FAIRNESS IN CLASSROOM ASSESSMENT: CONCEPTUAL AND EMPIRICAL INVESTIGATIONS

by

Amirhossein Rasooli

A thesis submitted to the Faculty of Education
In conformity with the requirements for
the Degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
August 2021

Copyright ©Amirhossein Rasooli, 2021
Abstract

This study aimed to investigate fairness in classroom assessment conceptually and empirically to contribute to the emerging research in this area. Conceptually, it critically reviewed the current conceptions of fairness in educational and classroom assessment and enriched these conceptions using a broad review of the key conceptions of fairness and justice in philosophy, sociology, psychology, economics, and education.

Empirically, it conducted a two-phase mixed methods study to investigate first year undergraduate students’ perceptions of fairness in classroom assessment in their secondary schools in Ontario, Canada. Phase I included 27 virtual interviews with students that ranged in length from 20-45 minutes and focused on students’ perceptions of un/fairness in classroom assessment. Drawing on the thematic analysis, eight themes were found: (1) overall perceptions of fairness; (2) fairness in groupwork; (3) fairness in exams; (4) fairness in cheating; (5) fairness in grading; (6) fairness in feedback; (7) socio-emotional environment; and (8) responses to perceptions of un/fairness. Overall, these themes coupled with social psychology theory provided the initial foundation to develop the underpinning construct for fairness in classroom assessment.

In Phase II, this underpinning construct were used to develop the Classroom Assessment Fairness Inventory. This inventory included demographic questions as well as five scenarios (i.e., groupwork, exam, grading, cheating, and feedback). The logical validity evidence for this inventory was investigated by collecting reviews from 10 international assessment expert panel and 10 graduate students with previous K-12 teaching experience. The empirical validity evidence (i.e., internal structure, and relationship to other variables) was collected by virtually administering the inventory to 217 participants and analyzing the data using factor analyses and multivariate regression analysis. The factor analyses supported a five-factor model including (a) unfairness in groupwork, (b) fairness in cheating, (c) fairness in grading, (d) unfairness in feedback, and (e) fairness in feedback. The multivariate regression analysis
showed evidence linking students’ personal belief in a just world with perceptions of fairness in grading ($\beta = .262, p < .05$), fairness in feedback ($\beta = .294, p < .05$), and unfairness in feedback ($\beta = -.19, p < .05$).
Acknowledgements

A dark midnight and the terror of wave and a whirlpool so fearful/How can they know of our state of being the light-burdened of shores (Hafiz)

This dissertation is dedicated to all the kids out there in the world who have been born and grown at a margin in disadvantaged conditions. This dedication was a very difficult decision to make given that I had questions how the dedication of this dissertation to these kids can be of any real antidote to the pains and difficulties in their life while also acknowledging that I am the only one who will reap any fruits out of this work. I finally overcame this internal conversation and decided to do this because I could not associate with anyone else in this world as much as these kids. As a result of oppressive social, cultural, economic, and family structures, many of these kids may have internalized their disadvantages and may have continually attributed their failures to their life choices and themselves, further extinguishing the liveliness in their souls. These kids are unaware that, in fact, the social, cultural, economic, and family structures have created this disadvantage and it is not their freewill that has been in operation to select a disadvantaged life. I know that these structures would not allow you even to hear these words, but it is my dim hope that perhaps these words may touch your ears and could contribute to liberating you by making you aware of the huge influence of these hidden structures. A liberation path that took me many years to appreciate but I could also see the power that lies on the other side of this liberation. The power to turn the lessons learned and the scars left through pains from a life at a margin and disadvantage into a window to understand fellow human beings. Hope this dissertation is a pebble toward realizing this liberation for us all.

In 2017, I moved 10,000 kilometers to Canada to start the PhD journey. This journey was my first wander into a big unknown world. This journey was not an easy one for a contemporary student with an Iranian citizenship. Due to political issues, Canada did not have an embassy in Iran, making Iranian students move to neighboring countries to complete Visa process. After waiting for a next step in Visa
application for 8 weeks, I was finally invited to a Canadian VAC in Turkey for a five-minute fingerprint process. I waited in a line outside the Canadian VAC from 5:30am-2:30pm in a sultry day alongside 80 other spouses, mothers, fathers, and elderly just to give a five-minute fingerprint. Due to international sanctions, I was deprived of having international bank Visa or Master cards that I could use to reserve an accommodation in Kingston from Iran. Persuading the Airbnb owner to pay by cash upon arrival, I arrived at Kingston after 25 hours of travel to an unknown world where I just know Dr. Chris DeLuca. Chris picked me up in Kingston, drove me to a sandwich place to eat, and gave a ride to the Airbnb. Knock, Knock, the owner was not at the house. I did not have a Canadian sim card. Chris stayed with me until 12am so that we finally found the owner and he came to allow me in. Thank you, Chris, for coming that night—a memory that will always be with me.

