and established expertise in the field of educational measurement and evaluation.

Growing Success

Interestingly, *Growing Success* does not define fairness in assessment explicitly, yet it refers to fairness in several sections of the document. For example, it cites Brock University professor of education and researcher Louis Volante's view that "fairness in assessment and evaluation is grounded in the belief that all students should be able to demonstrate their learning regardless of their socio-economic status, ethnicity, gender, geographic location, learning style, and/or need for special services" (p. 34). This statement is consistent with the way the government defines equity in another document as "fair, inclusive, and respectful treatment of all people" (Ontario Ministry of Education, 2014, p. 86). To achieve fairness in assessment, *Growing Success* suggests that teachers should use multiple sources of evidence of student achievement, emphasize recent achievement, provide accommodations to students when needed, and apply professional judgment in determining grades. *Growing Success* also suggests that learning skills and work habits which are not tied to curriculum expectations should not influence assessment (Ontario Ministry of Education, 2010). Fair assessment, therefore, is defined as student ability to demonstrate their learning related to curriculum expectations, regardless of other variables.

Standards for Educational and Psychological Testing¹

¹ Note that a new edition of *The Standards* is expected in 2026. Graduate students and professionals from the National Council on Measurement in Education recommend redefining

In the *Standards for Educational and Psychological Testing*, hereafter referred to as *The Standards*, fairness is defined as the extent to which the assessment process is free from bias and provides equitable opportunities for all students regardless of their individual characteristics, to demonstrate their understanding of the construct that is being measured. Bias is understood as systematic errors or construct irrelevant aspects that differentially affect the performance or outcomes for different groups of students (APA, AERA, & NCME, 2014). When this bias exists, the validity of assessment is called into question. *Valid assessment* is defined as the degree to which accumulated evidence supports a specific interpretation of a result (APA, AERA & NCME, 2014). While the terms valid and fair are sometimes used interchangeably, they are not the same. Validity refers to whether an assessment accurately measures what it is intended to measure. Fairness includes this accuracy but goes further, additionally encompassing ethical considerations such as equity.

The distinguishing feature of fair assessment according to *The Standards* is that it is free from bias as it provides the opportunity to learn for all students and removes construct irrelevant barriers, so outcomes are meaningful and comparable across groups (APA, AERA, & NCME, 2014). Since opportunity to learn and construct irrelevant variance are key concepts in fair assessment, each is described in turn in the next two sections of this chapter.

Section 2: Opportunity to Learn

fairness to emphasize inclusivity and transparency. They argue the revised Standards should explicitly address systemic racism in educational and psychological measurement and support decolonization efforts (Araneda et al., 2023).

Students' opportunity to learn is essential for fair assessment. *The Standards* define opportunity to learn as "the extent to which individuals have had exposure to instruction or knowledge that affords them the opportunity to learn the content and skills" (APA, AERA, & NCME, 2014, p. 56). Pamela Moss, from the University of Michigan, and colleagues researched how validity and equity in assessment relate to opportunity to learn. Moss et al. (2008) explained that because assessment shapes instructional content, resources, or learning processes it can either expand students' educational opportunities or constrain them, reinforcing existing social inequalities. This is consistent with *The Standards* which states that fairness is compromised when students lack adequate opportunities to learn or demonstrate their learning (APA, AERA, & NCME, 2014).

There are several reasons why students may not have an opportunity to learn. For example, students may not have the opportunity to learn if they have reduced engagement due to factors such as absence from school or marginalization. Socio-economic factors can influence opportunities to learn because students may be hungry, unstably housed, or need to work to help support the family, which may limit their participation in school. There may be cultural reasons that impact students' opportunities to learn, such as Indigenous students returning to their communities for special events or hunting season or deeply rooted systemic patterns of marginalization. Opportunities to learn might be reduced if the student has a disability or special learning needs. Finally, student opportunities to learn can be affected by teacher practice. The research literature for each of these topics is reviewed below.

Student Engagement: Student Absences

Attendance is often discussed as a measure of student achievement, but it may be a better measure of engagement since absence from school does not always equate to missed learning.

Research shows that unexcused absences are linked to lower academic performance, while excused absences, such as family trips, may not hinder learning if students receive support (Gottfried, 2009; Hancock et al., 2018). The critical issue is whether all students have the opportunity to learn before being assessed. *The Standards* note that holding students accountable for content they had unequal access to raises fairness concerns (APA, AERA, & NCME, 2014). Teachers' interpretations of and responses to student absences directly impact fairness. Reasons for student absences may influence whether students receive the support or time needed to catch up, affecting their opportunity to learn and thus how fairly they are assessed.

Student Engagement: Marginalized Groups

Language, ethnicity, culture, socio-economic status and differences in ability can affect student engagement. Sometimes students who are not part of the dominant group become marginalized and engagement in school is negatively impacted. John Ogbu, a Nigerian-American anthropologist and professor at the University of California at Berkeley, posited that historical experiences can shape minority students' relationships with school. He distinguished *voluntary* minorities, such as immigrants, from *involuntary* minorities, including those affected by colonization or enslavement. During his ethnographic research, he noted that involuntary minorities may disengage if meeting school expectations feels like abandoning their cultural identity or aligning themselves with an oppressive system (Ogbu, 1998, 2003). Recognizing sociocultural dynamics such as these helps educators understand how opportunities to learn can be limited for marginalized students. Scholars emphasize that schools must honour students' cultures to foster engagement by creating learning environments where students can succeed without compromising who they are.

Student Anxiety

Anxiety can affect students' academic performance. It can limit their opportunity to learn and introduce a construct irrelevant factor that compromises the validity and fairness of assessment. In a longitudinal study involving more than 700 students transitioning from middle to high school, Goldstein et al. (2015) found that adolescents who reported higher levels of stress also reported greater academic difficulties. Stress reduces cognitive capacity and interferes with memory and executive functioning (Feldman, 2019; Hammond, 2015). Extensive research demonstrates that anxiety disrupts working memory processes, including attention control, language comprehension, reasoning, planning, problem solving, and task switching (Chuderski, 2015; Eysenck et al., 2007; Moran 2016). These cognitive processes are essential for meaningful engagement in classroom learning and assessment. Research further suggests that the effects of anxiety are not uniform across students; rather, they vary according to demographic characteristics such as age, gender, socioeconomic status, or other individual factors. For example, Hammond (2015) observed that students of colour often experienced school as an unwelcoming environment where subtle imbalances of power and privilege heightened anxiety and hindered learning. Fisher et al. (1996) found that students with learning disabilities experienced higher anxiety than peers without disabilities, while Tramonte and Willms (2010) reported that students with lower self-confidence and from lower socio-economic backgrounds tended to experience more stress. Teachers also played a key role in shaping students' opportunity to learn. In a four-year study of children in low-income neighbourhoods in Montreal, Kurdi and Archambault (2020) found that supportive, conflict-free teacher-student relationships helped students manage anxiety and persist with learning tasks. Taken together, these studies indicated that anxiety was a meaningful source of construct irrelevant variance that could reduce

opportunities to learn, raising important concerns about fairness, validity, and consistency in assessment.

Culture

Ignoring cultural differences can perpetuate unfair assessment. For instance, Moss et al. (2008) found that while Massachusetts Comprehensive Assessment System scores rose overall, African American and Latinx students consistently lagged behind their White and Asian peers. Aggregated data originally masked these gaps that highlighted inequities in opportunities to learn or demonstrate learning. Examining why achievement gaps persist, particularly when achievement levels seem to be stratified by a factor such as ethnicity, can help challenge practices that limit fairness. While systemic change is essential, individual teachers also play a key role in addressing barriers. Ogbu (1998) emphasized that teachers can build trust and improve engagement by honouring students' cultures and experiences. For instance, acknowledging culture on personal and systemic levels when making decisions in classrooms can improve opportunities for all students to engage in learning, thus promoting fairness in assessment of that learning.

Socio-Economic Factors

Socio-economic hardship often creates a barrier to engagement and opportunities to learn. A study of over 35,000 Florida elementary students found that longer periods in low-income households correlated with lower grades (Morrissey et al., 2014). A child's socio-economic status affected their opportunities to learn in school, influencing everything from access to books and technology to nutritious food, safe housing, and time for study. Two students cannot have the same opportunities to learn if one grows up with economic stability and access to learning resources while the other does not. Schools must acknowledge these disparities and create

equitable conditions so that all students have opportunities to engage and be fairly assessed on what they can do, not what resources they lack.

Disability

The presence of a learning disability can impact a student's opportunity to learn. For example, boys with learning disabilities have reported higher anxiety than peers with similar academic skills, which can lead to increased absenteeism, disengagement, and reduced opportunities to learn (Fisher et al., 1996; Hancock et al., 2018). Appropriate supports or accommodations can level the playing field for students with disabilities so they can better engage and show what they know, thus promoting fairness in assessing what they know or are able to do (Moss et al., 2008) as suggested in *Learning for All* (Ontario Ministry of Education, 2013) and *The Standards* (APA, AERA, & NCME, 2014). Teachers might address factors affecting equitable opportunities to learn through the use of formative assessment.

Formative Assessment

Formative assessment is vital for identifying student needs and guiding accommodations. *Growing Success* notes that it helps teachers adjust instruction and provide feedback while students are still learning (Ontario Ministry of Education, 2010). Lorrie Shepard, professor and researcher at the University of Colorado, urges teachers to use formative assessment to adapt instruction and support equitable participation (Shepard et al., 2018). The primary purpose of formative assessment is not to generate a grade but to address barriers early, such as issues with engagement, and increasing opportunities for all students to learn. I explore formative assessment further in the next section and describe how a teacher's understanding and use of formative assessment practices can significantly influence students' opportunities to learn and the overall fairness of assessment.

Teacher Practice

Teachers are among the most important people in providing students opportunities to learn. Work done by Moss et al. (2008) suggests that teachers' knowledge of content and pedagogy and their commitment to educating all students equitably, directly contribute to students' opportunities to learn. This is consistent with the work of Joe Feldman, educator and author of *Grading for Equity*, who argues that teachers should do more than collect data: they must deeply understand what data reveal about each student's learning journey and use it to foster hope and growth for all learners (Feldman, 2019). To illustrate teachers' powerful role in student achievement, Feldman highlights John Hattie's influential research. In his analysis of over 500,000 studies, Hattie (2003) found that while students themselves account for about half of the variation in achievement, teachers are the next most significant factor, particularly through the feedback they provide. However, it is important to note that not all feedback improves opportunities for students to learn. Feldman and other equity advocates caution that assessment practices and feedback can unintentionally create barriers rather than address them. To ensure fairness, teachers must continuously reflect on whether their assessment and communication methods genuinely support every student's opportunity to learn. A good example of this is their use of formative assessment.

Although teachers are typically aware that formative assessment improves student learning and engagement, they may not be using it well. Without skillful, evidence-informed practice, feedback can just as easily limit learning as support it (Black et al., 2004; Hattie, 2007; Moss et al., 2008; Shepard et al., 2018). Voerman et al. (2012) define feedback as "information provided by the teacher concerning the performance or understanding of the student, with reference to a goal and aimed at improving learning" (p. 1108). In their empirical research with

classroom teachers on feedback intervention, they found that clear feedback about gaps between current and desired performance helps students improve.

Not all feedback is equally useful to students, however. Butler and Nisan (1986) found that sixth graders who received only grades were less satisfied and less able to improve than those who received written comments, suggesting that grades alone rarely help students understand how to improve. Decades of research demonstrates the potential impact formative assessment can make on student achievement; however, teachers still rely heavily on summative grading methods or evaluation (Brookhart et al., 2016). Although educators may see grades as feedback, Feldman (2019) and Shepard (2019) warn that grading work intended to be formative often ends the feedback cycle prematurely, as students tend to ignore comments once a grade appears. In their work with elementary educators, Black et al. (2004) noted that much of the time teachers spent grading did little to advance student learning.

Changing the assessment practices of teachers is difficult. Deneen and Boud (2014) found that even when faculty at the University of Hong Kong welcomed better learning activities, they resisted changing how they assessed students. History shows this resistance to change is longstanding. In the 1930s, the Eight Year Study in the United States brought together educators and social scientists to experiment with progressive approaches to curriculum and instruction in more than 30 secondary schools. Educators quickly learned that progressive teaching reforms would not succeed without corresponding shifts in assessment. Only when more than 300 colleges agreed to accept new forms of assessment for incoming students could schools fully adopt innovative teaching and assessment approaches, improving students' opportunities to learn (Tyler, 1986). Yet decades later, secondary schools still cling to traditional grading and testing methods. Pressure to focus on student credit accumulation, graduation rates,

and post-secondary admission help to keep outdated practices in place (Moss et al., 2008). Black et al. (2010) also highlight this tension. They found that even when teachers recognized flaws in their methods leading to unfair or inaccurate judgments of student learning, they often clung to old habits for comfort and perceived control.

In an extensive review of research on grading practices over a hundred years, Brookhart et al. (2016) noted that policy shifts alone rarely change practice. Taken together, these studies suggest that even when teachers value formative assessment and are aware of its potential to improve opportunities for students to learn, institutional pressures, deeply rooted beliefs, entrenched habits and fear of losing control can make it difficult for educators to abandon familiar methods.

Overall, opportunity to learn is a foundational condition for fair assessment. The literature shows that students' ability to demonstrate their learning is shaped not only by instruction, but also by factors such as engagement, cultural and socio-economic context, anxiety, disability, and teacher practice. When these conditions limit access to learning, assessment results risk reflecting unequal opportunities rather than true achievement. While formative assessment and responsive teaching can help address these disparities, entrenched practices and systemic pressures often constrain their effective use. Assessment outcomes may be influenced by many factors raising additional concerns about construct irrelevant variance and how well assessment methods measure what they are intended to measure.

Section 3: Construct Irrelevant Variance

In this third section, I explore how teachers' grading practices and student-related factors can introduce bias or errors when assessing student learning. *The Standards* describe construct irrelevant variance as factors that distort grades by reflecting influences unrelated to the skills or knowledge being measured (APA, AERA, & NCME, 2014). Therefore, grades that are shaped by influences outside the intended learning goals do not represent valid measures of achievement, calling the fairness of such assessments into question.

Researchers have suggested that grades can be unfairly impacted by poorly designed lessons, ineffective use of formative assessment, non-cognitive factors like behaviour, or the ways teachers record, calculate, and interpret scores (Brookhart et al., 2016). Examining measurement theory and the nuances of professional judgment that teachers exercise when interpreting grades helps us to gain a deeper understanding of how construct irrelevant variance impacts fairness in assessment. I therefore explore the grading practices of teachers and ways in which they introduce potential sources of construct irrelevant variance.

Grading Practices: Mathematical Processes

The methods teachers use to record, calculate, and interpret scores can produce grades that may not accurately reflect student knowledge or ability. Because grading practices are rarely monitored, teachers have substantial autonomy in assigning marks (Brookhart et al., 2016; Feldman, 2019). Black et al. (2010) found that in many schools, "there were no formal procedures in place to review overall patterns in test scores or the records in teachers' mark books" (p. 220), meaning that grades can reflect individual teacher practices and not solely student achievement. Grades that reflect something other than the knowledge or ability being measured are not fair assessment.

In Ontario high schools, grading practices involve averaging the grade from multiple learning activities to obtain a final grade. Despite ministry guidance that non-academic factors like effort or behaviour should not influence grades, teachers may include them either unintentionally, or because they strongly believe that certain behaviours should count. Some researchers describe this as “hodgepodge grading” (Brookhart et al., 2016) or “grade pollution” (Green et al., 2007). These factors may introduce bias into the assessment process and inflate or deflate grades, ultimately contributing to construct irrelevant variance and undermining fairness.

When low grades received early on are included in averages, final marks are disproportionately lower and misrepresent student growth. For example, imagine a student who scores 10%, then 60%, and finally 80% on assessments of the same curriculum expectation. Averaging these scores results in a final grade of 50%, even though the student’s most recent performance is much stronger. Because there is often “very little specific guidance for teachers on how to calculate grades based on a body of evidence” (DeLuca et al., 2017, p. 184), teachers must ask themselves whether a grade of 50% truly represents this student’s current level of achievement.

Reliance on grading software can exacerbate the problem in assigning a fair final grade. Guskey (2002), a professor and researcher in education at the University of Kentucky, explains that many teachers rely on computer programs to calculate averages because it makes it easy for them to explain and defend their grading practices, however, he cautions that “the mathematical precision they offer does not make the grading process any more objective or any fairer” (p. 777). For these reasons, many researchers and experienced educators argue that relying on averages as the main measure of student achievement is flawed and should be abandoned (Guskey, 2002; Guskey & Jung, 2016; O’Connor, 2011; Reeves, 2000). Depending on the system

teachers use to record individual scores, they may not realize that the mathematical processes they use to calculate grades include sources of construct irrelevant variance which can be misleading and ultimately unfair.

Assigning zeros for missing work is particularly problematic, as this practice contributes to construct irrelevance and inaccurate representations of knowledge or achievement. Assigning a zero can drastically lower averages, making other work count less than it should, inaccurately reflecting student understanding (Feldman, 2019; Reeves, 2000, 2004). Consider a student who receives a mark of 80% on a writing assignment and then receives a 0% on the next assignment because they did not turn it in. This brings their average grade to 40%. Not handing work in, causes work that was completed to be devalued so that the grade is no longer an accurate reflection of what the student knows. *Growing Success* describes both sides of the zero debate—that assigning a zero for missed work can be considered as an accountability measure, and that this practice is punitive. The document states that teachers will “use their professional judgment to determine the student’s report card grade” (Ontario Ministry of Education, 2010, p. 39), offering no explicit guidance on whether zeros should be assigned or not. Reeves (2004), an American researcher and author, contends that assigning zeros lacks logical and mathematical validity, as it can permanently distort a student’s overall grade. Although some teachers view zeros as a natural consequence for incomplete work, Reeves (2000) argues that “avoiding zeros is not an act of leniency but a matter of ensuring mathematical accuracy in grading” (p. 12). Unlike other consequences that end, a zero continues to affect a student’s average forever.

Alternatives, such as minimum grading, have been shown to preserve mathematical accuracy. Minimum grading is an assessment practice that prevents teachers from assigning grades below a pre-determined threshold. While some educators apply the minimum grade to

individual assignments throughout a reporting period, it is more commonly used to adjust a students' overall term or semester grade after a grading period has concluded (Carifio and Carey, 2009). For instance, *Growing Success* states that individual boards determine the lower limit for overall grades below 50% that will be recorded on the report card (Ontario Ministry of Education, 2010). A seven-year study of nearly 11,000 high school students in the U.S. found evidence that contradicts concerns about minimum grading causing grade inflation. Rather than artificially inflating grades, minimum grading practices have been shown to provide a more realistic picture of student achievement (Carifio & Carey, 2012, 2013). Although there is no evidence that assigning zeros or low marks as punishment improves student behaviour, some teachers continue to believe that a zero is an appropriate consequence for unfinished work. Regardless of personal beliefs, research shows that including zeros in grade calculations distorts averages, rendering inaccurate grades due to construct irrelevant variance (Feldman, 2019; Reeves, 2000, 2004). In light of this evidence, practices such as minimum grading may enhance fairness in assessment by reducing construct irrelevant variance and rendering grades that are a more accurate reflection of achievement.

Another way that teachers might deal with missing work is to exclude zeros entirely by leaving a task blank in their gradebook, often called a 'no mark.' This means that zeros are not included in the average and a student's grade is based only on completed work. Some teachers see this as unfair, since a student with three completed tasks might earn the same average as a student who completed five. This practice might be utilized inconsistently from classroom to classroom, bringing in additional concerns about fairness. Ultimately, these decisions about how missing work is handled can introduce mathematical inconsistencies that add construct irrelevant variance to grades, weakening their validity and fairness.

A review of grading policies and practices in Canada by DeLuca et al. (2017) found that “the nature of these policies is generally not prescriptive, leaving room for wide variation in how assessment is carried out across schools” (p. 184). Flexible grading policies respect teacher professional judgment but can introduce significant construct irrelevant variance when teachers include non-cognitive factors, zeros, and rely heavily on averages when determining overall grades. These practices not only distort the true measure of student achievement but also contribute to significant inconsistencies between teachers, highlighting areas where grading practices can vary widely and unfairly. Next, I describe variance among teachers keeping in mind that fair assessment is free from bias and provides equitable opportunities for all students to demonstrate what they know or are able to do.

Variance from Teacher to Teacher

Assessment is one area in education where there is little professional consensus between teachers. In fact, there is substantial evidence that “great variation exists in the grades teachers assign to student work” (Brookhart et al., 2016, p. 806). Variance in this context is due to the differences or inconsistencies in assessment practices between teachers. This raises important questions about the fairness of grades, as students may not be assessed by the same standards or have comparable opportunities from one classroom to another.

In the early 1900s, researchers Starch and Elliot (1912) observed significant variance in grades from teacher to teacher. In one experiment, the same high school English paper received 30 different grades from over a hundred teachers, ranging from 64% to 98%. This raises serious questions about validity, such as: which mark within this range truly reflects what the student knows or can do? While some dismissed the findings by arguing that assessing writing is subjective, a follow up experiment found even greater variation in math, where the same

geometry paper earned grades from 28% to 95% (Feldman, 2019). Some may argue that grading consistency has improved over time, but Brimi (2011) found similar results nearly a century later. In 2007/2008, 73 English teachers trained in the same grading system gave the same essay marks ranging from 50% to 96%. Likewise, Brookhart et al. (2016) found that grading practices still vary widely, even within the same school. Because fair assessment relies on both validity and equity, the wide variation in grading across teachers raises serious concerns about the fairness and consistency of current evaluation practices.

One reason for variation in grading from teacher to teacher is that some teachers use grades to reward or punish student behaviours unrelated to the learning objective. For example, Feldman (2019) identified teacher grading practices such as giving extra credit, late penalties, punishing cheating, rewarding effort, and manual adjustments as common sources of variability, noting that teachers often alter scores based on what they believe a student ‘deserves’. Similarly, McMillan’s (2001) survey of 1,400 Virginia teachers revealed wide variation in how cognitive and non-cognitive factors such as behaviour, neatness, or effort are weighted. These factors reflected personal student circumstances more than academic achievement, introducing construct irrelevant variance and raising fairness concerns. Teachers’ views differ greatly in how much student effort etc., should count toward final grades, thus compounding the issue of variance from teacher to teacher.

Having examined ways in which teachers may unintentionally contribute to construct irrelevant variance in assessment, I turn now to an examination of teacher worldview as a central lens through which beliefs about assessment and practices are formed.

Section 4: Teacher Worldview

Fairness in assessment cannot be separated from the worldviews that teachers bring to their classrooms. Cultural values, social norms, and lived experiences shape how teachers interpret student learning and what they consider evidence of achievement. When these beliefs are shaped by dominant or Eurocentric perspectives, they can perpetuate inequities in how students are taught and learning is assessed. Understanding these underlying perspectives helps uncover ways in which prejudice (preconceived judgments or attitudes) is a source of bias and how power dynamics can influence assessment decisions individually or systemically, thus impacting fairness. In this fourth and final section I examine the critical role of teacher worldview in shaping assessment through its influence on power dynamics, the use of professional judgment, and the social context of education, specifically teacher-student relationships.

Perceptions of fairness in assessment are shaped by values, beliefs, experiences and the social context in which educators find themselves. As Baniyadi et al. (2023) argue, the social, cultural, and relational dimensions of classrooms must be taken into account when designing and implementing assessments. In an article by researcher and professor in the field of education at Stanford University, Solano-Flores (2023) argues that culture and society deeply influence how people think, and that fairness concerns often emerge when assessments fail to consider students' diverse cultural, linguistic, gender, ability, or socio-economic backgrounds. Teachers' personal values and lived experiences influence their internalization of criteria that is important for assessment and grading, meaning their beliefs can either reinforce barriers to fairness or help remove them.

Researchers and scholars posit that Eurocentrism can shape fairness in classroom assessment by influencing what is taught, how it is taught, and how learning is measured (Battiste, 2005; Hunn, 2004; Sundberg, 2009). Eurocentrism refers to the tendency to interpret the world through European or Anglo-American values and perspectives (Merriam-Webster, n.d. Definition 1). Scholars describe Eurocentrism as an attitude that positions Europe as the source of universal values and progress, justifying prejudice and domination over non-European cultures. Historically, this belief in superiority has supported the imposition of European norms, knowledge, and language onto others, often through systems like education (Battiste, 2005; Hunn, 2004; Sundberg, 2009). This dominant worldview privileges certain ways of knowing while dismissing others, thus restricting some students' opportunities to learn and introducing construct irrelevant variance that can distort valid assessment, contributing to a process that is unfair.

Advocates for change argue that schools continue to act as assimilating agents, embedding Eurocentric, white, middle-class norms in curricula, policies, school culture, and assessment. Jamila Lyiscott (2019), a social justice scholar, illustrates this in *Black Appetite. White Food*. She cites a 2015 *Washington Post* piece noting that Eurocentric curricula still dominate “in a society that is increasingly brown” (p. 3). She recounts the experience of a Native American student who was sent to the principal's office for questioning why a history textbook devoted only a single paragraph to Native American history. The story illustrates how schools often expect students to accept narrow narratives without challenging omissions or prejudice. Science of learning research shows that prior knowledge is critical, yet dominant systems often dismiss knowledge rooted in non-Eurocentric cultures. Hammond (2015) explains how these attitudes marginalize the skills and experiences many students bring to the classroom. When

cultural prejudice, power imbalances, and historical inequities go unexamined, they reinforce barriers to fairness.

Centralizing the teachers' worldview

Brazilian educator and philosopher Freire (1970, 2000) argued that traditional education silences marginalized voices by treating students as passive recipients of knowledge. Instead, he called for a dialogical approach that values students' lived experiences and empowers them to challenge oppressive systems, including unfair classroom practices. This is consistent with Moss et al. (2008) who noted that teachers often struggle to recognize worldviews different from their own, perpetuating deficit thinking and stereotypes. This is reflected in Gershenson et al.'s (2016) large-scale study of over 1,600 teacher-student dyads, which found that non-Black teachers hold significantly lower expectations for Black students than Black teachers do, shaping course choices and long-term outcomes.

Codrington and Fairchild (2012) argue that despite *Brown v. Board of Education*, subtle forms of segregation persist. For instance, Black students are still 1.5 to 3 times more likely than others to be referred to special education. Grades can also sort students into tracks that limit opportunities. In Ontario, high school students can be placed in courses at different levels according to their academic ability or predicted post-secondary pathway. In grade 10, these are defined as academic, applied or essential (locally developed). In grades 11 and 12, the tracks or streams are referred to as university, college, or workplace level. Terrin and Triventi's (2023) meta-analysis showed that tracking or streaming (i.e., placing students into separate groups based on perceived academic achievement), while often justified as efficient, actually heightens inequality. Local data echoes this. York University's Black Demographic Advisory Committee reported that Black students in Toronto are twice as likely as peers to be streamed into applied or

essential courses that restrict access to college or university (James & Turner, 2017). These pathways, often referred to as ‘lower track’ have lower graduation rates and poorer performance on provincial assessments, raising clear fairness concerns because the hierarchical nature of the system limits the power of historically marginalized students to shape their own outcomes.

Teachers’ worldview of assessment of student learning is often based on their own lived experiences. Teachers continue to use the same assessment methods they once experienced without questioning how these practices might reinforce inequities (Feldman, 2019). Brookhart et al. (2016) observe that most teachers “today are not well trained, typically grade alone, and rarely seek help from colleagues to check the reliability of their grading” (p. 836). When the lived experience of the teacher is like the lived experience of the student, then they are more likely to be fairly assessed. For example, based on their research with Maori students and teachers in New Zealand, Kerr and Averill (2021) found that some students may be placed in a privileged position if their culture aligns with the teacher’s beliefs and skills regarding assessment. When teachers hold different worldviews than their students, it creates the potential for assessment to become a tool to penalize students whose behaviour or views are inconsistent with their own cultural view (Feldman, 2019; Kerr & Averill, 2021; Schuwirth, 2013). Sharing power, partly by respecting students’ voices and experiences, can challenge the hierarchical roles that perpetuate problematic perceptions and judgments (Holman et al., 2021; Kerr & Averill, 2021). Continuing to use practices that many scholars refer to as ‘traditional’ may be reinforcing the types of power dynamics that are rooted in Eurocentrism, likely creating barriers to fair assessment practices.

Teacher Professional Judgment

A teacher's professional judgment is essential for fair assessment (Brookhart et al., 2016; Ontario Ministry of Education, 2010). *Growing Success* defines this judgment as “informed by professional knowledge of curriculum expectations, context, evidence of learning, methods of instruction and assessment, and the criteria and standards that indicate success in student learning” (Ontario Ministry of Education, 2010, p. 152). Different worldviews can influence how evidence of learning is interpreted and weighted. Even with the same evidence, teachers may arrive at different conclusions based on their beliefs about the purpose of assessment and their understanding of fairness. As Moss et al. (2008) noted, assessment interacts with classroom culture and opportunities to learn, shaping student trajectories. Therefore, evidence is best interpreted considering the classroom culture including opportunities for learning and the ways in which students are able to use those opportunities. Kerr and Averill (2021) remind us that educators use assessment to “adapt learning so that all learners’ potential is realized” (p. 241). This cannot be done solely by crunching numbers.

Although professional judgment is essential in assessment, it opens the door for potential prejudice to creep in. Hammond (2015) explains that unconscious attitudes and stereotypes, often referred to as ‘implicit bias’, can shape teachers’ judgments. In a paper that highlights techniques for redistributing power with school based psychoeducational assessments, Holman et al. (2021) emphasize that educators must stay alert to how prejudice or preconceived notions about students’ abilities can skew decisions. Research shows that teachers often favour students who fit stereotypical expectations or who display effort in ways teachers value (Brookhart et al., 2016; Kelly, 2008). Even the help students receive can be unfair. For instance, students who know how to ask for help may get more support than those who do not. Favouritism or discrimination are

types of bias that can have a differential impact on students' opportunities to engage in classroom activities and demonstrate learning, thus rendering assessment practices unfair.

While professional judgment is necessary, it can undermine fairness when dominant practices and assumptions go unquestioned, echoing Freire's (2000) warning that unexamined habits sustain the status quo. To truly advance fairness, teachers must reflect critically on how their judgments and underlying beliefs or attitudes shape their assessment practices and consequently students' opportunities. Having explored the necessity as well as the pitfalls of professional judgment, which is largely impacted by a teacher's worldview, I turn now to the social context of education to further our understanding of teachers' perceptions of fairness in assessment.

Social Context

All learning and assessment happen within a social context, which is a key factor in assessment. Too often, classrooms are assumed to be culturally neutral spaces, but Hammond (2015) explains that they reflect the teacher's worldview. Lyiscott (2019) describes how curriculum and pedagogy lack diversity and students in the 2SLGBTQ+ community have echoed this emphasizing that a lack of representation leads to feeling devalued (Peter et al., 2021). Brookhart et al. (2003) remind us that classroom assessment is deeply tied to the social and relational environment, especially in the context of the teacher-student relationship.

The quality of teacher-student relationships has an impact on the fairness of assessment and the learning experience of students. Teachers hold significant power to shape not only what is assessed but how students feel about their capacity to succeed. Roorda et al. (2011) analyzed over 99 studies and found that positive teacher-student relationships significantly boost student

engagement and achievement, while negative relationships have the opposite effect. Notably, these relationships have an even stronger impact when grades are used as a measure of achievement. It is important to note that this meta-analysis consistently showed that student engagement tends to decline with age, along with the quality of their relationships with teachers. This matters for students from low socio-economic backgrounds, ethnic minority groups, or with learning difficulties, who are affected by the quality of connection they have with their teachers more than their peers in different circumstances. Supportive relationships enable teachers to close gaps that unfair systems create and actively dismantle barriers to fairness, creating opportunities for all students to fully engage in and demonstrate learning.

Sources of identity and socialization strongly influence what students learn and the opportunities they have to learn. Therefore, it is essential to acknowledge the social and cultural pressures students face, rather than pretending that school exists in a racial or cultural vacuum (Gordon, 1995; Holman, 2021). Having examined favouritism as a form of prejudice often rooted in a teacher's worldview, albeit unintentional, I now turn to how such views and attitudes can erode trust and harm teacher-student relationships, creating inequities that result in unfair assessment.

According to the Ontario College of Teachers (n.d.) trust is a fundamental ethical standard for the teaching profession; one that “embodies fairness”. Research indicates that teachers care about ethical practices that promote fairness, but maintaining trust requires more than good intentions. Favouritism impacts teacher-student relationships and is a fairness concern. Aydogan's (2008) study of nearly 900 Turkish high school students found that teachers were perceived to grade differently based on gender, socio-economic status, attractiveness, parental influence, and ideology. Similar results appeared in Pakistan, where university students reported

that teachers awarded higher marks to favoured students, often due to wealth or appearance (Quraishi et al., 2021). Such favouritism, whether real or perceived, undermines fairness because it hinders relationships. In the context of secondary school, this means that fair assessment depends not only on the knowledge and skills that are being measured, but on the quality of teacher-student relationships that shape how students are engaged, supported and able to demonstrate their learning.

Synthesis of the Four Sections of the Literature into a Conceptual Visual Model

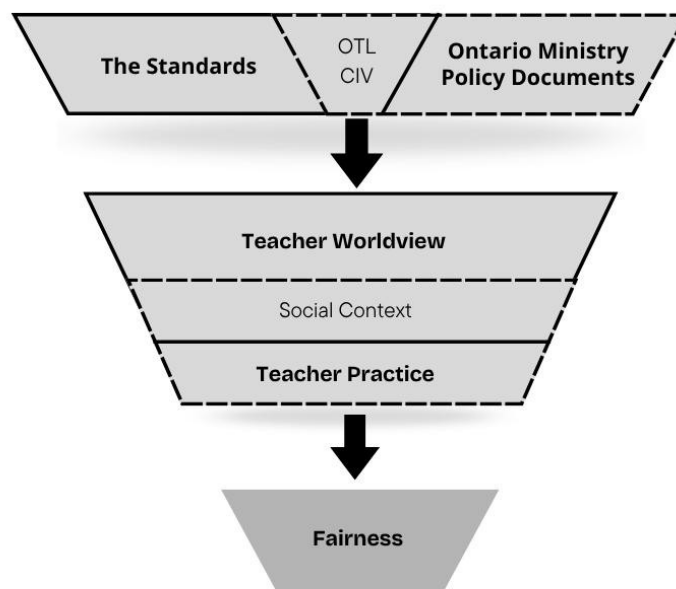
In this final section, the main ideas from the literature are synthesized to show how guiding policy documents and teacher worldview influence teachers' perceptions and practices, thus impacting fairness in assessment. The literature emphasizes that fairness is not only a technical concern impacted by educators' professional knowledge and experience, but an ethical one shaped by teachers' beliefs, relationships, and systemic conditions. Research reveals significant inconsistencies in assessment practices and highlights barriers such as prejudiced views that contribute to inequitable opportunities, and grading methods that may compromise fairness.

The relationships among key dimensions of fair assessment are represented in a conceptual framework illustrated in Figure 1. Since ways in which theories and policies are enacted seem to be filtered through one's worldview, I chose to use a funnel model. *The Standards for Educational and Psychological Testing* along with the Ontario Ministry of Education's policy document, *Growing Success*, which is a foundational source for understanding the nature and purpose of assessment are at the top of the funnel as guides for teacher practice. The documents both address the concepts of Opportunity to Learn (OTL) and Construct Irrelevant Variance (CIV) highlighting fairness in assessment as a key concern. The

documents fall short in terms of the social context of assessment as it relates to fairness, however, research literature on teachers' assessment practices fills that gap in understanding the context of how theory is put into practice. Teacher worldview and practice are shown to overlap with social context in the middle to visually demonstrate the impact of worldview on practice and that views and practices are all influenced by and enacted in a social context. A teacher's approach to fairness in classroom assessment and their interpretation of these guiding documents are shaped by their personal worldview and the broader social context in which teachers and students find themselves.

Figure 1

A Working Model of Relationships Between Factors Affecting Fairness in Assessment



This model leads to some testable expectations in the research. For instance, if opportunities to learn are mediated by a teacher's worldview, we would expect to see differences

in how student absences are treated across different demographic groups of students. For example, teachers whose worldview is that athletics are valuable may show more leniency toward absences of student athletes than absences for other reasons. If construct irrelevant variance is influenced by worldview, we may expect to see differences in how teachers arrive at optimal grades that reflect student achievement. If social contexts impact how teachers put policy into practice and this is also mediated by worldview, we can expect variation among teachers in their perceptions of what is fair and what is actually done in day-to-day practice.

Contribution of this Research to the Literature

By focusing on Ontario high school teachers' perspectives and linking measurement theory with the social dimensions of classroom practice, this study seeks to extend the literature and provide a deeper understanding of how fairness is interpreted and enacted in secondary school assessment. The goal was to illuminate discrepancies between teachers' perceptions and fairness principles in the guiding documents through an examination of how experienced educators grapple with the ethical and social dimensions impacting fairness while expressing a desire to improve practices and promote equity and fairness for all students.

Chapter 3: Method

To answer the research question “To what extent do secondary teachers' responses to hypothetical assessment scenarios align with the principles of fairness in assessment, as defined in the *Standards for Educational and Psychological Testing* and Ontario Ministry of Education policies and relevant research?” an explanatory sequential mixed-methods design was used, where quantitative data were collected and analyzed first, followed by the collection and analysis of qualitative data. This design was chosen because it allowed qualitative research to provide context for interpreting the initial quantitative results.

The study was conducted in two phases. The first phase involved quantitative data collection, which aimed to understand how secondary teachers' responses to hypothetical assessment scenarios align with key principles of fairness, as defined in *The Standards for Educational and Psychological Testing*, Ontario Ministry of Education policies, and relevant research literature. The second phase focused on qualitative data collection through interviews with educators, which were used to further explore and explain quantitative findings. These qualitative data provided a more detailed understanding of teachers' perceptions of fairness in assessment.

The mixed methods approach in this study was implemented in two ways. First, the qualitative data collection protocols were developed based on quantitative results, creating a clear link between the two phases. Second, the two phases were integrated during the interpretation phase, where quantitative and qualitative findings were compared and discussed to identify overarching themes and implications. The research plan is outlined in Table 1.