Tomorrow morning, I started officially my PhD at the Faculty of Education. Not only had I difficulties with learning the academic culture but also the social culture out there. There were a lot of challenges (a lot!) from simple things as how to find my cheese among many types of cheese, how to survive (and thrive) a TA course, how to pass PhD courses, how to successfully navigate new team work environment and accomplish my RA roles, how to learn bits and pieces of this new culture, and how to find myself, my identity, and my future success among a group of domestic students that talk English fast (I couldn’t understand), talk about SSHRC (in my head what is SSHRC?), and many other topics that I was blank. Thank you, Chris, that you were very helpful throughout this process and tried to support me in any way you could know. I appreciate many friends like Gwan Im, Alice Johnston, Jackson Pind, Andrew Coombs, Barbara Laing, Kristin Kinnard, Matt Drabenstott, Tahereh Firoozi, Lizzie Yan, and Stephen MacGregor who helped me during this familiarization process to appreciate this new culture and work toward an evolving identity.

Second year, I had a better feeling as I did not have many of the previous cultural challenges. However, the challenges of the past overshadowed the current life. The challenges that I was not even
aware of due to the structures I lived with before. Thanks to Chris who also directed me to the resources that contributed to supporting my well-being.

Third year, up to now I have worked hard. Of course, I am biased. I knew that I could not go to conferences in the US (because of Iranian citizens being included in the Muslim ban by Mr. Trump administration) and could not afford to go to conferences in Europe. Being used to limitations in life, I did not even have the energy to complain! Accepting the limitations, I continued working with the limited opportunities that could be offered to an international student and building new paths. Thanks, Chris, for supporting me to do an internship in Australia using Globalink MITACS Program. Brisbane was very beautiful, and I worked with beautiful people including Drs. Claire Wyatt-Smith, Joy Cumming, and Lenore Adie that taught me other ways of being an academic. During this year, I also worked in different committees in our faculty and learned more about ethical research in EREB committee (Thank you Drs. Pamela Beach and Ian Matheson for constructive environment of this group) as well as other committees, where I worked with other scholars at our faculty that implicitly and explicitly taught me about academics, liberty, and politics embrace. Thank you, Dr. Ted Christou, for our honest conversations about philosophy, life, and history. I also need to appreciate Erin Rennie and Natalie Lefebvre for their administrative support. The third year was a turning point.

Fourth year, the most wonderful year! While Covid-19 happened, and I was unaware of how to collect data in a new culture, I started this research with a lot of hesitations and anxiety. Within limitations, creativity emerges, and I am so much appreciative of the young Canadian students who supported this research with their kind participation and sharing their experiences. In the qualitative part, I tried to talk less and let your voice project. I also appreciate the local and international expert panel scholars who kindly supported this research. Thank you to all of you. While Covid-19 was a difficult year for many of humans, I wondered it provided me with great time for reflection and reading. It is also time to write about two influential scholars who are also members in my dissertation committee. I am writing about them in the fourth year because I appreciated them better in this year. Dr. Liying Cheng and I
worked on projects for the last four years. I learned a lot from Liying about life and academia. Liying is really a courageous scholar who has been able to flourish in an environment that in my perspective is less accepting of other ways of knowing and being. She has provided tremendous support for international and domestic students and has wonderfully carried out her academic work and social responsibilities. Thank you Liying for teaching me implicitly by doing what you are doing. Dr. Amin Mousavi has been a blessing in my fourth year. Amin at least shares some similarities in background with me. Knowing a successful role-model who you can associate with is very significant and empowering for a student—an epiphany that I have realized this year. I am very glad that you have become a part of my committee and learned a lot from you during this short time. Hope, we all continue working together and creating supportive system for future scholars to grow in a safe environment that appreciates other ways of being, thinking, and behaving.

Last but not least, I need to appreciate Drs. Karim Mojtahedi, Abdolah Anvar, Gholam Hossein Ebrahimi Dinani, and Dariush Shayegan. I have never met these philosophers in person but have studied their books. Karim Mojtahedi taught me that an individual without roots would not have a future. A country without being aware of its history would not also have a future. While these statements are in the Hegelian spirit, its archeological approach to dig the past even on an individual level like what Freud did has been a source of awareness and liberation for me to develop my identity. In this research, I also dug into students’ past lived experiences of fairness in assessment to contribute to building the fairer future experiences in assessment. Thank you especially to Dr. Mojtahedi who also taught me what it means to be a student—a Socratic conviction that I know that I do not know that directs the student to search deeply and stay a student forever. The purpose that one can live for.
# Table of Contents

Abstract..............................................................................................................................................ii
Acknowledgements.................................................................................................................................iv
List of Figures .........................................................................................................................................xi
List of Tables ..........................................................................................................................................xii
Chapter 1 Introduction ...........................................................................................................................1
  1.1 Purpose and Research Questions .................................................................................................7
  1.2 Significance .....................................................................................................................................8
  1.3 Definition of Key Terms ..............................................................................................................12
  1.4 Delimitations of the Study ..........................................................................................................14
  1.5 Structure of Dissertation ............................................................................................................15
Chapter 2 Literature Review ..................................................................................................................16
  2.1 Fairness in Assessment .................................................................................................................16
    2.1.1 Key Conceptions of Assessment ..........................................................................................16
    2.1.2 Key Conceptions of Fairness in Assessment ......................................................................20
      2.1.2.1 Psychometric Perspective ............................................................................................20
      2.1.2.2 Legal Perspective ..........................................................................................................22
      2.1.2.3 Philosophical Perspectives ............................................................................................23
      2.1.2.4 Social Psychology Perspective .......................................................................................26
  2.2 Justice Theory and Research in Social Sciences and Humanities .................................................28
    2.2.1 Justice in Philosophy ............................................................................................................28
    2.2.2 Justice in Sociology ..............................................................................................................30
    2.2.3 Justice in Psychology ...........................................................................................................32
    2.2.4 Justice in Economics ............................................................................................................36
  2.3 Justice in Education .......................................................................................................................37
    2.3.1 Justice Conceptions in Education Policy Literature ............................................................38
    2.3.2 Socially Just Pedagogies ........................................................................................................41
    2.3.3 Group Perspective to Justice in Education ..........................................................................42
    2.3.4 Empirical Research on Perceived Justice in Education ........................................................42
  2.4 Toward Conceptual Frameworks for Assessment Fairness ............................................................47
    2.4.1 Critical Review of Key Conceptions of Fairness in Assessment ..........................................47
    2.4.2 Theoretical Bases to Reconceptualize Fairness in Classroom Assessment .........................53
  2.5 Summary ........................................................................................................................................58
Chapter 3 Methodology ........................................................................................................... 61
  3.1 Context................................................................................................................................. 61
  3.2 Research Ethics ..................................................................................................................... 63
  3.3 Mixed Methods Research Design ........................................................................................ 63
  3.4 Phase I .................................................................................................................................. 65
    3.4.1 Recruitment .................................................................................................................... 65
    3.4.2 Participants .................................................................................................................... 66
    3.4.3 Data Collection ............................................................................................................. 67
    3.4.4 Data Analysis .............................................................................................................. 69
  3.5 Phase II ................................................................................................................................ 73
    3.5.1 Classroom Assessment Fairness Inventory Development Process .............................. 73
    3.5.2 Theoretical and Measurement Underpinnings for Classroom Assessment Fairness Inventory ........................................................................................................ 74
    3.5.3 Initial Item Development ............................................................................................... 77
    3.5.4 Validity Evidence ............................................................................................................ 85
      3.5.4.1 Evidence of Validity Based on Test Content: Expert-Panel and Pilot Testing Review ........................................................................................................... 86
      3.5.4.2 Evidence of Validity Based on Internal Structure and Relationship to Other Variables .... 89
    3.5.5 Recruitment .................................................................................................................... 89
    3.5.6 Participants ..................................................................................................................... 90
    3.5.7 Data Collection ............................................................................................................. 91
    3.5.8 Data Analysis .............................................................................................................. 92