Table 1*Methodological Overview of the Investigation into Ontario Secondary Teachers' Understanding of Fairness in Assessment*

Heading	Description
Research question	To what extent do secondary teachers' responses to hypothetical assessment scenarios align with the principles of fairness in assessment, as defined in the <i>Standards for Educational and Psychological Testing</i> and Ontario Ministry of Education policies and relevant research?
Source of data	The primary data sources were 33 secondary school teachers in northern Ontario who participated in the first phase, and a subset of 8 teachers who participated in the second phase.
Method of data collection	The study used an explanatory sequential mixed methods design conducted in two phases. Phase 1 involved an anonymous online questionnaire featuring 19 items, including multiple-choice and Likert-scale questions. Phase 2 consisted of semi-structured interviews conducted via Zoom to clarify and elaborate on the questionnaire results.
Data analysis approach	Quantitative data from the questionnaire were analyzed using descriptive statistics, specifically frequencies and percentages. Qualitative data from participant comments and interview transcripts were analyzed using thematic analysis. The two data types were integrated during the interpretation phase to identify patterns of convergence or divergence.
Timeline/milestones	Key milestones included obtaining approval from the Research Ethics Board and the school board. Phase 1 required participants to complete the online survey within a two-week window. Phase 2 followed within a month of survey completion, with interviews scheduled at the participants' convenience. Interviews and data collection took place in June and July 2024.

Participants

The participants were secondary school teachers in Northern Ontario. After receiving approval from the Research Ethics Board at Lakehead University and one school board, an invitation to participate was sent via email to all secondary teachers (approximately 165). A total of 33 teachers consented to participate in Phase 1, the online questionnaire. At the end of the online questionnaire, participants were invited to share their email addresses if they wished to participate in Phase 2 interviews. Fifteen participants expressed interest, but due to scheduling conflicts or theme saturation, eight teachers participated in the interviews.

Instruments for Data Collection

Online Questionnaire

An online questionnaire (in SurveyMonkey) was used to collect quantitative data in this study. The questionnaire was developed to assess teachers' perceptions of fairness in classroom assessment through hypothetical scenarios. The scenarios were developed using the themes and codes identified during the literature review. Teachers were asked to respond to hypothetical scenarios by indicating what they thought most teachers would do, rather than what they themselves would do or have done. This approach aimed to encourage responses that reflected broader professional norms. The goal was for participants to be able to answer honestly without fear of repercussions, by specifically stating something they had done, for instance. Participants also had the option to provide explanations for their answers. The full questionnaire is provided in Appendix A, along with references to relevant documents in the literature review.

Prior to administration, the questionnaire was piloted with a small group of educators to get an indication of whether educators would participate in completing this type of questionnaire,

as well as to ensure clarity, relevance and appropriateness of items. Feedback from the pilot was used to refine the questions and improve the overall validity of the instrument.

The questionnaire was anonymous, and responses were confidential. Teachers were informed that their participation was voluntary and that they could withdraw at any time without penalty. The questionnaire was distributed electronically with a link in an email that went out to all secondary teachers in one school board, and participants were asked to complete the survey within 2 weeks.

Semi-Structured Interview

Semi-structured interviews were conducted with a subset of participants from Phase 1 to gather qualitative data on teachers' perceptions of fairness in assessment. The interview protocol was developed based on the initial survey results, with questions designed to explore and clarify participants' responses to items in the questionnaire. It is included in Appendix C. The interviews were conducted via Zoom, recorded with participants' consent, and transcribed verbatim for analysis.

Confidentiality was emphasized during the interviews, and participants were encouraged to speak candidly about their experiences. Given the sensitive nature of the topics discussed, individual interviews were chosen over focus groups to ensure a comfortable and private setting for each participant.

Data Analysis

Quantitative Data Analysis

Responses to questions on the questionnaire were analyzed using descriptive statistics. Frequencies and percentages were calculated to identify trends in teachers' perceptions of fairness in assessment.

Qualitative Data Analysis

The qualitative data, including participant comments on the questionnaire and interview transcripts, were analyzed using thematic analysis. The initial coding process was guided by themes identified during the literature review, but new themes also emerged inductively from the data. Codes were applied to passages from the interview transcripts, and these were organized into categories to capture key ideas. For example, the initial code 'calculations and measures of central tendency used to determine and interpret student grades' was broadened to 'mathematical processes' after multiple interviews described the methods used in day-to-day practice which differed when compared to the thought processes participants described in considering the hypothetical scenarios presented. Due to the scale of variance as well as the gap between perception and practice observed, these instances were later grouped under two higher-level themes: construct irrelevant variance and worldview, specifically teacher professional judgment. Preliminary codes are provided in Appendix B and I describe the final list of codes in the Results and Discussion chapters of this thesis.

Integration of Quantitative and Qualitative Data

The integration of quantitative and qualitative data occurred during the interpretation phase. The quantitative findings were compared with qualitative themes to identify patterns of convergence or divergence. This process allowed for a deeper understanding of the results, where

patterns from one data type helped to explain or challenge those identified in the other. In the Discussion section, I elaborate on some specific insights and how they are integrated.

Design Rationale and Limitations

The decision to use hypothetical scenarios, where participants were asked what they thought most teachers would do, was inspired by the work of Childs and Umezawa (2009), who employed a similar approach to study dilemmas associated with administering EQAO tests. Some scenarios were based on examples used to explore or demonstrate commonly held beliefs about assessment practices in Feldman's (2019) book *Grading for Equity*.

Asking *what most teachers would do* was meant to reduce social desirability (individuals responding to questions in a way they believe would be viewed favourably by others, rather than answering completely honestly), shifting the focus away from the participants' own behaviour and moral identity. My intent was to provide a context for teachers to share their perceptions and freely discuss topics such as prejudice and other sources of bias as well as how optimal grades are determined without fear of judgment or negative consequences. However, this may have also contributed to one of the limitations of this study. It is possible that this approach blurred the distinction between participants' own practices and their perceptions of norms on a wider scale.

Ethical Considerations

An application for ethical approval was submitted to Lakehead University's Research Ethics Board, and a formal request was sent to the school board for permission to conduct the study. The study posed no known or anticipated risks to participants, and all teacher participation was voluntary.

The online questionnaire was designed to ensure anonymity. The information letter and consent form were included at the beginning the questionnaire (see Appendix A). In the interviews, confidentiality was maintained by using pseudonyms for all participants. Every effort was made to protect participants' identities throughout the study. In fact, some interview excerpts could not be included due to the potential for identification. Teachers were also able to withdraw from the study at any time.

Chapter 4: Results

Phase One: Quantitative Results

The results below include the analysis of data from the individual participant questionnaires. Descriptive statistics are presented as frequencies, indicating patterns and tendencies in responses. A total of 33 secondary teachers completed the questionnaire. Participants did not respond to all questions. The number of participants who responded to each question is indicated in the row at the bottom of each table. The results are separated based on the themes presented in the hypothetical scenarios to compare response rates, make observations, and attempt to draw conclusions from the frequencies of participant responses.

The results from the first three questions on the questionnaire are illustrated in Table 2. In all three of these scenarios, a student missed completing two assessment tasks because they were absent from school. The variable in each question is the reason for their absence: illness, hockey, or an Indigenous student goes hunting. Participants were asked to indicate what they thought most teachers would do if each of these students completed these assessments when they returned to school. The reasons for the absences and the participant responses are summarized in Table 2.

Table 2*Participants' Responses (n, %) to Scenarios Related to Student Absences*

Responses	The student absence is:		
	Due to Illness	Due to Hockey	Due to Hunting
Assess the two assignments the same as students who completed the assignments on time.	31 (94%)	24 (73%)	24 (73%)
Assess the assignments and remove marks for being late.	2 (6%)	8 (24%)	8 (24%)
Assign zeros for the missed assignments. Most teachers do not accept late assignments.	0 (0%)	0 (0%)	1 (3%)
Leave it blank in the gradebook.	0 (0%)	1 (3%)	0 (0%)
Give the student more difficult assessments to complete when they return.	0 (0%)	0 (0%)	0 (0%)
Other	0 (0%)	0 (0%)	0 (0%)
Total Responses	33 (100%)	33 (100%)	33 (100%)

Results in Table 2 demonstrate some variance in the practice of deducting late marks. For example, if a student was absent due to hockey or an Indigenous student went hunting, 73% thought that teachers would not deduct marks for late work, compared to 24% who felt that teachers would. Although the variance in the responses to this question is not large, they may merit further exploration because 25% disagreement among teachers suggests that some students may be impacted by differences between teacher grading decisions and that contributes to unfair assessment of student work. In contrast, a high degree of consensus indicates that participants

thought teachers would show students who have been ill more leniency when it comes to assessing work that is completed late. For example, only 2 (6%) indicated that teachers might deduct marks for work being completed late in this case.

One participant provided rationale for deducting marks based on expectations that students who miss for extracurricular reasons should have work completed when they return because “they know about the event beforehand.” Another participant noted that in some cases teachers “might give harder [tests]” when students return; an observation that appeared only in relation to hockey or hunting absences, not illness.

Responses also differed within absences not related to illness. For an Indigenous student going hunting, one participant indicated that most teachers would assess the work in the same way but added that “the issue is when students DO NOT complete the classwork.” This concern did not appear in responses related to hockey or illness.

The next two questions examined teacher responses when marking a test and it became apparent that a student had not learned some of the assessed material. The variable is whether the student or the teacher is responsible for the missed learning. Results are presented in Table 3.

Table 3*Participants' Responses (n, %) to Scenarios Related to Tests and Students' Opportunity to Learn*

Responses	The teacher recalls that the:	
	Student missed some lessons	Teacher rushed teaching some material
Leave it blank in the gradebook.	1 (3%)	0 (0%)
Grade the student's test the same as the other students.	26 (79%)	6 (18%)
Omit some test items from the grade calculation for this one student's test.	2 (6%)	3 (9%)
Omit some test items for all students who did the test.	0 (0%)	22 (67%)
Allow this one student to rewrite the test.	2 (6%)	0 (0%)
Allow all students to rewrite the test.	0 (0%)	2 (6%)
Other	2 (6%)	0 (0%)
Total Responses	33 (100%)	33 (100%)

The results in Table 3 indicate that decisions about grading consistency and test modification are dependent on context. While nearly half believed that most teachers would generally grade all tests consistently, this perception shifted depending on the source of the learning gap. Of the participants (N = 33), 26 (79%) indicated that teachers would grade tests consistently when students had missed lessons, compared to only 18% when instruction had been rushed. One participant noted that in the latter case, some teachers might “change the total value

of the test, realizing that it may have been too difficult for most,” suggesting greater willingness to adjust assessment when responsibility is perceived to rest with the teacher.

Decisions about removing test items further illustrate this variability. Omitting items for individual students was widely viewed as unlikely, regardless of circumstance. In contrast, 67% indicated that most teachers would remove items for all students if they realized during marking that some content had not been taught or covered thoroughly, whereas none believed teachers would do so if students had missed lessons. One participant explained that when a pattern emerges across multiple students, “most teachers might consider omitting that question for everyone.” Test rewrites were rarely seen as an option, with only two respondents indicating that teachers would consider them.

Reasons for student absences emerged again as influential. Participants suggested that they believe teachers differentiate based on why instruction was missed, with illness more likely to result in opportunities to learn the material and “potentially retake the test.” Together, these patterns indicate that teachers’ efforts to ensure fairness rely on context sensitive professional judgment rather than uniform assessment practices.

In order to explore teachers’ perceptions about formative assessment, participants were asked about the use of feedback. The variables included providing students with feedback and the opportunity to improve an assignment before it is graded, grading student work with no written feedback, and giving students feedback that could only be used on future assignments. The findings are presented in Table 4.

Table 4*Participants' Responses (n, %) to Scenarios Related to Feedback on a Formative Assessment*

Responses	Students can use feedback to improve before work is graded.	Student's work is graded with no written feedback.	Students are given specific feedback that can only be used on future assignments.
The practice is fair	25 (76%)	6 (18%)	29 (88%)
The practice is unfair	0 (0%)	17 (52%)	1 (3%)
This practice is unrelated to fairness	7 (21%)	9 (27%)	3 (9%)
Other	1 (3%)	1 (3%)	0 (0%)
Total Responses	33 (100%)	33 (100%)	33 (100%)

Most participants indicated that they thought teachers would view providing students with feedback as fair, while withholding written feedback would be considered unfair. Although 76% reported that teachers would see it as fair to give feedback and allow students to revise their work before grading, participants questioned how commonly this practice occurs. One noted that it “is not done enough because it causes an increased workload for the teacher, as they must now assess the assignment twice,” suggesting a gap between beliefs about fairness and actual practice.

Responses to the scenario involving grades without written feedback were more divided. While 52% indicated that most teachers would view this practice as unfair, 18% believed it

would be seen as fair, particularly if a rubric or clear success criteria were included with the assessment. One emphasized consistency, stating “If it is a consistent practice, it’s probably unfair as students wouldn’t be getting the kind of feedback they need to progress, but potentially assignments are taken up as a class.”

There was stronger agreement around feedback that supports future learning. One participant explained that this approach is common, “it’s almost exclusively how we mark,” but acknowledged that revisiting earlier grades to reflect improvement “takes time and is often overlooked.” Notably, fewer participants (77%) viewed allowing students to revise work before grading as fair, compared to 88% who see it as fair to use feedback only to improve future assignments, again pointing to workload constraints.

Overall, participants largely agreed that feedback is important for fairness in assessment, but their responses reveal a tension between beliefs and practice. Time and workload emerged as significant factors shaping assessment practices, suggesting that even when teachers value feedback as it contributes to fairness, constraints may limit their ability to enact it consistently.

The next three questions on the questionnaire asked participants what they thought teachers would view as the optimal grade that would represent a student’s knowledge or performance. No personal characteristics were provided about the student(s) in these scenarios. The variables were different sets of scores earned by a student in chronological order. The scenarios and results are presented in Table 5.

Table 5*Participants' Responses (n, %) to Scenarios About Determining the Optimal Grade*

Responses	A student has the following four scores in chronological order:		
	0%, 60%, 60%, 80%	82%, 50%, 82%, 85%	62%, 50%, 62%, 85%
The average score.	6 (18%)	10 (31%)	21 (66%)
The most recent score.	2 (6%)	2 (6%)	2 (6%)
The mode and median score.	22 (67%)	19 (60%)	9 (28%)
Other	3 (9%)	1 (3%)	0 (0%)
Total Responses	33 (100%)	32 (100%)	32 (100%)

The degree of variance in these results is remarkable. Across all three grading scenarios, there is no consensus on the optimal grade, with plausible marks spanning as much as 30 to 35 percentage points. In the first scenario, where a zero was present, participants indicated overall grades ranging from 50% to 80%, with most indicating that teachers would rely on the median or mode rather than the average. In the second scenario, which did not include a zero, participants were more evenly divided between using the average and the median or mode, and the range of grades narrowed to 75% to 85%. In the third scenario, the majority of participants (66%) indicated that most teachers would use the average, yet overall grades still ranged widely, from 62% to 85%.

Participants appeared to struggle with this section of the questionnaire and offered comments explaining their reasoning. In the first scenario, one participant noted that the “average of the three scores without the zero” would reflect the optimal grade, yet observed that “most just put all marks in.” In the second scenario, a participant who believed the mode would best represent student performance remarked that “unfortunately our marks programs would assign the average.” Both comments highlight the common practice of entering grades into computerized systems that calculate averages based on weighted categories of student work. In the final scenario, a participant stated, “marks programs would give the average, and I am not sure that most teachers would look at individual marks like this,” offering further insight into actual grading practices.

These patterns indicate that teachers’ determinations of an overall grade vary considerably depending on the distribution of scores and the presence of extreme values (outliers), reflecting differences in how mathematical processes and professional judgment are applied. This scale of variability demonstrates how construct irrelevant variance can emerge purely from mathematical choices themselves before any student characteristics or social contexts are considered. However, these results also suggest a gap between beliefs and practice in terms of fair assessment. While participants indicated here that the average does not always best represent student knowledge or ability, they also explained calculating the average remains a standard approach in day-to-day practice.

The next five questions in the questionnaire focus on two students with different individual characteristics. The main variables are the profiles of these learners. The first question asks participants to consider whether the same percentage grade for both students truly

represents their achievement. The findings from the first in the series of questions are presented in Table 6a.

Table 6a

Participants' Responses (n, %) to a Scenario Where Different Students Earn the Same Grade

Thomas is a 'dream student'. He participates in class and always puts forth effort. Thomas does very well on classwork and assignments, but poorly on tests. Dylan is a 'challenging student'. When he comes to class, he rarely participates and when he does, it seems like it is his mission in life to cause trouble. His classwork and assignments show very little effort and are done poorly, but he does extremely well on tests. After applying category weightings for classwork, assignments and tests, both of these students have an average grade of 75%. What observations would most teachers make regarding their assessment practices?	
Responses	
Classwork and assignments should be weighted higher because Thomas' score should be higher than Dylan's.	4 (13%)
Tests should be weighted higher because Dylan's score should be higher than Thomas'.	0 (0%)
Change nothing. This is how grades work out sometimes.	27 (84%)
Other	1 (3%)
Total Responses	32 (100%))

When presented with this scenario, 84% of respondents indicated that they felt most teachers would not change their grading methods, accepting the grades as they stand. This is in contrast to the results in the previous section which indicated a preference in using professional judgment to arrive at an optimal grade rather than solely relying on the average. Here, with only student profiles and overall grades but no individual scores, participants expected teachers to default to averages, suggesting a shift from judgment-based assessment toward calculation-based grading. It is possible that due to their professional judgment, participants believe that difficult behaviour should not affect grading.

A few participants suggested teachers might adjust the weighting so Thomas would have a higher grade than Dylan, while none thought that changes should favour Dylan. One participant selecting “other” explained they “would use a point system and have in-class activities that would benefit students who came [attend], like Thomas,” indicating a desire to reflect work habits and behaviour in Thomas’ grade.

The next four questions refer to the same student profiles for Thomas and Dylan, but both students have an overall grade of 78% and participants were asked if they felt teachers would bump their grade to an 80%. In addition to the learner profiles, new variables include different circumstances in order to gauge how assessment practices might be impacted by societal pressures. The details of the circumstances and responses to these four questions are summarized in Table 6b.

Table 6b*Participants' Responses (n, %) to Scenarios Related to Professional Judgment and Relationships*

Responses	Acceptance to post-secondary		Mom is an old friend	
	Thoma s	Dylan	Thomas	Dylan
Keep the marks as they are. Teachers don't adjust marks without evidence.	5 (15%)	11 (33%)	14 (42%)	15 (46%)
Increase student's mark up to 80%.	19 (58%)	9 (27%)	12 (36%)	10 (30%)
Adjust the weighting of an assignment or two in order to bring student to a legitimate 80.	9 (27%)	8 (24%)	3 (9%)	4 (12%)
Increase all student grades 2%.	0 (0%)	2 (6%)	1 (3%)	1 (3%)
Other	0 (0%)	3 (9%)	3 (9%)	3 (9%)
Total Responses	33 (100%)	33 (99%*)	33 (99%*)	33 (100%)

Note: There was a typo in the questionnaire with the scenario related to bumping Dylan's mark up to make him eligible for post-secondary admission. The option for the response for Dylan said, "increase Thomas' mark up to 80%". Seven participants mentioned the error in the comments and answered the question as if the second option in the responses said "Dylan".

*The total for responses is 99% due to rounding error.

When asked about adjusting student marks for post-secondary prospects, participants consistently expected teachers to favour students with positive classroom behaviour. Most

believed Thomas' mark would be increased (85%), compared to only about half for Dylan (51%), indicating a clear disparity based on perceived effort or likability. When asked about adjusting marks for a parent who is an old friend, responses were more evenly split, with roughly 42-45% predicting a grade increase, highlighting the potential influence of social relationships on grading decisions. One participant who thought teachers would adjust Thomas' mark stated that "this practice is unfair to other students who do not ask for extra marks and who are not 'liked' by teachers." They explained that they chose this response because, they "believe this practice happens often."