Chapter 4 Findings .................................................................................................................. 94
  4.1 Findings for Phase I ........................................................................................................... 94
    4.1.1 Theme 1: Overall Perceptions of Fairness ................................................................. 95
    4.1.2 Theme 2: Fairness in Groupwork ................................................................................ 97
    4.1.3 Theme 3: Fairness in Exam ......................................................................................... 101
    4.1.4 Theme 4: Fairness in Cheating .................................................................................. 103
    4.1.5 Theme 5: Fairness in Grading .................................................................................... 106
    4.1.6 Theme 6: Fairness in Feedback .................................................................................. 110
    4.1.7 Theme 7: Socio-emotional Environment ..................................................................... 111
    4.1.8 Theme 8: Responses to Perceptions of Unfairness .................................................... 115
  4.2 Findings for Phase II .......................................................................................................... 119
    4.2.1 Descriptive Statistics for Classroom Assessment Fairness Inventory ......................... 119
    4.2.2 Exploratory Factor Analysis ....................................................................................... 124
    4.2.3 Confirmatory Factor Analysis .................................................................................... 128
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.4 Structural Equation Modeling</td>
<td>131</td>
</tr>
<tr>
<td>4.2.5 Subgroup Analysis</td>
<td>132</td>
</tr>
<tr>
<td>Chapter 5 Discussion</td>
<td>134</td>
</tr>
<tr>
<td>5.1 A Unifying Theory of Fairness in Assessment</td>
<td>138</td>
</tr>
<tr>
<td>5.1.1 Overall Perceptions of Fairness</td>
<td>140</td>
</tr>
<tr>
<td>5.1.2 Fairness in Groupwork</td>
<td>141</td>
</tr>
<tr>
<td>5.1.3 Fairness in Exams</td>
<td>142</td>
</tr>
<tr>
<td>5.1.4 Fairness in Cheating</td>
<td>143</td>
</tr>
<tr>
<td>5.1.5 Fairness in Grading</td>
<td>144</td>
</tr>
<tr>
<td>5.1.6 Fairness in Feedback</td>
<td>146</td>
</tr>
<tr>
<td>5.1.7 Socioemotional Environment</td>
<td>146</td>
</tr>
<tr>
<td>5.1.8 Psychological and Behavioral Responses</td>
<td>149</td>
</tr>
<tr>
<td>5.2 Limitations of This Research</td>
<td>150</td>
</tr>
<tr>
<td>5.3 Implications</td>
<td>153</td>
</tr>
<tr>
<td>5.4 Conclusion</td>
<td>154</td>
</tr>
<tr>
<td>References</td>
<td>156</td>
</tr>
<tr>
<td>Appendix A Ethical Clearance Letter</td>
<td>182</td>
</tr>
<tr>
<td>Appendix B Ethics Amendment Approval Letter</td>
<td>183</td>
</tr>
<tr>
<td>Appendix C Ethics Amendment Approval Letter</td>
<td>184</td>
</tr>
<tr>
<td>Appendix D Phase I Interview Consent Form</td>
<td>185</td>
</tr>
<tr>
<td>Appendix E Phase II Expert Panel Consent Form</td>
<td>187</td>
</tr>
<tr>
<td>Appendix F Phase II Pilot Testing Consent Form</td>
<td>188</td>
</tr>
<tr>
<td>Appendix G Phase II Survey Consent Form</td>
<td>190</td>
</tr>
<tr>
<td>Appendix H Interview Guide</td>
<td>191</td>
</tr>
<tr>
<td>Appendix I Expert Panel and Pilot Testing Instruction</td>
<td>193</td>
</tr>
<tr>
<td>Appendix J Recruitment Scripts for Phase I</td>
<td>206</td>
</tr>
<tr>
<td>Appendix K Recruitment Scripts for Phase II</td>
<td>207</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. The geographic map of the quantitative studies on fairness in classroom-based research. *Note.*

The total is 102, which is more than 96, because a few studies were cross-cultural. ...........................................47

Figure 2. The link between pragmatics epistemology of psychometrics, validity, and fairness ...............50

Figure 3. Schematic overview of research design...............................................................................................65

Figure 4. Scree plot for the Classroom Assessment Fairness Inventory .................................................................125

Figure 5. Standardized SEM results for belief in a just world and classroom assessment fairness ..............133

Figure 6. A framework for students' perceptions of fairness in classroom assessment .................................137
List of Tables

Table 1. Justice dimensions and their principles.................................................................................. 44
Table 2. Overview of phase I student participants’ demography (n=27).............................................. 67
Table 3. Summary of themes, codes, examples, and frequencies .......................................................... 71
Table 4. Descriptions of dimensions and principles of justice and their representations in Classroom Assessment Fairness Inventory .................................................................................. 79
Table 5. Representation of each fairness dimension within and across five scenarios ....................... 80
Table 6. Classroom Assessment Fairness Inventory- scenario 1 blueprint......................................... 81
Table 7. Classroom Assessment Fairness Inventory- scenario 2 blueprint......................................... 82
Table 8. Classroom Assessment Fairness Inventory- scenario 3 blueprint......................................... 83
Table 9. Classroom Assessment Fairness Inventory- scenario 4 blueprint......................................... 84
Table 10. Classroom Assessment Fairness Inventory- scenario 5 blueprint........................................ 85
Table 11. The expert ratings of alignment across justice principles and scenario actions............... 88
Table 12. Overview of student participants’ demography (n=217).................................................... 90
Table 13. A general overview of themes and associated codes ......................................................... 94
Table 14. Descriptive statistics for students’ responses to the first scenario .................................... 121
Table 15. Descriptive statistics for students’ responses to the second scenario............................... 121
Table 16. Descriptive statistics for students’ responses to the third scenario.................................... 122
Table 17. Descriptive statistics for students’ responses to the fourth scenario................................. 122
Table 18. Descriptive statistics for students’ responses to the fifth scenario ..................................... 123
Table 19. Five-factor model using maximum likelihood with direct oblimin rotation.................... 125
Table 20. Correlations across factors of the Classroom Assessment Fairness Inventory ............... 128
Table 21. Goodness of fit and comparative indices for confirmatory factor analysis ..................... 129
Table 22. Standardized factor loadings from confirmatory factor analysis...................................... 130
Table 23. Students’ perceptions of assessment research and fairness conceptualizations and measurement .................................................................................................................................................. 135
Chapter 1

Introduction

“Injustice anywhere is a threat to justice everywhere. We are caught in the inescapable network of mutuality, tied in a single garment of destiny. Whatever affects one directly, affects all indirectly.” Martin Luther King