These patterns suggest that professional judgment is sometimes influenced by student effort and behaviour as well as social contexts, rather than solely by demonstrated knowledge or ability. Although the results related to the scenario about increasing Dylan's mark may have been impacted by the error in one of the available responses to the question, the comments give us some insight into teachers' practices and thought processes and the variance that can happen. For instance, one participant commented that they thought teachers would, "bump up Dylan's mark to 80", but another said, "Dylan doesn't work hard to deserve to get bumped." Taken together, these results show that social factors including teacher perceptions of work ethic, personal relationships, and power dynamics can introduce construct irrelevant variance, complicating the fair and ethical use of professional judgment in assessment.

In the final two questions on the questionnaire, participants were asked about accommodations for a student without an Individual Educational Plan (IEP) who struggles with test anxiety. The scenario, questions and responses are summarized in Table 7.

Table 7*Participants' Responses (n, %) to Scenarios Related to Alternative Methods as Accommodations*

Responses	Alternative Assessment Method	Alternative Location
Teacher A because an alternative assessment method/alternative location is considered differentiated support.	6 (18%)	17 (52%)
Teacher A because the student's score will be negatively affected by their anxiety.	2 (6%)	8 (24%)
Teacher A for a reason other than those listed above. (comment below)	0 (0%)	0 (0%)
Teacher B because this student would have an unfair advantage over other anxious students who are writing the test.	19 (58%)	6 (18%)
Teacher B because writing a test in standardized conditions is an important part of fairness.	4 (12%)	2 (6%)
Teacher B for a reason other than those listed above. (comment below)	0 (0%)	0 (0%)
Other	2 (6%)	0 (0%)
Total Responses	33 (100%)	33 (100%)

Most participants (76%) felt that teachers would feel that providing an alternate location was a fair accommodation for a student with test anxiety, while only 24% viewed an alternative assessment method (assignment) as appropriate. Furthermore, 58% indicated that using an alternative assignment would be an unfair advantage, suggesting that participants see location adjustments as fair, but alternative assessment methods as potentially inequitable. A few

participants noted that teachers should work with students to manage anxiety with one person stating that “anxiety is part of life, and testing experience will help the student get over it.”

These results indicate that beliefs about fairness differ depending on the type of accommodation, potentially reflecting variation in teacher worldviews, assessment philosophies or professional knowledge.

At the end of the questionnaire, participants were asked whether they agreed or disagreed with a statement designed to explore teacher practices and beliefs about factors related to fairness in assessment. The primary goal was to examine how teachers’ perceptions align with or diverge from fairness principles described in *The Standards*, ministry documents, and current research literature. To observe the extent of consensus (or lack thereof) in teacher perceptions and practices as they relate to key fairness principles, results are presented from the statements in order showing the greatest amount of variance to those showing the greatest consensus. A summary of the results is presented in Tables 8a, 8b and 8c.

Table 8a

Participants' Responses (n, %) Related to Assessment Practices Demonstrating Maximum Variance

	Responses		
	Agree	Disagree	Total Responses
Most teachers are aware of systemic bias and attempt to address it in their classroom.	16 (50%)	16 (50%)	32 (100%)
Most teachers decide unilaterally how students will demonstrate knowledge in their classroom.	22 (67%)	11 (33%)	33 (100%)
Most teachers use grades as a reward for good behaviour.	11 (33%)	22 (67%)	33 (100%)
Most teachers believe that there are some students who don't have the ability to learn what they're teaching.	21 (64%)	12 (36%)	33 (100%)

The greatest variance emerged around systemic bias with roughly half of participants indicating that most teachers are aware and attempt to address bias, while the other half did not share this view. Since the definition of fairness states that assessment should be free from bias and provide students with equitable opportunities, teachers' views on systemic bias are likely to impact their perceptions of fairness in assessment.

Similarly, participants were split regarding views about student ability with 64% indicating that teachers believe some students lack the ability to learn. Although we don't know why this view is expressed, it might coincide with teachers' views about streaming or recommending course changes, which are explored in Table 8b.

Most did not think teachers use grades as rewards. The 33% who think teachers do reward students with grades aligns with results from previous scenarios where teachers demonstrated a desire to reflect work habits and behaviour in grades.

Responses to these statements had rates of agreement or disagreement between 50% and 67% meaning that 33% to 50% thought differently, showing the highest degree of variance in teachers' views and practices. These findings highlight substantial differences in teacher beliefs about fairness, assessment practices, and student potential. Take together, these results indicate that perceptions of systemic bias, decisions about how students demonstrate knowledge, and beliefs about ability can contribute to variance in assessment practices potentially leading to inequities.

Table 8b

Participants' Responses (n, %) Related to Assessment Practices Demonstrating Moderate Variance

	Responses		Total Responses
	Agree	Disagree	
Most teachers will adjust their methods if some students are not doing well, rather than recommend a course change.	7 (21%)	26 (79%)	33 (100%)
Most teachers communicate information about an assessment at a level of detail appropriate to students.	25 (76%)	8 (24%)	33 (100%)
Most teachers think that streaming or tracking students will impact their learning opportunities negatively.	9 (27%)	24 (73%)	33 (100%)
Most teachers have good knowledge of fair assessment practices.	23 (70%)	10 (30%)	33 (100%)
Most teachers don't have time to think about fairness in assessment.	10 (30%)	23 (70%)	33 (100%)

Participants showed strong consensus (79%) that teachers would recommend a level change rather than adjust their methods when a student is struggling, and most (73%) did not view streaming as negatively impacting learning opportunities. This pattern aligns with earlier findings that some teachers perceive that not all students as having the ability to learn what they are teaching. From these results, it is difficult to tell if this is due to systemic reasons such as the historical practice of streaming, or an indication of teachers' personal views on students' ability to learn.

Many respondents (76%) indicated that most teachers communicate information about assessment with a level of detail that is appropriate for students, and 70% agreed that teachers have good knowledge of and time to think about fair assessment practices. While positive, these results indicate that there is still room for improvement as approximately one-quarter to one-third of participants reported that communication and knowledge of fair practice in terms of assessment is lacking.

Responses to these statements had rates of agreement or disagreement between 70% and 79% meaning that 21% to 30% thought differently, showing moderate variance in teachers' views and practices. Despite ongoing efforts to improve equity in systems of education, these moderate levels of variance in terms of views about student ability and knowledge of fair assessment practices, indicate that traditional practices, a lack of professional knowledge or systemic factors may continue to influence teacher's views and decision-making impacting fairness in assessment.

Table 8c

Participants' Responses (n, %) Related to Assessment Practices Demonstrating Minimum Variance

	Responses		Total Response s
	Agree	Disagre e	
Most teachers respond more favourably to students who follow classroom rules.	33 (100%)	0 (0%)	33 (100%)
Most teachers connect more easily with students with whom they have things in common.	31 (94%)	2 (6%)	33 (100%)
Most teachers use grades as a punishment for bad behaviour.	4 (12%)	29 (88%)	33 (100%)
Most teachers assess students fairly.	27 (82%)	6 (18%)	33 (100%)
Most teachers replicate grading systems that they know (i.e., grading systems that are similar to what they experienced in school).	27 (82%)	6 (18%)	33 (100%)

The strongest consensus among participants was observed in the context of teacher-student relationships. All (100%) participants agreed that teachers respond more favourably to students who follow the rules, and nearly all (94%) believed teachers connect more easily with students they share commonalities with. While this level of agreement shows reduced variability in teacher views and practices, it also raises serious concerns. For instance, it is remarkable that many participants (82%) indicated feeling that teachers assess fairly, while at the same time acknowledging these types of favouritism. These results indicate the potential for inequities that could affect fairness.

Few participants (12%) believed teachers use grades as punishment, however these results offer no indication of whether participants would view assigning zeros for missed work or deducting late marks as punishments or not.

Responses to these statements had rates of agreement or disagreement between 82% and 100% meaning that only 0% to 18% thought differently, showing the greatest amount of consensus in teachers' perceptions. These results indicate a disconnect between perceived fairness and interpersonal biases, suggesting that social factors may contribute to inequities. The finding that most teachers (82%) replicate grading systems they experienced as students themselves indicates that traditional methods of assessment continue, which can reinforce entrenched historical beliefs. Taken together, these results suggest that when teachers form stronger connections with certain students, and rely on traditional methods, it may inadvertently undermine equity and fairness in assessment.

Phase Two: Qualitative Results

The qualitative data in phase two were obtained from individual semi-structured interviews. Eight secondary school teachers from various disciplines of study were interviewed via Zoom. All participants are identified with the following pseudonyms: P1, P2, P3, P4, P5, P6, P7, P8. They agreed that their gender and years of experience could be discussed in the results where convenient or relevant, if confidentiality was protected.

The themes, codes, and factors influencing fairness in assessment identified in the literature were explored to determine how teachers' perceptions align with or diverge from those described in *The Standards*, ministry documents, and related research. Additional factors emerged during the interviews that were not emphasized in the literature review. Most notably,

participants identified class size, teacher workload, and cheating or student use of AI as barriers to fairness particularly in the ways they affect teacher practice. Teachers reported that increasing class sizes intensify workload and time constraints, which in turn affect instructional and assessment practices and limit students' opportunities to learn. Cheating and the use of AI were described as contributing to variability in student scores, thereby introducing other sources of construct irrelevant variance. The interview findings are organized into the following three sections: Opportunity to Learn, Construct Irrelevant Variance, and Teacher Worldview.

This format mirrors the literature review and focusses on the factors the educators perceived as the greatest impacts on fairness in assessment.

Opportunity to Learn

Factors affecting opportunities to learn, such as student engagement, student absences, marginalized groups and teacher practice, were topics of discussion in the interviews. Class sizes, although not examined in the literature, emerged as another factor impacting fairness in assessment when it comes to teacher practice and students' opportunities to learn.

Student Engagement: Student Absences

Educators consistently highlighted student absences as a major challenge for maintaining fairness in assessment. Many rely on virtual platforms to keep students connected, yet tracking missed work remains difficult, especially for students absent 20 to 30 days. P1 noted, "It's really hard to keep track...attendance is horrific... that's a lot of school." Several teachers admitted to not strictly following school policies for deducting marks on late assignments, citing heavy workloads and concerns that teachers are not enacting the policy consistently, making it unfair. P4 explained, "If kids are away for most of the unit, you want to make sure that the assessment is

still fair for them when they have time to get caught up.” Many acknowledged the tension between accommodating student needs and reinforcing the importance of deadlines, particularly when students may be intentionally delaying submissions in the absence of consequences.

These findings illustrate the inconsistencies in handling absences (and late work) as teachers try to balance individual student needs while maintaining practices that are consistent and fair. Inconsistent practices may contribute to construct irrelevant variance and differential opportunities for students.

Student Engagement: Marginalized Groups

Three of eight educators interviewed highlighted the challenges of fairly assessing students from marginalized groups, particularly when accommodations or IEPs (Individual Education Plans) are involved. P7 described feeling disconnected from the supports already in place for these students.

It would feel nice to be more, I think, connected as a team. When there are extra supports that are already in place, I think it would feel nice to actually kind of feel a part of that. I don't right now necessarily...I don't feel like I'm a part of that. I feel like I'm, you know, off to the side.

Another emphasized the need for more guidance on addressing special needs and communicating information to classroom teachers about mental health histories.

Despite these challenges, participants articulated their strong commitment to honouring student differences and improving engagement. P2 felt that, “whether it’s background, whether it’s cultural, whether it’s like life events, I think our whole goal is for the betterment of the student.” Several educators shared practical strategies, such as offering oral assessments,

involving students in setting success criteria and deadlines, and providing choices in assignments, which they found enhanced participation.

These findings suggest that teachers' awareness of student identity, culture, and ability impacts student engagement and opportunities to learn, and that professional judgment and the appropriate use of accommodations play a key role in enacting fair assessment.

Student Anxiety

Participants identified student anxiety and personal circumstances as contributors to unexplained variance in assessment outcomes. Although this area was not explored directly in follow-up interviews, educators raised concerns that traditional assessment methods may disadvantage students experiencing anxiety and may not accurately capture their learning.

Several participants described experimenting with alternative assessment approaches and perceive these as yielding more valid representations of student understanding, particularly for anxious learners. Others emphasized the importance of contextual knowledge being shared with classroom educators, noting that temporary life stressors can affect performance and complicate judgments about student achievement. These comments suggest that teachers believe that student anxiety introduces construct irrelevant variance and challenges their ability to assess student learning consistently and fairly.

Teacher Practice

Educators highlighted that systemic constraints, time pressures, and large class sizes significantly shape assessment practices. Most participants (7 of 8) described entering grades directly into computerized systems without reviewing patterns on paper, largely due to workload and policy restrictions. P8 noted, "most people would just... blindly put it [a grade] into the

program... it is fair because that's what everyone does." Some participants expressed concern about whether grades are entered accurately and whether errors are noticed when using this method, highlighting limitations imposed by existing systems and routines.

Time pressures, such as providing interim reports, further constrain teachers' ability to engage in meaningful formative assessment. Six of eight participants described that grades alone, without feedback, is often the result of workload. As P1 explained, "We're super constrained in high school by time... often students don't even read the feedback you give them." Others explained that feedback might be read if student work was ungraded, but current procedures require marks to be entered into the system. P7 described "a feeling like I just have to get more in; it's quantity, it's not quality." Participants described merging formative and summative assessment or defaulting to summative measures as common strategies to address these pressures. Although some questioned the fairness of these methods, P8 described feeling a "familiarity with things you've always done."

Growing class sizes exacerbate challenges in providing equitable learning opportunities and fair assessment. As P1 reflected:

When I first started teaching, I had a class of 22 students. Now I regularly have classes of 26, 27, 28; 16 or 17 of them with IEPs and I have zero support. Like it is not a recipe for success for anyone and so it's unfair. There's no way that you can follow those IEPs, there's no way that you can give 17 students 'proximity to the instructor'. Our students have less access to their teacher. By basically any measure, our students' learning conditions, I think are worse than they were 20 years ago."

Difficulties in accommodating students' individual needs, monitoring attendance, and providing timely feedback, indicate that structural pressures can undermine teacher practice limiting opportunities for all students to have equitable access to meaningful support in order to learn and demonstrate learning.

Throughout the interviews all participants described feeling restricted by current practices or policies yet continually expressed a desire to know if what they thought was "right" and wanted to know how they might improve their current assessment practices. P4 expressed dislike for current practices.

Having to put a percentage and number on a student is hard... what's the difference between an 83 and an 84? But, we have to do a percent... it limits how we have to mark; we have to put a number at the end of the day on something.

The interviews revealed a general openness to discussing current practices and exploring ways to enhance fairness, showing passion for the profession, and a willingness to adapt both individually and collectively.

Participants emphasized transparency, clear expectations, and ongoing communication as key strategies for promoting fairness. Teachers described ensuring students understand expectations, providing opportunities to ask questions, and returning graded work in a reasonable time frame. Some specific strategies included offering oral assessments, allowing multiple attempts to demonstrate learning, and giving feedback prior to assigning grades. P3 explained: "My strategy with marks is to be constantly showing students how they're doing," while P8 described moving away from traditional methods stating, "More marking doesn't mean more learning."

These findings illustrate that while teacher practice is heavily influenced by systemic and structural pressures, professional judgment and adaptive strategies can mitigate construct irrelevant variance and inequities imposed by such pressures. Although it is evident that there is a need and a desire to improve professional knowledge, current practices like transparency, ongoing communication, and flexible assessment approaches are being enacted to support equitable opportunities for students and enhance fairness in assessment.

Construct Irrelevant Variance

Mathematical Processes

During the interviews, participants were asked to review the questionnaire results concerning a student who received grades of 0, 60, 60, and 80 in chronological order on individual assessments. Every participant expressed surprise that some teachers might use the mode or median to calculate a student's grade, noting that averages are typically used to calculate overall grades. P4 reflected, "I'm kind of surprised some people said 80, the most recent score, because I don't feel like that's what we do at all." Others highlighted the tension between the numerical average and what they felt would best represent a student's learning trajectory, with P8 noting, "it looks like this student is doing what we want them to do, so why should they be penalized?" P6 was troubled that "many of us are technically giving this 50 because of the program we're using." These reflections reveal a conflict between common grading procedures and professional judgment, illustrating how mathematical processes can create construct irrelevant variance if they fail to represent student achievement accurately.

All participants reported that teachers generally enter grades directly into computerized systems without reviewing trends on paper. P2 explained, "the mark is the mark," noting that averages are simple to calculate and explain. "You go by that straight numerical average because

you don't want to be perceived as being unfair." Some mentioned that this is easier to explain to parents and it is the process that they are expected to follow. Some teachers shared that they can often tell when students understand the material, through conversations and daily participation, even when their marks fail to reflect that understanding. While convenient and standardized, participants acknowledged that the current common practice may not accurately reflect student learning, highlighting the limitations of solely crunching numbers.

Assigning zeros emerged as a complex issue, highlighting significant differences in beliefs and practices among teachers. Participants explained that zeros can act as warnings, motivators, or signals of missing work, but acknowledge that it does not reflect a student's true knowledge or ability. "The zero does not show a picture of their grasp of the subject or content," P8 stated. P6 agreed that when using averages, "it completely skews that mark." Five of eight participants said that teachers will input a zero and give the student a chance to complete the work. "That compassion piece comes in," P1 explained, "like maybe they're not legitimate reasons, as in, they're not excused absences, but a lot of our students have big life things going on." Three participants thought that most teachers would end up changing a zero to a 'no mark' if the work is never completed, meaning that the zero would not be included in the calculation of the student's average. However, several non-entries in the gradebook can become a problem. Teachers are required to give a number to reflect a student's achievement at some point, so they might enter zeros if the student does not complete the work. "It doesn't mean that they don't know, or that they couldn't, but they've been given the opportunity to, and they didn't," P1 explained. Teachers described balancing fairness for one student with consistency for the class, often weighing whether a zero should remain, be converted to 'no mark,' or if they would allow a makeup assignment. P5 said, "I need to be fair and equitable. If I'm going to give you a

makeup assignment, I need to open it up for the whole class”. These conversations make it evident that teachers are concerned about how zeros are handled but lack awareness of alternatives like minimum grading.

Overall, participants highlighted a tension between procedural mathematical practices (using averages, computerized systems, and assigning zeros) and the professional judgment needed to represent student learning fairly. Decisions about which measures to use, whether to include zeros, and how to interpret trends demonstrate that systemic constraints, lack of professional knowledge and teacher beliefs about assessment contribute to construct irrelevant variance, thus impacting the fairness.

Variance from Teacher to Teacher

Differences or inconsistencies in assessment practices among teachers often stem from their individual perceptions of fairness, which can differentially impact students’ experiences. When prejudice influences assessment or students are denied equitable opportunities, fairness is compromised. Exploring teachers’ perceptions of fairness in assessment can offer valuable insights into the roots of these variations and their impact on fair assessment.

Participants consistently (7 out of 8) emphasized that fairness is subjective and context dependent. Although P3 explained that “fair means different things in different situations,” they expressed surprise at the degree of disagreement revealed in questionnaire responses, particularly around whether providing a grade without feedback is fair. This highlights variability in how teachers conceptualize fairness.

Participants reacted strongly to the wide range of grades (50-80%) teachers might assign in the same hypothetical grading scenario. P8 remarked, “Barely a pass and an Ontario scholar,

like that's huge." Another questioned the fairness of such variance among teachers teaching the same course. Although participants acknowledged the necessary role of professional judgment and articulated that fairness can be subjective, they expressed discomfort with the range of inconsistency in how judgment is exercised and the scale of construct irrelevant variance this introduces.

Consistency emerged as a central concern related to fairness. When asked whether they believed most teachers shared a common approach, seven of eight interviewed said, "no." P6 explained that "we try to be on the same page, but some teachers like to give students breaks." Most participants believed that teachers do not share a common approach to practices such as assigning zeros or determining student workload. For example, participants explained that the number of assignments across sections can vary, highlighting the lack of consistency in the same school and even within departments. While teachers valued consistency, they also cautioned that uniformity alone does not guarantee fairness. "It's hard when everyone's kind of doing their own thing and then, you know, you kind of mask it under this professional discretion... consistency is important as long as that consistency is actually rooted in [knowledge of] fair assessment," P2 explained. Participants expressed an understanding that sameness, or certain types of consistency do not necessarily contribute to fairness.

A lack of shared understanding, stemming from limited training and inconsistent guidance, was described as a barrier to fair assessment by all interview participants. Several (6 of 8) noted that they had never been formally taught how to assess or grade, and others expressed uncertainty about what fair practice looks like in concrete terms. P4, an educator with more than fifteen years of experience said, "No one has ever shown me how to mark a test." Many suggested that professional learning focused on assessment purposes and best practices could

reduce variance and support more equitable grading decisions and fairness overall in assessment practices.

Although participants were aware that teachers' conceptions of fairness and assessment practices vary, they were surprised by the magnitude of variance revealed in the results of the questionnaire. These findings suggest that differing interpretations of fairness, inconsistent application of professional judgment, and limited guidance encouraging shared philosophies and practices negatively impact fairness in assessment by contributing to construct irrelevant variance and inequitable student experiences and outcomes.

Cheating or Student Use of AI

Participants identified academic dishonesty, including student use of AI, as an extraneous factor that can distort the meaning of grades. Educators expressed concern about whether submitted work accurately reflects a student's own learning, particularly when assessments are completed outside of supervised settings. P6 asked, "Who are we evaluating? Is there even a purpose to evaluate at this point?" This was expressed a factor impacting fairness because, "I always think of the one who worked really hard and ended up with a 70. And then there's this person who didn't do anything other than go to AI and get the answer."

Several participants questioned the validity of evaluating work when authorship is unclear, noting the fairness implications for students who complete work independently. Others highlighted the difficulty of distinguishing between legitimate support and unethical assistance when students complete work at home particularly when they have been absent. Cheating and student use of AI introduce construct irrelevant variance and thus are important concerns related to fairness in assessment.

Teacher Worldview

Participants identified teacher authority and control over assessment as a key factor influencing fairness, particularly for marginalized students. Several educators acknowledged that assessment decisions are shaped by personal values and implicit prejudice, often without conscious awareness. P5 explained,

I guess fairness just really boils down to the teacher's own personal values and what they believe is fair. Or maybe, if they were a student and they were assessed in a certain way, maybe they've stuck with that or done the complete opposite because they felt maybe they were unfairly treated by a teacher. Who knows, but I just think it boils down to your personal values and beliefs and philosophies.

Participants' reflections highlight that perceptions of fairness in assessment are deeply personal and often shaped by educators' own experiences and personal views. Their beliefs inform what they perceive as fair suggesting that worldview functions as a lens through which fairness principles and knowledge of assessment is perceived before practices are enacted.

The use of grades as a form of control or punishment emerged as a concern. When discussing a grading scenario that included a zero, participants questioned whether the mark reflected student learning or a type of punishment. Participants explained that teachers tend to stick to what they know or fall back on what they have always done. P3 explained that "there's people that would look at this situation (0, 60, 60, 80) and say I'm not going to drop the zero because that's going to inflate the kids' mark. But there might be research; there might be tons of research that shows that that's not the case, and people are just hanging on to practices because that's all they know." Educators noted that such decisions are often justified through professional judgment that actually reflect habitual practices rather than sound professional knowledge

supported by research, potentially contributing to construct irrelevant variance and inequitable opportunities and outcomes for students.

Some participants contrasted these practices with efforts to redistribute power by incorporating student voice. Examples included negotiating deadlines and inviting student input into assessment decisions, which educators described as more responsive to students' lived circumstances. P1 acknowledged the importance of striving for fairness, noting that "education will never be totally fair, and that's reality. But I also don't think that means we shouldn't try as much as we can to make things more fair."

Together, these findings illustrate how power dynamics, shaped by worldviews and institutional norms, can influence assessment practices in ways that either perpetuate or mitigate inequities, thus impacting fairness. These power dynamics often come into play when teachers exercise professional judgment.

Teacher Professional Judgment

Participants described professional judgment as a reflective and context dependent process that shapes assessment decisions beyond the mechanical application of policy. Several educators emphasized the need to re-evaluate assessments when results suggest misalignment between instructional intent and student performance, such as revising or removing test items when widespread difficulty indicates an issue with the design of the assessment tool or method. "Let's say, on a test, kids took too long to write it. Some didn't finish. Let's make it out of 45 instead of 50 or something like that. I might say, oh I didn't really focus on that type of question, so I'm going to take it off the test," P4, explained. P6 echoed this sentiment, noting, "If a whole class is bombing something and they usually don't bomb on it, I would just reevaluate." These

examples illustrate professional judgment as an interpretive act aimed at improving the validity of evidence used to represent student learning.

Hypothetical grading scenarios involving limited information and outliers, particularly the presence of a zero, were identified as moments when professional judgment was most heavily exercised. In our conversations about the student with grades of 0, 60, 60 and 80, every interview participant expressed concern about the zero grade, wanting to understand its context, the purpose of the assessment, and whether more recent evidence (80%) should carry more weight. P4 felt that “if the 80 is an exam, it should be weighted more” and P7 felt that because there are so few grades, it’s hard to say what is most fair and stated that “perhaps the 80 is an outlier.” It is important to note that educators are aware that professional judgment only goes so far. P1 gave the example of bumping a grade from a 55 to a 70 and that a practice such as that should be questioned, stating “could I provide the evidence?” The teachers interviewed indicated that determining what is fair involves weighing the purpose of each assessment, understanding the reasons behind outliers, and acknowledging that no single approach fits all situations. These conversations demonstrate how professional judgment can both mitigate and contribute to construct irrelevant variance.

Participants described professional judgment as an individual responsibility that can be challenged by external pressures. Conflicts with administrative directives were noted in cases involving truancy, academic dishonesty, or credit accumulation, yet participants maintained that only the classroom teacher has access to the full range of evidence needed to justify a final grade, as P1 described:

At the end of the day, you really need to be able to say like this kid got a 40 because they just didn't demonstrate enough understanding of all of the expectations, or ok, well, I'm bumping this up to a pass because I think that even though they didn't do a lot of these things, they did demonstrate the understanding, whether it be like orally in class or not. But you're the only one who knows that. Right? Like the principal can't tell you that; like nobody else can tell you that. Like you're their teacher in class.

They expressed an overwhelming feeling that the responsibility of fair assessment really falls on the classroom teacher.

Several educators characterized their assessment practices as evolving, particularly as they gain experience. P3 explained:

The more you teach and the more different students you come across, the more you realize you can't...you can't do things the same for every kid. Sometimes we feel like maybe we can't make as many choices as we can, especially when you're a younger teacher.

They noted that mentorship and training may support more confident and consistent use of judgment.

The findings show that professional judgment is central to fair assessment. Teachers use it to interpret scores, consider the purpose of assessments, and address outliers or unexpected results, moving beyond numerical averages to understand student learning in context. While judgment varies among educators and can be influenced by personal experiences or external pressures, classroom teachers bear the primary responsibility for making informed decisions. Educators indicated that mentorship and training can support more consistent and confident use

of judgment. Despite its potential to introduce variability, reflective and thoughtful professional judgment is essential for accurately representing student achievement, thus promoting fairness.

Social Context

Assessment takes place within a social environment, and teachers' perceptions of fairness are influenced by interactions with students, parents, colleagues, and institutional expectations. Several participants emphasized the importance of being perceived as fair, noting that social pressures, such as requests to adjust grades, can affect grading decisions. P3 explained that "you never want to seem like you're favouring a certain student," while P4 described balancing individual student needs with the perception of fairness by deciding when to extend deadlines or when students might just need an extra push.

Every kid's different and every kid needs different things. And kids don't always see it that way. So, they'll just see this kid got two extra days [to prepare for a test, or finish an assignment] and say, but I also wanted two extra days and that's not fair. So sometimes I think maybe we just extend it for everyone; just being fair. But that might not necessarily be the best thing for the kid.

Differences between teacher and student perceptions was described as another reason that teachers tend to rely on averages for calculating student grades. "My practice has always been their mark is their mark," P5 said. These examples illustrate how social context impacts professional judgment challenging teachers' abilities to ensure fairness in assessment practices.

Teachers expressed concern that personal views and adherence to traditional practices may unintentionally compromise fairness. For instance, participants described how some students might benefit from leniency due to parent involvement or personal connections thereby contributing to inequities. Several educators noted that professional autonomy and comfort with

traditional practices can discourage critical examination of the ethical implications of assessment decisions, even when those decisions might favour or disadvantage certain students. P2 stated, “when people maintain a certain belief and that impedes their approach to fairness...their own personal bias could be a challenge to equity.” These findings indicate that teachers believe that unexamined social and cognitive influences can contribute to barriers to fairness.

Collaboration with colleagues emerged as a strategy to navigate social pressures and improve consistency and fairness. Participants reported that discussing scenarios and sharing classroom experiences helped them rethink assessment decisions and challenge the status quo. P6 noted, “I think we need more conferencing between staff about what is fair” and P7 added that such conversations “get me thinking differently and about things I’d like to do.” These reflections highlight that the social context for educators themselves—through collaboration and professional dialogue—can support equitable assessment practices.

Overall, the findings suggest that fairness in assessment is negotiated within a complex social environment, where teacher-student relationships, social expectations, and collaboration among educators all interact to influence both perceptions and enactment of fair assessment.

Teacher-Student Relationships

Teachers’ relationships with students are central to fair assessment, as trust, transparency, and responsiveness shape both engagement and opportunities to learn. Participants emphasized that explaining their perspectives while considering students’ needs helps build credibility and supports fairness. P3 explained, “When you are as transparent as possible...taking into account their point of view and their needs...you are trying to be as fair as possible.” P6 noted that conferencing with students reduces complaints and strengthens relationships, supporting both

fair assessment and classroom climate. They felt this was important because they “have seen some pretty scary things where kids are like, well, I’m not that person’s favourite. I don’t want to be in their class.” These examples illustrate how relational practices directly influence perceptions of fairness in assessment. These results emphasize that fair assessment is not only about methods and marks, but also about the relationships teachers cultivate with students.

Participants also reflected on how teachers’ views related to student behaviour or shared interests can have inequitable impacts on student experiences. While reviewing questionnaire results showing that most teachers respond more favourably to compliant students or those with whom they have something in common, P7 commented, “If you’re not comfortable...if you don’t have a good environment where you feel accepted...how can you be your best learner?” P8 highlighted that “the onus is on us to find something in common with them and have a relationship with them.” They further explained that informal assessments, like quick conversations, can reveal more about student understanding than a quiz, but only when rapport exists, emphasizing that teacher-student relationships enable more accurate and meaningful assessment because students are more willing to open up and engage.

These findings suggest that teacher-student relationships function as a critical social and relational context for assessment. Fair assessment is not solely determined by methods or marks used to determine the optimal grade, but is deeply shaped by the trust, transparency, and engagement in relationships that are cultivated between teachers and students. These are essential foundations for making meaningful connections that support both learning and well-being, thus improving engagement and opportunities for all.

Chapter 5: Discussion

This study explored secondary school teachers' perceptions of fairness in assessment. The research question asked: "To what extent do teachers' responses to hypothetical assessment scenarios align with the principles of fairness in assessment, as defined in the *Standards for Educational and Psychological Testing* and Ontario Ministry of Education policies and relevant research?" By combining quantitative survey data with qualitative interview insights, the study provides a nuanced view of the complex variables that influence teachers' perceptions of fair assessment practices in Ontario's secondary classrooms.

Summary of Key Findings

Teachers' responses to hypothetical assessment scenarios showed notable areas of divergence from fairness principles. Research results revealed that teacher's beliefs about assessment and their practices can vary based on personal worldviews, their professional knowledge and experience, and the social contexts in which teachers find themselves. Specifically, this study shows that (1) inequities in opportunities to learn are created by decisions teachers make, (2) the mathematical processes teachers use to determine grades include sources of construct irrelevant variance, and (3) a teacher's worldview can introduce bias particularly in terms of how they exercise their professional judgment and form relationships with students. These findings are examined in relation to existing literature, highlighting both where they align with and diverge from established fairness principles, while offering new insights into how teachers interpret and enact fairness in assessment contexts.

Inequities in Opportunities to Learn

A central principle of fair assessment, as outlined in *The Standards* and Ontario Ministry of Education policy documents, is that students must have equitable opportunities to learn the

knowledge and skills being assessed. This principle assumes that alignment between instruction, support, and evaluation can minimize the influence of contextual inequities on student performance (APA, AERA, & NCME, 2014; Ontario Ministry of Education, 2010, 2013, 2014). However, the findings of this study suggest that this alignment is not consistently achieved in practice. Rather than functioning as a stable precondition for fair assessment, opportunity to learn appears to be variably interpreted and enacted, raising questions about how effectively this principle can be operationalized in everyday classroom contexts.

Responses to absenteeism scenarios illustrate how opportunity to learn is mediated through teacher interpretation. Quantitative results in Table 2 show that teachers believe there is greater flexibility for illness-related absences than for extracurricular or cultural activities. While *The Standards* emphasize that students should not be assessed on content they have not had the opportunity to learn, they provide limited guidance on how teachers should evaluate differing forms of absence. This ambiguity appears to be resolved through informal distinctions about what constitutes a “legitimate” reason, introducing variability into assessment decisions. Although Brookhart et al. (2016) and Hattie (2003) position teacher judgment as central to responsive and effective assessment, these findings suggest that such judgment does not operate consistently across contexts. In this respect, the study challenges the assumption that professional judgment, in the absence of shared criteria, will reliably support equitable outcomes.

A similar pattern was found in approaches to late and missing work. Participants described a wide range of practices, including assigning zeros, deducting marks, or allowing opportunities for completion. The literature on assessment for learning emphasizes flexibility and responsiveness as mechanisms for supporting student achievement (Black et al., 2004; Shepard et al., 2018). However, the findings suggest that flexibility is not systematically applied

and may instead depend on individual teacher beliefs or situational factors. This complicates the notion that responsive assessment practices inherently promote equity; without clear structures, they may instead produce uneven access to opportunities to demonstrate learning. In this way, the study extends existing literature by highlighting how variability in enactment may undermine the equity such practices are intended to support.

The role of formative assessment further underscores the gap between principle and practice. Consistent with Black et al. (2004), participants expressed strong support for formative assessment but described limited implementation due to systemic pressures such as workload and time constraints. Hattie (2007) identifies feedback as a high-impact influence on student achievement, yet these findings suggest that its integration into everyday practice remains inconsistent. This study adds to the literature by foregrounding the implications of this gap for fairness: when formative assessment is not used to inform instruction or provide actionable feedback, students have fewer opportunities to improve prior to evaluation. This raises a broader concern that the literature may overestimate the extent to which high-impact practices are enacted in typical classroom conditions.

Variation in the understanding and application of accommodations provides an additional point of tension. While *The Standards* define accommodations as adjustments that preserve the construct being assessed, participant responses in the survey data and interviews indicate uncertainty about how to apply this distinction in practice, particularly for students with needs, such as test anxiety. This aligns with Moss et al. (2008), who argue that teachers often struggle to translate assessment principles into context-specific decisions. However, the findings also suggest that accommodations are sometimes evaluated through a perceived fairness to other students, rather than through their role in supporting valid measurement. This challenges a key

assumption in the literature that fairness is primarily a technical issue of aligning supports with constructs. Results of this study suggest that it is also shaped by relational and normative judgments about equity and advantage.

These findings suggest that opportunity to learn is not simply a structural condition that precedes assessment, but a principle that is actively constructed through teacher decision-making within specific contexts. While existing literature positions teachers as central agents in promoting equitable learning conditions, this study highlights the variability that emerges when such decisions are made without clear, shared frameworks. This variability suggests a limitation in current policy and research: both tend to articulate the importance of opportunity to learn without attention as to how it is interpreted and enacted in practice. As a result, differences in access to instruction, support, and flexibility may exist across classrooms, with important implications for the validity and equity of assessment outcomes.

Mathematical Processes as Sources of Construct Irrelevant Variance

Responses to questions about determining grades reveal a misalignment with a key fairness principle outlined in policy documents and the broader assessment literature. This misalignment arises, in part, from the mathematical processes teachers use to calculate grades, which can introduce construct irrelevant variance and undermine the validity of reported achievement. Brookhart et al. (2016) note that grades may be influenced by factors unrelated to the intended construct, including how scores are calculated, recorded, and interpreted. The findings of this study support this concern and point to additional sources of construct irrelevant variance, including reliance on averages and the use of zeros in grade calculations. The degree of variation observed across teacher responses further raises questions about the consistency and validity of grading practices.

When presented with hypothetical grade data and asked to determine an overall mark, teachers demonstrated considerable variability in their decisions. Consistent with findings from Starch and Elliott (1912) and Brimi (2011), different teachers assigned different grades to the same hypothetical piece of student work. At the same time, many participants indicated that a simple average does not always provide an accurate representation of student achievement. This highlights a key tension: while professional judgment can introduce variability, rigid reliance on averaging can obscure meaningful patterns in student learning. Scholars such as Guskey (2002, 2016), O'Connor (2017), and Reeves (2000) argue that averaging may distort achievement by giving disproportionate weight to early or incomplete evidence of learning. Despite this, participants reported that averages remain the dominant method in practice, often due to digital grading systems that calculate final marks. This suggests that procedural consistency may be prioritized over interpretive judgment, allowing factors unrelated to the intended construct to influence final grades.

The practice of assigning zeros for missing work further complicates the fairness of grades. Because zeros disproportionately lower averages, they can introduce variance that reflects student compliance or circumstance rather than achievement. Participants indicated that approaches to missing work vary widely, with some teachers assigning zeros and others allowing opportunities for completion. Reeves (2004) and Carifio and Carey (2012, 2013) suggest that alternative approaches, such as minimum grading, may better preserve the integrity of achievement data. However, findings from this study suggest that many teachers are not familiar with or do not use alternatives to zeros, often relying instead on default options within grading systems. In the absence of clear policy guidance, these decisions are made individually, contributing to inconsistency across classrooms.

Two additional findings further illustrate how grading practices may introduce construct irrelevant variance. First, some teachers reported including effort or behaviour in grades, despite policy guidance discouraging this practice. Second, external pressures—such as supporting student admission to post-secondary programs or responding to parent expectations—were reported to influence grading decisions. As noted by Feldman (2019) and McMillan (2001), these factors are not uncommon but vary widely across contexts, contributing to inconsistency. In these cases, grades may reflect a combination of academic and non-academic considerations, complicating their interpretation as indicators of achievement.

Overall, these findings suggest that commonly used grading practices—particularly averaging, assigning zeros, and incorporating non-academic factors—can introduce variance unrelated to the intended construct and produce inconsistent outcomes across teachers. While some of this variability reflects the exercise of professional judgment, the results indicate that current practices may not consistently support valid and fair representations of student learning.