As a response to the Civil Rights Movement during 1960s and 70s, educational researchers have focused on investigating fairness in educational assessment and measurement (Cole & Zieky, 2001; Randall, 2021). The persistent investigations on fairness have been reinvigorated with the recent 21st century social and educational movements towards equity, diversity, and inclusion for disadvantaged groups (Aitken, 2012; Camilli, 2013; Carvalho, 2013; Cowie, 2015; Dorans & Cook, 2016; Gipps & Stobart, 2009; Herman & Cook, 2019; Kane, 2010; Kunnan, 2018; Lantolf & Poehner, 2013). Paralleled with this conceptual and empirical literature, fairness has progressively been recognized as a foundation for gauging the quality of assessments in Standards for Educational and Psychological Testing (1974, 1985, 1999, 2014) and in Classroom Assessment Standards (Klinger et al., 2015; Rogers, 1996a, 1996b). This dissertation study aims to add to this literature base and contributes to conceptualizing fairness in assessment standards through investigating Canadian undergraduate students’ perceptions of fairness in classroom assessment during their secondary school experiences in Ontario.

Since the mid-1800s, educational and psychological tests were leveraged for various purposes including measuring school achievement, aptitudes, and intelligence as well as selecting individuals for university and vocational purposes (Newton & Shaw, 2014). Tests were epistemologically supported and designed by the psychometric view of assessment (Moss, Pullin, Gee, & Haertel, 2005). In this view, the intent is to make the testing procedure objective and standardized in an aim to treat everyone fairly (i.e., equally). The recent edition of Standards for
Educational and Psychological Testing (2014) characterizes fairness to represent this objective and standardized view that promotes equal treatment of individuals in a testing context. In the Standards, investigation of fairness includes such practices as eliminating culturally biased items (through Differential Item Functioning and Differential Test Functioning techniques), standardizing test administration conditions, providing equitable treatment for special education students and English language learners, providing access to the construct of the test, and presenting guidance on test score interpretation and use. These practices can be argued to be the spirit of discussions around fairness in the psychometric tradition (Camilli, 2006, 2013; Dorans & Cook, 2016; Xi, 2010).

Inspired by the psychometric tradition as ascribed in the Standards, classroom assessment scholars have also attempted to interpret psychometrically-based fairness practices into classroom assessment (Camilli, 2013; Herman & Cook, 2019). Herman and Cook argued that psychometric and classroom assessment processes are similar in that they fundamentally share the premise that fairness is a key foundation for examining the quality of assessment, but they also acknowledged that classroom assessment can entail a broader conceptualization of fairness. Aligned with the Standards, Herman and Cook discussed four core issues associated with fairness in classroom assessment: fairness in treatment during assessment, fairness as reducing measurement bias, fairness as access to the construct being measured, and fairness as an opportunity to learn. They also emphasized the accommodated practices for students with disability and English language learners as additional areas for research on fairness in classroom assessment contexts.

While the translation of psychometric principles of fairness into a classroom assessment context may be useful, the epistemological and contextual differences across these two contexts warrants interrogation. In the psychometric tradition, assessment (or measurement) is a process to represent learning through a set of operations (e.g., items). In this tradition, learning, as the measurement object, is either ontologically conceptualized to have ‘objective’ reality (i.e.,
independent from subjective mind) causing the variations in test scores or have been understood based on a set of operations that represent a sample of a domain behavior to which the results of test scores are extrapolated (Borsboom, 2005; Elwood & Murphy, 2015; Gipps, 1994; Michell, 1999). This kind of assessment largely considers test-takers as autonomous individuals and disregards the social, cultural, historical, economic, and family contexts as well as school and learning culture within which a student lives and takes a test (Elwood & Murphy, 2015; Maul, Irribarra, & Wilson, 2016). Accordingly, the psychometric view of fairness includes principles that are largely driven by prioritizing equal treatment across individuals (here students) and is rooted within a cultural and historical practice of standardized testing that allow interpreting fairness within the confines of its object of inquiry (i.e., a test context). Fairness in this view is largely conceptualized from measurement researchers’ perspectives and is investigated in relation to groups (over an individual) designated based on their common characteristics (e.g., age, race, gender).

However, classroom assessment, in its sociocultural conception (Cowie, 2015; Shepard, 2001), is interpreted through the belief systems of teachers and students. The assessment belief systems are shaped by an interplay of historical, cultural, social, economic, family, and policy milieu and are negotiated and enacted within particular contexts of curriculum, school, classrooms, and student and teacher demography (Brookhart, 2003, 2004, 2018; Cowie & Bell, 1999; Elwood & Murphy, 2015; LA Shepard, 2001). It is within the sociocultural conception of assessment that is imperative to explore the perceptions of assessment (fairness) as viewed by teachers and students within diverse classroom contexts and leverage these perceptions to develop classroom assessment theory and practice. Teachers’ and students’ perceptions of fairness in assessment can be analyzed by how their perceptions are influenced by group membership characteristics (e.g., race, gender) among other individual, social, and situational characteristics that constitute fairness views and practices.
A few conceptual and empirical studies have recently embarked on investigating fairness in classroom assessment as a phenomenon viewed by teachers and students embedded within a socio-cultural context (Murillo & Hidalgo, 2017, 2020; Rasooli, DeLuca, et al., 2019; Rasooli et al., 2018; Rasooli, Zandi, et al., 2019; Sonnleitner & Kovacs, 2020; Tierney, 2013, 2014).