Teacher Worldview Introduces Bias

The findings suggest that fairness principles outlined in *The Standards*, Ontario Ministry of Education policy documents, and related research may be compromised when teachers' professional judgment is shaped by underlying worldviews that introduce bias. In this study, teacher worldview refers to the constellation of beliefs, experiences, knowledge, and assumptions that inform how educators interpret and enact fairness in assessment. It operates as a lens through which policies and principles are understood, shaping both decision-making and interactions with students.

Although professional judgment is essential for responsive and context-sensitive assessment, the results indicate that it can also contribute to inconsistency when guided by differing assumptions about students and learning. For example, responses to absenteeism scenarios suggest that teachers vary in how they determine which reasons are legitimate, leading to uneven application of flexibility or penalty. While such differences may reflect contextual considerations, they also point to the role of underlying beliefs in shaping what is perceived as fair.

Teacher knowledge and experience, which are key components of worldview, also influence how fairness principles are enacted. Quantitative findings indicate that teachers generally viewed alternative locations for students with test anxiety as fair but were less likely to view alternative assessments in the same way. According to *The Standards*, accommodations may involve changes to format, setting, or administration conditions, provided the assessed construct is not altered (APA, AERA, & NCME, 2014). The variation in responses suggests that some teachers may not have a fully developed understanding of how to apply accommodations while maintaining construct validity. This gap in knowledge has implications for students who rely on such supports to demonstrate their learning.

Relational dynamics further illustrate how worldview can shape assessment practices. Results indicate that teachers tended to respond more favourably to students with whom they shared similarities or stronger connections. This aligns with Moss et al. (2008), who note that educators may find it challenging to recognize and value perspectives that differ from their own, and with Roorda et al. (2011), who highlight the impact of teacher-student relationships on student engagement and outcomes. In this study, factors such as likability, perceived effort, and social connection appeared to influence how professional judgment was applied. As a result,

students who are more closely aligned with teacher expectations may be more likely to benefit from discretionary decisions, while others may not receive the same consideration.

These patterns raise concerns about the consistency and equity of assessment practices. When relational proximity or perceived alignment influences grading decisions, assessment outcomes may reflect more than demonstrated learning. This has particular implications for students from marginalized backgrounds, whose ways of being and learning may not align with dominant norms in schooling. Culturally responsive scholarship (e.g., Hammond, 2015; Lyiscott, 2019; Solano-Flores, 2023) highlights how such misalignment can contribute to inequitable educational experiences if not explicitly addressed.

The findings suggest that these dynamics are not solely the result of individual prejudice. Teachers' responses to scenarios involving streaming, for example, indicate that many view course placement decisions as neutral or beneficial, despite evidence that streaming can reproduce inequities (James & Turner, 2017; Terrin & Triventi, 2023). This suggests that broader systemic norms and institutional practices also shape teacher worldview, reinforcing particular interpretations of fairness and student ability. In this way, individual decision-making is situated within a wider social and structural context.

Together, these findings indicate that teacher worldview plays a central role in mediating how fairness principles are interpreted and enacted. While professional judgment allows for flexibility, it also introduces variability that can affect students' opportunities to learn and demonstrate achievement. Addressing fairness in assessment, therefore, requires not only clearer policy guidance but also critical reflection on the beliefs, assumptions, and contextual factors that shape everyday decision-making in classrooms.

The Updated Model

The findings of this study indicate that current policy documents, such as *Growing Success*, provide insufficient guidance for consistently enacting fairness principles in practice. This is evident in the substantial variability observed in teachers' responses to identical scenarios, suggesting that knowledge of policy alone does not ensure consistent application. The updated model found in Figure 2 synthesizes these findings by illustrating why policy is insufficient to guarantee fair assessment. It conceptualizes fairness not as a direct outcome of policy knowledge, but as a dynamic product of the interaction between teacher worldview, professional judgment, and systemic pressures.

Why Policy Alone Is Insufficient

The model positions guiding documents at the top of the framework, representing the theoretical foundation for fairness principles such as opportunity to learn and the reduction of construct irrelevant variance. However, it also illustrates that these principles must pass through the teacher's interpretive lens before being enacted. This explains why teachers may demonstrate awareness of fairness principles yet still describe practices that contradict them. The variability observed in the study suggests that policy guidance is insufficiently prescriptive to resolve ambiguity in practice. This aligns with Moss et al. (2008), who note that teachers often struggle to translate broad principles into context-specific decisions. The findings extend this work by showing that such struggles are not merely due to gaps in knowledge, but also to the interpretive role of teacher worldview and the influence of systemic constraints.

The Centrality of Teacher Worldview

Teacher worldview is positioned as the central mediating element through which all assessment decisions are filtered. This reflects the study's finding that fairness principles are not

applied neutrally, but are interpreted through educators' beliefs, experiences, and assumptions about students and learning. The model highlights that professional judgment is not merely a technical application of rules but is deeply intertwined with the teacher's personal values and the social context of education. This challenges the assumption in much of the assessment literature that policy alignment is primarily a technical problem of knowledge transfer. Instead, fairness is also a matter of interpretation, shaped by worldviews that are not easily altered through policy dissemination alone.

Systemic Pressures as Catalysts for Divergence

The model also introduces systemic pressures as an external force acting on both professional judgment and assessment practices. These pressures are represented not as peripheral challenges, but as conditions that shape decision-making contexts. This reframes divergence from fairness principles not simply as a lack of professional knowledge, but as a reactive response to environmental stressors. In this way, the model extends the literature by showing how systemic constraints mediate the relationship between policy and practice.

Explaining Divergence from Fairness Principles

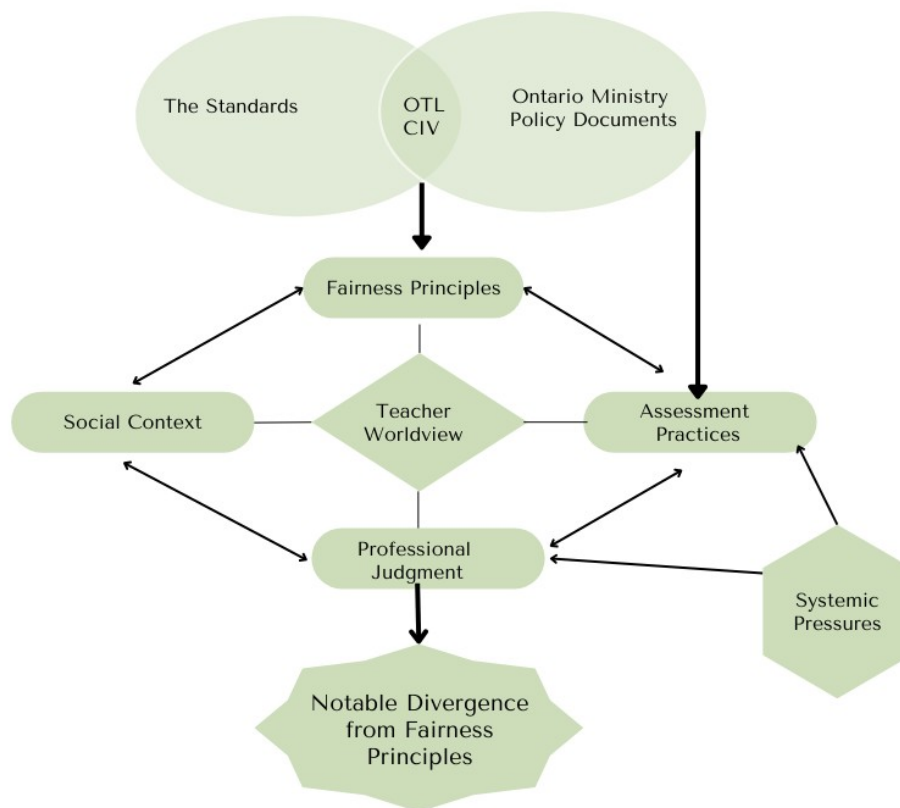
By mapping the interaction between worldview, systemic pressures, and professional judgment, the model provides an explanation for the variability observed in teachers' responses. Misalignment is not located in a single factor but emerges through the cumulative effect of how these elements interact in practice. This perspective shifts the focus from identifying isolated inconsistencies to understanding the conditions under which they arise.

Importantly, the model serves as both an explanatory and diagnostic tool. It suggests that improving fairness in assessment will require more than revised policy documents.

Strengthening policy guidance alone is unlikely to produce consistent practice without also addressing the interpretive role of teacher worldview and the systemic conditions that shape everyday decision-making. The model thus points to a more complex reality: achieving fairness likely requires attention to the dynamic relationships between policy, judgment, and context, rather than to policy alone.

Figure 2

Overview of the Dynamic Relationships Between Factors Affecting Fairness in Assessment



This model illustrates how fairness in assessment emerges from the interaction between teacher worldview, professional judgment, and systemic pressures. Teacher worldview functions as a central interpretive lens through which assessment principles and policies are understood

and enacted. Professional judgment operates as the mechanism through which decisions are made, while systemic pressures—such as workload, absenteeism, and grading structures—shape the conditions under which those decisions occur. Together, these elements influence assessment practices, including how opportunities to learn are structured and how grades are determined. The model demonstrates that divergence from fairness principles is not attributable to a single factor but arises through the combined and dynamic influence of these interacting components.

Unexpected Findings and Alternative Explanations

One unexpected finding was that some teachers appeared to endorse practices that conflict with the fairness principles outlined in policy and research, even when they demonstrated awareness of those principles. This suggests that knowledge of fairness does not necessarily translate into consistent enactment, and that teachers may rely on informal assumptions or locally developed habits when making assessment decisions.

A second unexpected finding was the way some absences were interpreted more sympathetically than others. Illness was more often viewed as a legitimate reason for flexibility than cultural or extracurricular activities. This pattern is notable because it suggests that the meaning of “valid” absence is not fixed but shaped by teacher interpretation and broader norms about which student experiences are seen as deserving accommodation.

These findings also invite alternative explanations. In some cases, inconsistencies may reflect workload, time pressure, or uncertainty about how to respond rather than deliberate prejudice. In other cases, they may reflect the persistence of institutional norms that shape what teachers see as reasonable or fair. The data suggest that these explanations are not mutually

exclusive: individual beliefs, practical constraints, and systemic expectations may all contribute to the decisions teachers make.

The unexpected findings reinforce the importance of the updated model. Rather than locating inconsistency in a single source, the model helps explain how teacher worldview, professional judgment, and systemic pressures interact to produce patterns of divergence from fairness principles.

Limitations of the Study

There are three limitations concerning the results of this study. First, it is limited by its relatively small and geographically concentrated sample, which may not fully represent the diverse range of assessment practices across Ontario.

A second potential limitation is that while hypothetical scenarios are useful for eliciting honest, reflective responses, it is possible that they may not perfectly reflect actual classroom behaviours. For instance, individuals may have answered differently on the questionnaire if only answering for themselves, but the survey asked what most teachers would do.

The third potential limitation is self-selection bias. Teachers who chose to participate may already have had a particular interest in assessment fairness, which may skew results toward greater awareness or sensitivity than exists broadly. Although these limitations are important to bear in mind, this research also yields some useful results that inform the implications for this study.

Implications for Educational Practice and Policy

This study's findings point to clear, actionable implications for educational practice and policy. Drawing on both quantitative and qualitative data, they highlight key areas for improving fairness in secondary school assessment. The implications are organized around three main areas: teacher learning and capacity, equity and culturally responsive practice, and policy reform.

1. Targeted Professional Development in Assessment

Survey and interview data revealed gaps between teachers' stated understanding of fairness and their actual practices, including inconsistent use of formative assessment, reliance on averages for final grades, and inconsistencies in assigning zeros and handling late work. Teachers also expressed uncertainty about implementing accommodations for students with anxiety or other special needs and about consequences for academic dishonesty. Professional development should therefore focus on:

- A. Applying fairness principles consistently across varied student contexts, particularly in relation to absences, late work, determining overall grades, and the use of formative assessment and accommodations.
- B. Recognizing and mitigating construct irrelevant variance in grading decisions by limiting the influence of non-academic factors such as behaviour, effort, and perceived deservingness.
- C. Using case-based exercises, such as the hypothetical scenarios used in this study, to help teachers reconcile fairness principles with everyday assessment decisions.

2. Embedding Equity and Culturally Responsive Practices in Assessment

Quantitative and qualitative findings indicate that students from marginalized groups may be disadvantaged due to implicit biases, inconsistent accommodations, or differential treatment.

Schools and teacher education programs should therefore:

- A. In consultation with representatives from cultural groups, integrate culturally responsive practices into assessment frameworks, explicitly addressing how cultural and extracurricular contexts influence opportunity to learn.
- B. Co-construct guidelines with representatives from cultural groups mentioned in A, that address particular concerns (such as absences or academic dishonesty) to improve consistency and reduce reliance on individual judgment.
- C. Provide teachers with practical strategies for accommodating individual student needs without compromising fairness or altering the assessed construct (APA, AERA & NCME, 2014).
- D. Include reflective exercises in professional development and pre-service teacher education that highlight how worldview, social context, and teacher-student relationships shape assessment decisions.

3. Review and Reform Assessment Policies

Survey data show that teachers often rely on averages, address late or missing work inconsistently, and adjust grades based on perceived effort or likability, rather than demonstrated learning (Tables 5–6). To reduce construct irrelevant variance and promote fairness, policy reforms should:

- A. Establish evidence-based grading policies at the board and school levels that align with fairness principles when ministry documents are not prescriptive.
- B. Provide explicit guidance on the use of zeros and late penalties in calculating grades, as well as appropriate methods for reporting effort or behaviour.
- C. Encourage and support formative assessment practices, including feedback that students can use to revise work before grading (Table 4).
- D. Incorporate accountability measures to monitor consistent application of fair assessment practices across teachers and classrooms.
- E. In the absence of specific direction from the ministry, school boards should create policies or guidelines that are evidence-based and align with principles of fairness.

Recommendations for Future Research

Much work remains to be done to gain a fuller understanding of how teachers perceive and enact fairness and how their perceptions and practice align or diverge from the principles set out in policy documents. Future research should explore a broader geographic range of teachers to improve generalizability. Including student perspectives on assessment and fairness could help to uncover additional dimensions and improve student experiences. Conducting classroom observations would help to compare stated beliefs with actual practices and to assess the effectiveness of any changes in practice.

In addition, research could include tracking and monitoring any changes implemented to examine whether professional development or policy changes contribute to fair assessment.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Araneda, S., Beiting-Parrish, M., Bottoms, B., La Torre, D. M., McIntee, K., Musa, S., Medinaceli, M. V. (2023, October 27). *Statement about the new edition of the standards for educational & psychological testing*. [Paper]. National Council on Measurement in Education. Graduate Student Issues Committee.
<https://doi.org/10.31235/osf.io/dv4pa>
- Aydogan, I. (2008). Favoritism in the classroom: A study on Turkish schools. *Journal of Instructional Psychology*, 35(2), 159-.
- Baniasadi, A., Salehi, K., Khodaie, E., Bagheri Noaparast, K., & Izanloo, B. (2023). Fairness in classroom assessment: A systematic review. *The Asia-Pacific Education Researcher*, 32(1), 91–109. <https://doi.org/10.1007/s40299-021-00636-z>
- Battiste, M. (2005). You can't be the global doctor if you're the colonial disease. In Tripp, P. & Muzzin, L.J. (Eds.), *Teaching as Activism* (pp. 121-133). Queen's University Press.
<https://doi.org/10.1515/9780773572348-013>
- 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., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2010). Validity in teachers' summative assessments. *Assessment in Education: Principles, Policy & Practice*, 17(2), 215–232. <https://doi.org/10.1080/09695941003696016>

- Brimi, H. M. (2011). Reliability of grading high school work in English. *Practical Assessment, Research & Evaluation*, 16(17), 17-.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803–848. <https://doi.org/10.3102/0034654316672069>
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78(3), 210–216. <https://doi.org/10.1037/0022-0663.78.3.210>
- Carifio, J., & Carey, T. (2009). A Critical Examination of Current Minimum Grading Policy Recommendations. *The High School Journal*, 93(1), 23–37. <http://www.jstor.org.ezproxy.lakeheadu.ca/stable/40363968>
- Carifio, J., & Carey, T. (2013). The arguments and data in favor of minimum grading. *Mid-Western Educational Researcher*, 25(4), 19-.
- Carey, T., & Carifio, J. (2012). The Minimum Grading Controversy: Results of a Quantitative Study of Seven Years of Grading Data From an Urban High School. *Educational Researcher*, 41(6), 201–208. <https://doi.org/10.3102/0013189X12453309>
- Childs, R. A., & Umezawa, L. (2009). When the teacher is the test proctor. *Canadian Journal of Education*, 32(3), 618+. https://link-gale-com.ezproxy.lakeheadu.ca/apps/doc/A221205321/AONE?u=ocul_lakehead&sid=bookmark-AONE&xid=f93338f3

- Chuderski, A. (2015). High intelligence prevents the negative impact of anxiety on working memory. *Cognition and Emotion*, 29(7), 1197–1209.
<https://doi.org/10.1080/02699931.2014.969683>
- Codrington, J. & Fairchild, H.H. (2012). *Special education and the miseducation of African American children: A call to action*. [Position Paper]. The Association of Black Psychologists.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches* (Fifth edition.). SAGE.
- DeLuca, C., Braund, H., Valiquette, A., & Cheng, L. (2017). Grading Policies and Practices in Canada: A Landscape Study. *Canadian Journal of Educational Administration and Policy*, 184, 4-.
- Deneen, C., & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment and Evaluation in Higher Education*, 39(5), 577–591.
<https://doi.org/10.1080/02602938.2013.859654>
- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and Cognitive Performance: Attentional Control Theory. *Emotion (Washington, D.C.)*, 7(2), 336–353. <https://doi.org/10.1037/1528-3542.7.2.336>
- Feldman, J. (2019). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms*. Thousand Oaks, California, Corwin, A Sage Company.
- Fisher, B., Allen, R., & Kose, G. (1996). The relationship between anxiety and problem-solving skills in children with and without learning disabilities. *Journal of Learning Disabilities*, 29(4), 439–446. <https://doi.org/10.1177/002221949602900412>
- Freire, P. (2000). *Pedagogy of the oppressed* (30th anniversary edition.). Continuum.

- Goldstein, S. E., Boxer, P., & Rudolph, E. (2015). Middle school transition stress: Links with academic performance, motivation, and school experiences. *Contemporary School Psychology, 19*(1), 21–29. <https://doi.org/10.1007/s40688-014-0044-4>
- Gordon, E. (1995). Myths and realities: African Americans and the measurement of human abilities. *The Journal of Negro Education, 64*, 211–385.
- Gottfried, M. A. (2009). Excused versus unexcused: How student absences in elementary school affect academic achievement. *Educational Evaluation and Policy Analysis, 31*(4), 392–415. <http://www.jstor.org/stable/25621592>
- Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review, 52*, 209–224. <https://doi.org/10.1016/j.econedurev.2016.03.002>
- Green, S. K., Johnson, R. L., Kim, D.-H., & Pope, N. S. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education, 23*(7), 999–1011. <https://doi.org/10.1016/j.tate.2006.04.042>
- Guskey, T. R. (2002). Computerized Gradebooks and the Myth of Objectivity. *Phi Delta Kappan, 83*(10), 775–780. <https://doi.org/10.1177/003172170208301013>
- Guskey, T. R., & Jung, L. A. (2016). Grading: Why you should trust your judgment. *Educational Leadership* (Vol. 73, Issue 7, pp. 50-). ASCD.
- Hammond, Z. (2015). *Culturally responsive teaching and the brain: Promoting authentic engagement and rigor among culturally and linguistically diverse students*. Corwin.
- Hancock, K. J., Gottfried, M. A., & Zubrick, S. R. (2018). Does the reason matter? How student-reported reasons for school absence contribute to differences in achievement

- outcomes among 14-15 year olds. *British Educational Research Journal*, 44(1), 141–174. <https://doi.org/10.1002/berj.3322>
- Hattie, J. (2003). *Teachers make a difference: What is the research evidence?* [Paper presentation]. Building Teacher Quality: What does the research tell us? ACER Research Conference. Melbourne, Australia.
- http://research.acer.edu.au/research_conference_2003/4/
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Holman, A. R., D’Costa, S., & Janowitch, L. (2021). Toward equity in school-based assessment: Incorporating collaborative/therapeutic techniques to redistribute power. *School Psychology Review*, ahead-of-print, 1–14.
- <https://doi.org/10.1080/2372966X.2021.1997060>
- Hunn, L. M. (2004). Africentric philosophy: A remedy for Eurocentric dominance. *New Directions for Adult and Continuing Education*, 2004(102), 65–74. <https://doi.org/10.1002/ace.139>
- James, C.E. & Turner, T. (2017). *Towards race equity in education: The schooling of black students in the Greater Toronto Area*. York University.
- Kelly, S. (2008). What Types of Students’ Effort Are Rewarded with High Marks? *Sociology of Education*, 81(1), 32–52. <https://doi.org/10.1177/003804070808100102>
- Kerr, B. G., & Averill, R. M. (2021). Contextualizing assessment within Aotearoa New Zealand: Drawing from Mātauranga Māori. *AlterNative: An International Journal of Indigenous Peoples*, 17(2), 236–245. <https://doi.org/10.1177/11771801211016450>

- Kurdi, V., & Archambault, I. (2020). Self-perceptions and engagement in low-socio-economic elementary school students: The moderating effects of immigration status and anxiety. *School Mental Health, 12*(2), 400–416. <https://doi.org/10.1007/s12310-020-09360-3>
- Lyiscott, J. (2019). *Black Appetite. White Food*. Routledge.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement, Issues and Practice, 20*(1), 20–32. <https://doi.org/10.1111/j.1745-3992.2001.tb00055.x>
- Merriam-Webster. (n.d.). Eurocentric. In *Merriam-Webster.com dictionary*. Retrieved August 20, 2023, from <https://www.merriam-webster.com/dictionary/Eurocentric>
- Moran, T. P. (2016). Anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological Bulletin, 142*(8), 831–864. <https://doi.org/10.1037/bul0000051>
- Morrissey, T. W., Hutchison, L., & Winsler, A. (2014). Family income, school attendance, and academic achievement in elementary school. *Developmental Psychology, 50*(3), 741–753. <https://doi.org/10.1037/a0033848>
- Moss, P.A., Pullin, D.C., Gee, J.P., Haertel, E.H., Young, L.J. (2008). *Assessment, Equity, and Opportunity to Learn*. Cambridge University Press.
- O'Connor, K. (2011). *A repair kit for grading: 15 Fixes for broken grades*. Pearson.
- Ogbu, J. U., & Simons, H. D. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology & Education Quarterly, 29*(2), 155–188. <https://doi.org/10.1525/aeq.1998.29.2.155>

- Ogbu, J. U. (2003). *Black American students in an affluent suburb: A study of academic disengagement*. L. Erlbaum Associates. Ontario College of Teachers. (n.d.). *Ethical Standards*. <https://www.oct.ca/public/professional-standards/ethical-standards>
- Ontario Ministry of Education. (2010) *Growing Success. Assessment, evaluation and reporting in Ontario schools. First edition, covering grades 1 to 12.*
- Ontario Ministry of Education. (2013). *Learning for all: A guide to effective assessment and instruction for all Students, kindergarten to grade 12.*
- Ontario Ministry of Education. (2014). *Equity and inclusive education in Ontario schools. Guidelines for policy development and implementation.*
<https://files.ontario.ca/edu-equity-inclusive-education-guidelines-policy-2014-en-2022-01-13.pdf>
- Peter, T., Campbell, C.P., & Taylor, C. (2021). Still in every class in every school: Final report on the second climate survey on homophobia, biphobia, and transphobia in Canadian schools. Toronto, ON: Egale Canada Human Rights Trust.
- Quraishi, U., Munir, H., Afzal, A., Saeed, M. (2021). A Study on Teachers' Favoritism and its Effects on Students' Academic Learning at University Level in Multan District. *Psychology and Education*, 58(3), 2304–2311.
<https://psychologyandeducation.net/pae/index.php/pae/article/view/4243>
- Reeves, D. B. (2000). Standards Are Not Enough: Essential Transformations for School Success. *NASSP Bulletin*, 84(620), 5–19.
<https://doi.org/10.1177/019263650008462002>
- Reeves, D. B. (2004). The Case against the Zero. *Phi Delta Kappan*, 86(4), 324–325.
<https://doi.org/10.1177/003172170408600418>

- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The Influence of Affective Teacher-Student Relationships on Students' School Engagement and Achievement: A Meta-Analytic Approach. *Review of Educational Research*, 81(4), 493–529.
<http://www.jstor.org/stable/41408670>
- Schuwirth, L., & Ash, J. (2013). Assessing tomorrow's learners: In competency-based education only a radically different holistic method of assessment will work. Six things we could forget. *Medical Teacher*, 35(7), 555–559.
<https://doi.org/10.3109/0142159X.2013.787140>
- Shepard, L. A. (2019). Classroom assessment to support teaching and learning. *The Annals of the American Academy of Political and Social Science*, 683(1), 183–200.
<https://doi.org/10.1177/0002716219843818>
- Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement, Issues and Practice*, 37(1), 21–34.
<https://doi.org/10.1111/emip.12189>
- Solano-Flores, G. (2023). How serious are we about fairness in testing and how far are we willing to go? A response to Randall and Bennett with reflections about the standards for educational and psychological testing. *Educational Assessment*, 28(2), 105–117.
<https://doi.org/10.1080/10627197.2023.2226388>
- Starch, D., & Elliott, E. C. (1912). Reliability of the Grading of High-School Work in English. *School Review*, 20(7), 442–457. <https://doi.org/10.1086/435971>
- Sundberg, J. (2009). Eurocentrism. *International encyclopedia of human geography*. (pp. 638-643). Elsevier, Oxford. <https://doi.org/10.1016/B978-008044910-4.00093-6>

- Terrin, É., & Triventi, M. (2023). The effect of school tracking on student achievement and inequality: A meta-analysis. *Review of Educational Research*, 93(2), 236–274.
<https://doi.org/10.3102/00346543221100850>
- Tramonte, L., & Willms, D. (2010). The prevalence of anxiety among middle and secondary school students in Canada. *Canadian Journal of Public Health / Revue Canadienne de Sante'e Publique*, 101, S19–S22. <http://www.jstor.org/stable/41995369>
- Tyler, R. W. (1986). Reflecting on the Eight-Year Study. *Journal of Thought*, 21(1), 15–23.
<http://www.jstor.org/stable/42589134>
- Voerman, L., Meijer, P. C., Korthagen, F. A. J., & Simons, R. J. (2012). Types and frequencies of feedback interventions in classroom interaction in secondary education. *Teaching and Teacher Education*, 28(8), 1107–1115. <https://doi.org/10.1016/j.tate.2012.06.006>
- What Did the Eight-Year Study Reveal? (1942). *The High School Journal*, 25(7), 305–309.
<http://www.jstor.org/stable/40367588>

Appendix A

Questionnaire, Themes, Codes, References to Literature and Consent

You are invited to participate in a research project entitled “Deepening our Understanding of How Secondary Teachers Perceive Fairness in Assessment Practices”. This research is being conducted under the supervision of Dr. Christina van Barneveld, Associate Professor in the Faculty of Education at Lakehead University as part of the requirements for a thesis in my work towards a Masters Degree in Social Justice in Education.

Taking part in this study is voluntary. Before you decide whether or not you would like to take part in this study, please read this letter carefully to understand what is involved. After you have read the letter, please email any questions you may have.

PURPOSE

This research will explore teacher perceptions about factors that impact fairness in assessment. Assessment plays a big part in the teaching and learning process. This study could provide a deeper understanding of teachers’ perceptions of fairness in the area of assessment, thus illuminating ways in which teachers impact fairness.

WHAT INFORMATION WILL BE COLLECTED?

An online questionnaire will collect data from you on your perceptions about assessment practices. You will be asked to respond to hypothetical scenarios to indicate what you think most teachers would do, not what you actually would do or have done.

You will have the option to enter your email address if you are interested in being entered in the draw for one of three \$50 Visa gift cards. This information will be collected using a separate link so there is no way of attaching your identity to your questionnaire responses.

WHAT IS REQUESTED OF ME AS A PARTICIPANT?

You will be asked to complete an anonymous online questionnaire. It will take about 25 minutes to complete multiple choice questions with an opportunity to add comments, if you wish.

Upon completion of the questionnaire, you are invited to volunteer for a 30 minute semi-structured interview. If you wish to participate in the interview, click on the link at the end of the questionnaire and provide your email address in the space provided. The researcher will contact about six participants who express interest, in order to set up a convenient time for the interview.

WHAT ARE MY RIGHTS AS A PARTICIPANT?

Please be assured that:

- You are under no obligation to participate and are free to withdraw at any time without prejudice to pre-existing entitlements.
- Your decision to participate will not affect your employment.
- You will be given, in a timely manner throughout the course of the research project, information that is relevant to your decision to continue or withdraw from participation.
- You have the right to request the withdrawal of your data from the study. However, the feasibility of the data withdrawal after completion of the online questionnaire will be limited, due to the anonymous nature of the questionnaire.

WHAT ARE THE RISKS AND BENEFITS?

- There are no foreseeable harms or risks for research participants.
- There is no deception as part of this research.
- There are no direct benefits of participation for those that volunteer. However, entering an email address will include you in a raffle to win one of three \$50 Visa gift cards. If you are randomly selected to win one of the cards, the researcher will email you to arrange for delivery of the gift card.
- Each participant who is interviewed will receive a \$50 Visa gift card after the interview.
- There may be benefits to faculty and future students in that the results of this research may inform dialogue and professional development around the pedagogical and ethical role of fair assessment practices.

HOW WILL MY CONFIDENTIALITY BE MAINTAINED?

The online questionnaire is anonymous for all participants who choose not to be interviewed. The information you provide, will be published in aggregate form only and your identity will be protected. No demographic data will be collected in the questionnaire, so there will be no way of identifying an individual person even through deduction. If you do provide an email address, it will only be used to contact you regarding receipt of a gift card or to set up an interview. Your email will be known to the researcher only, and will be kept confidential.

Please note that the online survey tool used in the study, SurveyMonkey, is hosted by a server located in the USA. The US Patriot Act permits U.S. law enforcement officials, for the purpose of anti-terrorism investigation, to seek a court order that allows access to the personal records of any person without the person's knowledge. In view of this we cannot absolutely guarantee the full confidentiality and anonymity of your data. With your consent to participate in this study, you acknowledge this.

WHAT WILL MY DATA BE USED FOR?

Data will only be seen by myself, Samantha Cullen and the supervisor on this project, Dr. Christina van Barneveld. The data will only be used for the research purposes stated above. Anonymity will be protected in the publication of the thesis and the dissertation as results will only be published in aggregate form. If the results appear in future publications or academic conferences, the same terms of confidentiality will apply.

WHERE WILL MY DATA BE STORED?

Survey information will be collected on SurveyMonkey. The digital data collected will be exported to spreadsheets and stored on the researcher's password-protected computer that is stored in a locked environment for 7 years before being permanently deleted.

HOW CAN I RECEIVE A COPY OF THE RESEARCH RESULTS?

The completed thesis will be available to anyone. You may obtain a copy of the research results by emailing the researcher, Samantha Cullen at slcullen@lakeheadu.ca.

WHAT IF I WANT TO WITHDRAW FROM THE STUDY?

If you wish to withdraw from the research, please contact the researcher, Samantha Cullen via email at slcullen@lakeheadu.ca. Please note that previously collected data on anonymous online questionnaires cannot be withdrawn.

RESEARCHER CONTACT INFORMATION:

Researcher: Dr. Christina van Barneveld, PhD. Email: cvanbarn@lakeheadu.ca

Student Researcher: Samantha Cullen, BSc. BEd. Email: slcullen@lakeheadu.ca

NOTE: There is no possibility of commercialization of research findings, and there are no real, potential or perceived conflicts of interest on the part of the researchers or their institutions.

RESEARCH ETHICS BOARD REVIEW AND APPROVAL:

This research study has been reviewed and approved by the Lakehead University Research Ethics Board. If you have any questions related to the ethics of the research and would like to speak to someone outside the research team, please contact Sue Wright at the Research Ethics Board at 807-343-8010 ext. 8283 or research@lakeheadu.ca.

If you agree to the terms in the information letter above and consent to participate in this research, PLEASE CLICK "NEXT" (below) to move to the survey questions.

Hypothetical Scenarios

For each of the hypothetical scenarios below, please select the option that most closely resembles **what you think “MOST TEACHERS” would do**, not specifically what you would do or have done.

There is a comment section for each scenario if you want to comment or explain your answer.

1. A student misses completing 2 assessments due to illness. When they return, they complete the two assessments. Most teachers would... ● Assess the two assignments the same as students who completed the assignments on time. ● Assess the assignments and remove marks for being late. ● Assign zeros for the missed assignments. Most teachers do not accept late assignments. ● Leave it blank in the grade book. ● Give the student more difficult assessments to complete when they return. ● Other (please explain below) Please comment (optional)	
Theme/Code Opportunity to Learn - Student Engagement Construct-irrelevant Variance - Mathematical processes - Variance in student scores Teacher Worldview - Professional Judgment - Relationships with students	Reference to Literature Gottfried (2009) Reeves (2004) Hancock et al. (2018)

2. A student misses completing 2 assessments due to hockey. When they return, they complete the two assessments. Most teachers would... ● Assess the two assignments the same as students who completed the assignments on time. ● Assess the assignments but remove marks for being late. ● Assign zeros for the missed assignments. Most teachers do not accept late assignments. ● Leave it blank in the grade book.

• Give the student more difficult assessments to complete when they return. • Other (please explain below) Please comment (optional)	
Theme/Code Opportunity to Learn - Student Engagement Construct-irrelevant Variance - Mathematical processes - Variance in student scores Teacher Worldview - Professional judgment - Relationships with students	Reference to Literature Gottfried (2009) Hancock et al. (2018)

3. An Indigenous student misses completing 2 assessments while hunting. When they return, they complete the two assessments. Most teachers would... • Assess the two assignments the same as students who completed the assignments on time. • Assess the assignments but remove marks for being late. • Assign zeros for the missed assignments. Most teachers do not accept late assignments. • Leave it blank in the grade book. • Give the student more difficult assessments to complete when they return. • Other (please explain below) Please comment (optional)	
Theme/Code Opportunity to Learn - Student Engagement - Marginalized groups Construct-irrelevant Variance - Mathematical processes - Variance in student scores Teacher Worldview - Professional judgment - Relationships with students	Reference to Literature Ogbu (1998, 2003) Hancock et al. (2018)

4. While marking a test, it becomes apparent to the teacher that a student has not learned some of the material. The teacher recalls that the **student missed a couple of lessons during the unit**.

What would most teachers do?

- Leave it blank in the gradebook.
- Grade the student's test the same as other students.
- Omit some test items from the grade calculation for this one student's test.
- Omit some test items for all students who did the test.
- Allow this one student to rewrite the test.
- Allow all students to rewrite the test.
- Other

Please comment (optional)

Theme/Code

Opportunity to Learn

- Student engagement (absences)

Construct Irrelevant Variance

- Variance from Teacher to Teacher
- Variance in Student Scores

Teacher Worldview

- Professional judgment
- Relationships with students

Reference to Literature

Gottfried (2019)

Brookhart et al. (2016)

Ontario Ministry of Education (2010)

5. While marking a test, it becomes apparent to the teacher that a student has not learned some of the material. The teacher recalls that they **rushed teaching** that material because they were short on time.

What would most teachers do?

- Leave it blank in the gradebook.
- Grade the student's test the same as other students.
- Omit some test items from the grade calculation for this one student's test.
- Omit some test items for all students who did the test.
- Allow this one student to rewrite the test.
- Allow all students to rewrite the test.
- Other

Please comment (optional)	
Theme/Code Opportunity to Learn - Teacher Practice Construct-irrelevant Variance - Variance from Teacher to Teacher Teacher Worldview - Power Dynamics - Professional Judgment	Reference to Literature Moss et al. (2008) Hattie (2003)

6. On a formative assessment, a teacher in your department offers students specific feedback about how to improve an assignment without putting a grade on it. **The students have an opportunity to use the feedback to improve their work before turning it in for a grade.**

Indicate how most teachers would feel about this practice.

- This practice is fair
- This practice is unfair
- This practice is unrelated to fairness
- Other

Comment (optional)

Theme/Code Opportunity to Learn - Teacher Practice - Assessment Practices (feedback)	Reference to Literature Black et al. (2004) Shepard et al. (2018) Hattie (2007)
---	---

7. On a formative assessment, a teacher in your department offers students **a grade (a number or letter) on assignments with no written or oral feedback.**

Indicate how most teachers would feel about this practice.

- This practice is fair
- This practice is unfair
- This practice is unrelated to fairness
- Other

Comment (optional)	
Theme/Code Opportunity to Learn - Teacher Practice - Assessment Practices (feedback)	Reference to Literature Black et al. (2010) Butler and Nisan (1986)

8. On a formative assessment, a teacher in your department offers students specific feedback that students can only use on future assignments (i.e., they cannot redo this particular assignment.) Indicate how most teachers would feel about this practice. • This practice is fair • This practice is unfair • This practice is unrelated to fairness • Other Comment (optional)	
Theme/Code - Teacher Practice - Assessment Practices (feedback)	Reference to Literature Black et al. (2010) Voerman et al. (2012)

9. For a given curriculum expectation, a student has the following four scores (in chronological order): 0%, 60%, 60%, 80%. What would most teachers say is the best numerical grade representation of this student's knowledge/performance for this curriculum expectation? • 50% (the average score) • 80% (the most recent score) • 60% (the mode and median score) • Other Comment (optional)	
Theme/Code Construct Irrelevant Variance - Mathematical Processes	Reference to Literature Reeves (2004) Ontario Ministry of Education (2010)

- Variance from Teacher to Teacher Teacher Worldview - Professional Judgment	Carifio and Carey (2015) Guskey and Jung (2016)
--	--

10. For a given curriculum expectation, a student has the following four scores (in chronological order): **82%, 50%, 82%, 85%**

What would most teachers say is the best numerical grade representation of this student's knowledge/performance for this curriculum expectation?

- 74.5% (the average score)
- 85% (the most recent score)
- 82% (the mode and median score)
- Other

Comment (optional)

Theme/Code Construct Irrelevant Variance - Mathematical Processes - Variance from Teacher to Teacher Teacher Worldview - Professional Judgment	Reference to Literature Black et al. (2010) Feldman (2019) Guskey and Jung (2016)
---	---

11. For a given curriculum expectation, a student has the following four scores (in chronological order): **62%, 50%, 62%, 85%**.

What would most teachers say is the best numerical grade representation of this student's knowledge/performance for this curriculum expectation?