Tierney (2014) interviewed six experienced English language arts teachers to gain an insight into how these teachers conceptualize fairness in their assessment practices. She reported that teachers conceptualized classroom assessment fairness in relation to issues such as constructive classroom environment, transparent communication, equity and equality, and reflective thinking. Drawing on these findings, she concluded that teachers’ conceptions of fairness in assessment were influenced by a complex interplay of moral beliefs, contextual factors, and theoretical and practical knowledge and experience (Tierney, 2014).

Murillo and Hidalgo (2020) focused on teachers' conceptions of fairness in classroom assessment in Spain. They have interviewed 30 teachers from primary and secondary schools with high and low socio-economic demography. They found that teachers interpreted fairness in assessment based on either equality or equity principles. Based on the equality principle, teachers raised issues of transparency in assessment expectations and criteria, inclusion of curriculum expectations in the exam content, and the need for classroom assessments to prepare students for standardized objective tests as to support students in their future path. Equality views of fairness were articulated by both primary and secondary teachers, with more frequency in high socio-economic schools. With respect to the equity principle, teachers articulated the need to make assessments individualized based on student need, context, and circumstances (i.e., considering student as a whole person), and the need for a continuous assessment throughout the term using multiple methods with attention to student effort, progress, and improvement. In this equity perspective, teachers also questioned solely valuing curriculum-focused assessment and grading and believed that a fair assessment needs to make students more responsible and inherently
motivated learners. Equity views of fairness were articulated by both primary and secondary teachers, with greater articulation by teachers in low socio-economic schools.

I along with colleagues, conducted a systematic review of the literature of fairness in assessment and education contexts that served as a foundation for conceptualizing fairness in classroom assessment (Rasooli, Zandi, & DeLuca, 2018). This review concluded that fairness in classroom assessment should go beyond considering only assessment processes and develop and interpret various dimensions of fairness in the interplay of assessment, teaching, and classroom interactions. Rasooli et al. (2019) extended this theoretical review by first reviewing the theories and definitions of fairness in assessment literature and then using organizational and social psychology theory of justice as a basis to revisit fairness in classroom assessment. They reported fuzziness and ambiguity in definitions of assessment fairness as well as scarce and disappointing attention to theory in the previous empirical studies. This finding was also echoed by Tierney (2013) in relation to classroom assessment literature and Nisbet (2017) in psychometric contexts.

I along with colleagues then used organizational and social psychological theory of justice to conceptualize how students can perceive fairness in the interplay of assessment, instruction, and interaction dynamics within classrooms (Rasooli, Zandi, & DeLuca, 2019). Organizational and social psychological theory of justice consists of three dimensions (i.e., distributive, procedural, and interactional justice) that attempt to outline principles of justice that individuals use to arrive at perceptions of fairness in diverse contexts including workplace, school, or legal settings. Specifically, distributive justice refers to individuals’ perception of fairness of outcome distributions (e.g., distribution of grades) based on three justice principles of equity, equality, and need (Adams, 1965; Deutsch, 1975). The principle of equity suggests that students would come up with perception of in/justice if they think they do not receive what they deserve by comparing their contributions (e.g., effort) to the outcome they receive (e.g., grade). Whereas, according to the principles of equality and need students would perceive in/justice if the
outcomes they receive is not equally distributed (e.g., teacher attention) or is not distributed based on prioritizing the needs.

Procedural justice refers to individuals’ perception of the fairness of procedures for outcome distributions. This dimension encompasses justice principles of consistency (i.e., consistent treatment across time and place), accuracy (i.e., accurate application of procedures), bias-suppression (neutrality in decision-making process), correctability (i.e., correcting wrong decisions), representativeness of voice (i.e., provision of voice), ethicality (i.e., decision-making based on sound moral standards), transparency (i.e., transparent application of procedures), and reasonableness (i.e., reasonable application of procedures) to evaluate the fairness of the procedures according to which outcomes are distributed (Kazemi & Törnblom, 2008; Leventhal, 1980; Thibaut & Walker, 1975).

Interactional justice refers to how individuals are treated interpersonally and how they are provided with information (Bies & Moag, 1986; Greenberg, 1993). Interactional justice has two facets: interpersonal justice and informational justice. Interpersonal justice encompasses justice principles of propriety, respect, politeness, dignity, and caring to evaluate the fairness of interactions among individuals and informational justice considers justice principles of adequacy, truthfulness, honesty, and timeliness to evaluate whether information was communicated among individuals fairly.