- 65% (the average score)
- 85% (the most recent score)
- 62% (the mode and median score)
- Other

Comment (optional)

Theme/Code Construct Irrelevant Variance - Mathematical Processes - Variance from Teacher to Teacher Teacher Worldview	Reference to Literature Black et al. (2010) Feldman (2019) Guskey and Jung (2016)
---	---

- Professional Judgment	
-------------------------	--

12. Thomas is a 'dream student'. He participates in class and always puts forth effort. Thomas does very well on classwork and assignments, but poorly on tests.

Dylan is a 'challenging student'. When he comes to class, he rarely participates and when he does, it seems like it is his mission in life to cause trouble. His classwork and assignments show little effort and are done poorly, but he does extremely well on tests.

After applying category weightings for classwork, assignments and tests, both of these students have an average grade of 75%.

What observation would most teachers make regarding their assessment practices?

- Classwork and assignments should be weighted higher, because Thomas' score should be higher than Dylan's.
- Tests should be weighted higher, because Dylan's score should be higher than Thomas'.
- Change nothing. This is how grades work out sometimes.
- Other

Comment (optional)

Theme/Code

Construct Irrelevant Variance
 - Variance from Teacher to Teacher
 Teacher Worldview
 - Professional Judgment
 - Relationships with Students

Reference to Literature

McMillan (2001)
 Brookhart et al. (2016)

13. Consider the students Thomas ('dream student') and Dylan ('challenging student') for this scenario. They both have an average of 78.

Bumping **Thomas'** mark up to an 80 would make him **eligible for acceptance at a post-secondary institution, but if his mark remains 78, he is unlikely to be admitted.**

Most teachers would likely...

- Keep the marks as they are. Teachers don't adjust marks without evidence.

● Increase Thomas's mark up to 80. ● Adjust the weighting of an assignment or two in order to bring Thomas to a legitimate 80. ● Increase all student grades 2% ● Other Comment (optional)	
Theme/Code Construct Irrelevant Variance - Variance from Teacher to Teacher Teacher Worldview - Professional Judgment - Relationships with Students	Reference to Literature Kelly (2008) Ontario Ministry of Education (2010)

14. Consider the students Thomas ('dream student') and Dylan ('challenging student') for this scenario. They both have an average of 78. Bumping Dylan's mark up to an 80 would make him eligible for acceptance at a post-secondary institution, but if his mark remains 78, he is unlikely to be admitted. Most teachers would likely... ● Keep the marks as they are. Teachers don't adjust marks without evidence. ● Increase Dylan's mark up to 80. ● Adjust the weighting of an assignment or two in order to bring Dylan to a legitimate 80. ● Increase all student grades 2% ● Other Comment (optional)	
Theme/Code Construct Irrelevant Variance - Variance from Teacher to Teacher Teacher Worldview - Professional Judgment - Relationships with Students	Reference to Literature Ontario Ministry of Education (2010) Feldman (2019) Kelly (2008)

15. Consider the students Thomas ('dream student') and Dylan ('challenging student') for this scenario. They both have an average of 78.	
---	--

Thomas' **mom is an old friend** and she would really like Thomas to receive an 80.

Most teachers would likely...

- Keep the marks as they are. Teachers don't adjust marks without evidence.
- Increase Thomas' mark up to 80.
- Adjust the weighting of an assignment or two in order to bring Thomas to a legitimate 80.
- Increase all student grades 2%
- Other

Comment (optional)

Theme/Code

Construct Irrelevant Variance

- Variance from Teacher to Teacher

Teacher Worldview

- Professional Judgment
- Relationships with Students

Reference to Literature

Aydogan (2008)

Baniasadi (2023)

Black et al. (2010)

16. Consider the students Thomas ('dream student') and Dylan ('challenging student') for this scenario. They both have an average of 78.

Dylan's **mom is an old friend** and she would really like Dylan to receive an 80.

Most teachers would likely...

- Keep the marks as they are. Teachers don't adjust marks without evidence.
- Increase Dylan's mark up to 80.
- Adjust the weighting of an assignment or two in order to bring Dylan to a legitimate 80.
- Increase all student grades 2%
- Other

Comment (optional)

Theme/Code

Construct Irrelevant Variance

- Variance from Teacher to Teacher

Teacher Worldview

Reference to Literature

Aydogan (2008)

Baniasadi (2023)

Black et al. (2010)

- Professional Judgment - Relationships with Students	
--	--

17. In the staffroom, two teachers are arguing about a student who struggles with test anxiety.
The student does not have an IEP.

Teacher 'A' suggests that an **alternative method (an assignment)** might be more fair than a test for this student.

Teacher 'B' argues that changing the method for one student would be unfair to other students.

Most teachers are more likely to side with...

- Teacher A because an alternative assessment method is considered differentiated support.
- Teacher A because the student's score will be negatively affected by their anxiety.
- Teacher A for a reason other than those listed above (Please comment below)
- Teacher B because this student would have an unfair advantage over other anxious students who are writing the test.
- Teacher B because writing a test in standardized conditions is an important part of fairness.
- Teacher B for a reason other than those listed above (Please comment below)
- Other

Comment (optional)

Theme/Code

Opportunity to Learn

- Teacher Practice

Construct Irrelevant Variance

- Variance from Teacher to Teacher
- Student Anxiety

Reference to Literature

Moss et al. (2008)

Goldstein et al. (2015)

Ontario Ministry of Education (2013)

18. In the staffroom, two teachers are arguing about a student who struggles with test anxiety.
The student does not have an IEP.

Teacher 'A' suggests that an **alternate location** might be more fair for this student.

Teacher 'B' argues that changing the location for one student would be unfair to other students.

Most teachers are more likely to side with... • Teacher A because an alternate location is considered differentiated support. • Teacher A because the student's score will be negatively affected by their anxiety. • Teacher A for a reason other than those listed above (Please comment below) • Teacher B because this student would have an unfair advantage over other anxious students who are writing the test. • Teacher B because writing a test in standardized conditions is an important part of fairness. • Teacher B for a reason other than those listed above (Please comment below) • Other Comment (optional)	
Theme/Code Opportunity to Learn - Teacher Practice Construct Irrelevant Variance - Variance from Teacher to Teacher - Student Anxiety	Reference to Literature Moss et al. (2008) Shepard et al. (2018) Ontario Ministry of Education (2013)

19. Indicate your agreement or disagreement with each of the following statements.

	Agree	Disagree
Most teachers replicate grading systems that they know (i.e. grading systems that are similar to what they experienced in school).	O	O
Theme/Code Opportunity to Learn - Assessment Practices Teacher Worldview - Power Dynamics	Reference to Literature Deneen and Boud (2014) Black et al. (2010)	
Most teachers decide unilaterally how students will demonstrate knowledge in their classroom.	O	O
Theme/Code Construct-irrelevant Variance - Variance from Teacher to Teacher Teacher Worldview - Power Dynamics	Reference to Literature Black et al. (2010) Brookhart et al. (2016)	

Most teachers communicate information about an assessment at a level of detail appropriate to students.	O	O
Theme/Code Opportunity to Learn - Assessment Practices	Reference to Literature Voerman et al. (2012) Hattie (2007)	
Most teachers believe that there are some students who don't have the ability to learn what they are teaching.	O	O
Theme/Code Opportunity to Learn - Marginalized Groups Teacher Worldview - Power Dynamics - Social Context	Reference to Literature Gershenson et al. (2016) Codrington and Fairchild (2012)	
Most teachers are aware of systemic bias and attempt to address it in their classroom.	O	O
Theme/Code Opportunity to Learn - Marginalized Group Teacher Worldview - Social Context	Reference to Literature Kerr and Averill (2021) Shepard et al. (2019)	
Most teachers respond more favorably to students who follow classroom rules.	O	O
Theme/Code Construct-irrelevant Variance - Variance from Teacher to Teacher Teacher Worldview - Professional Judgment	Reference to Literature Kelly (2008) McMillan (2001)	
Most teachers will adjust their methods if some students are not doing well, rather than recommend a course change.	O	O
Theme/Code Opportunity to Learn - Assessment Practices	Reference to Literature Moss et al. (2008) Ontario Ministry of Education (2014)	
Most teachers think that streaming or tracking	O	O

students will impact their learning opportunities negatively.		
Theme/Code Teacher Worldview - Power Dynamics	Reference to Literature James and Turner (2017) Terrin and Triventi (2023)	
Most teachers use grades as a reward for good behaviour.	O	O
Theme/Code Construct-irrelevant Variance - Variance from Teacher to Teacher Teacher Worldview - Professional Judgment	Reference to Literature Kelly (2008) Brookhart et al. (2016)	
Most teachers use grades as a punishment for bad behaviour.	O	O
Theme/Code Construct-irrelevant Variance - Variance from Teacher to Teacher Teacher Worldview - Professional Judgment	Reference to Literature Shepard et al. (2018) Feldman (2019)	
Most teachers connect more easily with students with whom they have things in common.	O	O
Theme/Code Opportunity to Learn - Marginalized Groups Teacher Worldview - Power Dynamics - Relationships with Students	Reference to Literature Holman et al. (2021) Hammond (2015) Solano-Flores (2023)	
Most teachers have good knowledge of fair assessment practices.	O	O
Theme/Code Opportunity to Learn - Teacher Practice	Reference to Literature Moss et al. (2008) Hattie (2003)	
Most teachers assess students fairly.	O	O
Theme/Code Opportunity to Learn	Reference to Literature Black et al. (2010)	

- Teacher Practice Construct-irrelevant Variance - Variance from Teacher to Teacher Teacher Worldview - Power Dynamics - Professional Judgment - Social Context - Relationships with Students	Moss et al. (2008)	
Most teachers don't have time to think about fairness in assessment.	O	O
Theme/Code Opportunity to Learn - Teacher Practice	Reference to Literature Brookhart et al. (2016) Black et al. (2010)	

Comment (optional)

Thank you

Thank you for completing the questionnaire.

Would you like to enter a raffle to win one of three \$50 VISA gift cards? If yes click [HERE](#) and enter your email. The three winners will be randomly selected. If you are randomly selected to win one of the cards, the researcher will email you with the good news and arrange for delivery of the gift card.

CONSENT FOR INTERVIEW

I'd love to hear more from you! **Are you willing to participate in a follow-up interview?** If you are willing to participate in a 30-minute interview please click [HERE](#) and enter your email address. I'll email you to set up a time that is convenient for you. **All volunteers who complete the interview will receive a \$50 gift card.**

NOTE: Both these links go to the same place so you can enter both the raffle and also volunteer for an interview, if you wish.

Appendix B

Themes and Codes Related to Fairness and Assessment

Table B-1

Theme	Codes	Considerations
Opportunity to Learn	student engagement (student absences)	attendance, excused vs unexcused absences
	marginalized groups	ethnicity, socio-economic status, learning disability, accommodations, formative assessment
	teacher practice	knowledge of content and pedagogy, impact on achievement and OTL, difficult pedagogical practice to change, effective use of feedback, formative assessment
Construct-irrelevant Variance	mathematical processes	teacher autonomy (grading systems not monitored), averages, use of zeros
	variance from teacher to teacher	standards and procedures, perceptions (subjectivity), adjusting grades, including non-cognitive factors (academic enablers) in grades
	variance in student scores	measuring circumstance rather than knowledge or ability, intermittent catastrophic failures or early failures, reliance on averages rather than judgment (most recent, most consistent)
	student anxiety	difficult to assess cognitive ability (working memory), may contribute to variance in student scores, marginalized students, learning disabilities
Teacher Worldview		

power dynamics (history of education systems and eurocentrism)	marginalized students historically underserved and impacted, systemic bias, power imbalance, demographic mismatch, subtle forms of segregation, student tracking, teachers replicate grading systems they know
teacher professional judgment	fair/valid grades require interpretation, bias (discrimination/favouritism, punishment/reward), differing practices based on beliefs about purpose of assessment (use of zeros)
social context	classroom culture/environment, barriers and opportunities, student voice, contextualizing learning, bias, lack of representation for marginalized groups
relationships with students	acknowledge social/cultural pressures, favouritism/discrimination (worldviews differ, nepotism, attitudes or habits), classroom/school discipline, positive relationships (not established with all) improve outcomes

Appendix C

Purpose of the Interviews

These interviews will deepen our understanding of how secondary teachers perceive fairness in assessment practices.

Estimated Length

30 – 45 minutes

Opening Script

Thank you so much for taking the time to meet with me. I would like to confirm that you have read the information/consent letter. Do you have any questions about this research project? I'd like to remind you that this interview will be recorded for transcription purposes. Your confidentiality will be protected. I won't discuss your particular interview with anyone other than my supervisor. You have the right to 'pass' (not answer) any questions posed during the interview. Would it be ok if your gender or years of experience are mentioned in the results? A pseudonym will be used for any quotes that are published in the results.

Interview Questions

1. Participants were shown a bar graph of the results of question 7 (about formative assessment) from the questionnaire and were asked "What do you think is happening?"
2. Participants were shown a bar graph of the results for question 9 (where individual grades were 0, 60, 60 and 80) from the questionnaire and were asked "Do these results surprise you?" Some follow up probe questions included: "What do you think teachers are thinking?" and "Do you think teachers track grades on paper and look for patterns (vs entering marks in the system)?"

3. “How do you think teachers handle zero grades?” A follow up probe question included “Do you think most people are on the same page?”
4. “What factors do you think teachers should consider when determining whether assessment is fair or not?”
5. “What challenges do you think most teachers face when trying to ensure fairness in assessment?” A follow up probe question included “Are there any strategies you’ve found effective in promoting fair assessment practices with your own students?”
6. “Can you describe a specific assessment task, practice or strategy you belief exemplifies fairness?” A follow up probe question included “How do you think teachers could ensure that assessments are fair within the context of their own school?”
7. Participants were invited to share any final thoughts they might have regarding fairness in assessment. A follow up probe question included “What changes or support would enhance your ability to enact fair assessment practices?”

Closing Script

Thank you for your time and valuable insights. Would you like to check the transcription or themes taken from the transcription? Would you like to receive a copy of the results when they are available? You can contact me by email if you have any questions or need clarification. Confidentiality will be protected. Nothing from the transcript that could potentially be used to identify will be published in the results. Thank you again for your time.

The following letter was provided to participants at the end of the online questionnaire. It details all information that was provided to participants about the interview protocol and consent to participate.

Information Letter and Consent Form to Participate in An Interview

Do you want to participate in a 30-minute interview? ALL interview participants will receive a \$50 gift card.

Read the information below and **ENTER YOUR EMAIL AT THE BOTTOM OF THE PAGE**. If you consent, you will be contacted by the researcher to set up a time that is convenient to you.

Dear Potential Interview Participant:

You are invited to participate in a research project entitled “Deepening our understanding of how secondary teachers perceive fairness in assessment practices”. This research is being conducted by Samantha Cullen, M.Ed. candidate, under the supervision of Dr. Christina van Barneveld, Associate Professor in the Faculty of Education at Lakehead University. Taking part in this study is voluntary. Before you decide whether you would like to take part in this study, please read this letter carefully to understand what is involved. After you have read the letter, please ask any questions you may have.

PURPOSE

The purpose of this research is to explore teacher perceptions about factors that impact fairness in assessment. Assessment plays a big part in the teaching and learning process. This study could provide a deeper understanding of teachers perceptions of fairness in the area of assessment, thus illuminating ways in which teachers impact fairness.

WHAT INFORMATION WILL BE COLLECTED?

If you opt to be interviewed, you will be asked to input your email address so the researcher may contact you to set up an interview at a time convenient to you. Interview questions will focus on themes in the research literature, e.g., gauging teacher perceptions about fairness in assessment, opinions on the interpretations of aggregated results from the online questionnaire, and asking about challenges in trying to ensure fair assessment of students. The interview will be recorded and transcribed by Zoom.

WHAT IS REQUESTED OF ME AS A PARTICIPANT?

We ask you to volunteer for a 30-minute semi-structured personal interview with the researcher via Zoom. If you opt to be interviewed, you will be asked to share your email address so the researcher may contact you to set up an interview at a time convenient to you.

WHAT ARE MY RIGHTS AS A PARTICIPANT?

- You are under no obligation to participate, are free to withdraw at any time without prejudice to pre-existing entitlements.
- Your decision to participate will not affect your academic status/employment.
- You will be given, in a timely manner throughout the course of the research project, information that is relevant to your decision to continue or withdraw from participation.
- You have the right to request the withdrawal of your data from the study.

WHAT ARE THE RISKS AND BENEFITS?

- There are no foreseeable risks to interview participants.
- There is no deception as part of this research.
- There are direct benefits of participation for those that volunteer. Every interviewee will receive a \$50 VISA gift card once the interview is completed.
- There may be a benefit to faculties of education and future students, in that results of this research will inform dialogue and professional development around the pedagogical and ethical role of fair assessment practices.

HOW WILL MY CONFIDENTIALITY BE MAINTAINED?

The interview is confidential but not anonymous to the researcher since she will be interviewing you. Only the researcher and her supervisor will see your interview data. Your privacy will be maintained in all reports by using a pseudonym of your choice and you will not be identifiable directly or by deduction.

WHAT WILL MY DATA BE USED FOR?

Your data will be analyzed, and a report of aggregated results will be included in the MEd thesis. A report of results will also be prepared for scholarly publication. Results may be shared at an academic and/or professional conference. Only the

researcher and the supervisor will have access to the data. Anonymity will be protected in all research reports (written, oral, or otherwise).

WHERE WILL MY DATA BE STORED?

Data will be stored for a minimum of 7 years following completion of the project in the principal researchers secured office on her desktop computer. Interview data will be downloaded and deleted from the Zoom application, then stored securely at Lakehead University. The data files and computer each will be password-protected.

HOW CAN I RECEIVE A COPY OF THE RESEARCH RESULTS?

In research reports to the Faculty of Education, scholarly publications, and conference presentations, participants will be not identified directly or indirectly. You may obtain a copy of the research results by emailing the researcher, Samantha Cullen at slcullen@lakeheadu.ca.

WHAT IF I WANT TO WITHDRAW FROM THE STUDY?

If you wish to withdraw from the interview portion of the research, please contact the researcher, Samantha Cullen at slcullen@lakeheadu.ca.

RESEARCHER CONTACT INFORMATION

The student researcher is Samantha Cullen and she may be reached at slcullen@lakeheadu.ca

The principal researcher, Dr. Christina van Barneveld, may be contacted via email at cvanbarn@lakehead.ca

NOTE: There is no possibility of commercialization of research findings, and there are no real, potential or perceived conflicts of interest on the part of the researchers, their institutions, or the research sponsor(s).

RESEARCH ETHICS BOARD REVIEW AND APPROVAL

This research study has been reviewed and approved by the Lakehead University Research Ethics Board. If you have any questions related to the ethics of the research and would like to speak to someone outside of the research team, please contact Sue Wright at the Research Ethics Board at 807-343-8010 ext. 8283 or research@lakeheadu.ca.

MY CONSENT

I agree to the following:

- I have read and understand the information contained in the Information Letter
- I agree to participate

- I understand the risks and benefits to the study
- That I am a volunteer and can withdraw from the study at any time (except for anonymous online questionnaires), and may choose not to answer any question
- That the data will be securely stored for a minimum period of 7 years following completion of the research project
- I understand that the research findings will be made available to me upon request
- I will remain anonymous [unless I volunteer to share my email address]
- All of my questions have been answered
- I allow video recording and transcription of the interview on ZOOM
- By consenting to participate, I have not waived any rights to legal recourse in the event of research-related harm.

If you agree to the terms of the information above and consent to participate in the interview, please enter your email so we can contact you to set up a time that is convenient for you.

Please note that ALL interview participants will receive a \$50 VISA gift card.