Drawing on these three dimensions, justice is defined as students’ perception of a teacher’s or peer’s adherence to principles of justice (e.g., equity) and fairness then can be defined as students’ overall evaluation of a teacher’s conduct based on the combined interaction of distributive, procedural, and interactional principles as well as related key variables (e.g., student personality).

To provide empirical support for the conceptualization of fairness based on social psychology theory of justice, Rasooli et al., (2019) collected 502 written accounts from higher
education students about their experiences of fairness in assessment, teaching, and classroom interactions. The findings of this study supported the argument that students’ experiences of fairness in classroom assessment, teaching, and school interactions can be adequately interpreted based on distributive, procedural, and interactional justice dimensions. More importantly, they observed that more than 60% of students’ fairness experiences were directly relevant to assessment issues. This observation was consonant with the findings of previous research in classroom fairness literature (Horan, Chory, & Goodboy, 2010; Houston & Bettencourt, 1999).

Despite these initial efforts to conceptualize fairness in classroom assessment, the field has a long road ahead to develop and interpret fairness as situated within the historical, cultural, social, and economic contexts of classrooms and reflective of the conceptions of key stakeholders including teachers and students. While previous research has focused on examining students’ perceived fairness in classroom assessment in higher education, it has yet to explore students’ perceptions of fairness in relation to their secondary schools. Cross-cultural research is also needed to explore students’ perceptions of fairness in assessment. These efforts to theorize fairness predicated on teachers’ and students’ conceptions will serve to expand the psychometric-based view of fairness and provide fertile ground for developing a theory of assessment fairness specific for classroom contexts.

1.1 Purpose and Research Questions

The purpose of this study is to extend the program of research on investigating students’ perceptions of fairness through the theory of social psychology of fairness. Specifically, this study aims to explore Ontario-based Canadian first-year undergraduate students’ perceptions of fairness in classroom assessment during their secondary school experiences to further contribute to developing a theory of fairness in classroom assessment. Constructing theory predicated on students’ lived experiences – as is the goal of this research- is important because the quality of formative and summative classroom assessments depends largely on students’ interpretations of
the fairness of these assessments that subsequently influence their agency and actions as a result of assessment information. Students who interpret assessment process as unfair would develop negative attitudes towards the teacher, learning process, and outcomes of formative and summative assessment, conversely those with positive fairness perceptions would more readily value their teachers, learning experiences, and assessments (e.g., Murillo & Hidalgo, 2017; Rasooli et al., 2018, 2019). In addition, this study will contribute to geographic gaps in research as there are currently few studies published on the topic in the Canadian context (Scott et al., 2014) and few studies focusing on classroom assessment fairness in the K-12 context across the world (Bempechat et al., 2013).

To this end, this study will be guided by the following research question:

1. How do first-year undergraduate students perceive fairness in classroom assessment in relation to their secondary school experiences in Ontario, Canada?
2. What psychological and behavioral responses are provoked by the experiences of fairness in classroom assessment?

1.2 Significance

Strong public, policy, and research premises support research on students’ perceptions of fairness in classroom assessment (Herman & Cook, 2019; Lerner, 1980; Ontario Ministry of Education, 2010; Resh & Sabbagh, 2016). From a public perspective, individuals in the society inherently value fairness as they have a basic need to believe in a just world and believe that everyone gets what they deserve (Ellard, Harvey, & Callan, 2016; Lerner, 1980). This inherent and innate value of fairness is also empirically supported by the evolutionary studies of fairness in animals and human. Evolutionary studies have found that humans and animals such as monkeys are engineered with the innate property of fairness that has evolved to manage the costs and benefits induced by cooperation (Baumard, 2016; Debove, 2015). Across diverse societal spheres, including legal, organizational, health and education contexts, individuals were found to
respond in various forms to unfairness, showing the significance of fairness as a moral value (Adams, 1965). Building on this inherent moral drive, students would prefer to be treated fairly within classrooms and would react cognitively, emotionally, and behaviorally in various forms when confronted with unfairness to restore fairness even in the face of personal loss. Within the context of this study, this morally driven need for fairness implies that teachers and students inherently implicate fairness in the cycles of assessment, instruction, and interpersonal relationships within classrooms. This view of fairness as a moral end supports the conduct of this study to investigate how students conceptualize fairness in classroom assessment.

From the policy perspective, international and local policies (Melbourne Declaration on Educational Goals for Young Australians, 2008; OECD, 2015; Ontario Ministry of Education, 2010; UNESCO, 1974, 2014) prioritize the justice movement in education as a principal focus to ensure that all students from diverse backgrounds are provided with opportunities to attain their educational aspirations and goals on an equitable basis. Therefore, justice research supports the policy demands by investigating justice from macro- and micro-levels. The macro-level justice perspective discusses and critiques issues such as distribution of access and funding opportunities to education, distribution of and access to educational places and facilities, and the choice of curriculum content that recognizes the diversity within a society. One dominant line of macro-level justice research in sociology of education relies on using data from large-scale assessments to examine the inequalities in student achievement of curriculum objectives across country, state, district, and school levels.

The micro-level justice perspective discusses issues related to the fairness of assessment, instruction, and interaction dynamics within the classroom (Resh & Sabbagh, 2016). The micro-level perspective focuses largely on students’ and teachers’ perceptions of fairness and is worthy of attention for three reasons. First, it nicely aligns with the fundamental value of democratic policy-making that encourages focusing on student voice and prioritizing fairness as the leading
quality for student assessments (Ontario Ministry of Education, 2010, 2016). Second, while students’ perceptions of fairness are peculiar to their individual beliefs and values, their perceptions would inevitably incorporate collective values of fairness that have been learned in the communities and society that students have lived, not to mention the innate nature of fairness in humans. By studying students’ personal perceptions of fairness, we can inform policymakers of the social construction and mechanisms that have somewhat brought about individual perceptions of unfairness. The third reason touches upon the research premise that is described below. In this way, this study focuses on the micro-level justice perspective and explores student voice in their conceptualizations of fairness in classroom assessment as an initial empirical basis to inform assessment policy.

From the research perspective, empirical studies show that fairness perception can have significant consequences for students. In schools, students experience and learn about fairness and good citizenship and how they might support these virtues in society in their future lives (Gorard & Smith, 2010; Resh & Sabbagh, 2016). Experiences of fairness in schools provide the foundation for students’ perceptions of fairness and influence their civic and democratic attitudes (Gorard & Smith, 2010; Resh & Sabbagh, 2014a). Empirical research has also provided support for the association of students’ perception of fairness and outcome variables including academic learning (Holmgren & Bolkan, 2014), psychological engagement (Berti, Molinari, & Speltini, 2010), self-efficacy (Vallade, Martin, & Weber, 2014), student well-being (Dalbert & Stoeber, 2006), teacher satisfaction (Wendorf & Alexander, 2005), legitimation of and identification with teacher (Di Battista, Pivetti, & Berti, 2014; Gouveia-Pereira, Vala, Palmonari, & Rubini, 2003), motivation (Chory-Assad, 2002), aggression and hostility (Horan, Chory, Carton, Miller, & Raposo, 2013), cheating (Lemons & Seaton, 2011), truancy (Ishak & Fin, 2013), and political trust (Abdelzadeh, Zetterberg, & Ekman, 2015). These findings show that students’ experiences of fairness in the cycles of assessment, teaching, and classroom interactions link with significant
consequences for students within and beyond classrooms. This study aims to explore the
Canadian students’ articulated experiences of fair assessment consequences on their
psychological and behavioral reactions.

Finally, the equity, diversity, and inclusion movement within contemporary education
makes it imperative for classroom assessment practices to be fair, equitable, and inclusive for
students from diverse backgrounds. Recent statistics illustrates the diversity in the demography of
education system and society in Ontario: 40% students with disability in schools, 14% English
language learners in schools, 18.8% children living in poverty, and 29% visible minority
individuals (Statistics Canada Census, 2016; Child and Family Poverty in Ontario, 2017; People
for Education Annual Report on Ontario’s Publicly Funded Schools, 2015). The diversity in the
Canadian schools and society is increasingly growing in light of Canada’s immigration policies;
Statistics Canada reports that by 2036, 25% to 30% of Canada’s population will consist of
immigrants. This growth in diversity in contemporary classrooms provides a solid ground for
classroom assessment research to strive for conceptualizing and enacting fair and equitable
assessment and accommodated practices to enhance student learning. As this study draws on
students’ perceptions of fairness in classroom assessment, it might capture the diverse fairness
conceptualizations of students and contribute to a more comprehensive and diverse interpretations
of fairness in contemporary classroom assessment contexts (Herman & Cook, 2019).

Through all these efforts, the results of this study will take another step toward
suggesting implications for enhancing students’ experiences of fairness in classroom assessment
in Ontario schools. As “injustice anywhere is a threat to justice everywhere” (Martin Luther
King), it can be assumed that moving toward fairer experiences in schools not only directly
impacts positive outcomes within schools, but also bears beneficial fruits for the Canadian society
in general. In contrast, students’ unfair experiences would not only negatively influence their
academic and character education but are also a threat to establishing fairness in other spheres of Canadian society.

1.3 Definition of Key Terms

**Fairness and Justice**: There are traditions of thoughts in the assessment and social justice literature that distinguish fairness from justice. More elaboration on the distinctions between fairness and justice based on these traditions of thoughts will be reviewed in the literature chapter. However, for the purpose of this study, justice and fairness will be distinguished based on the tradition of social psychology theory as the key theoretical framework for this study. Social psychology theory of fairness consists of three dimensions: distributive, procedural, and interactional justice. **Distributive justice** refers to individuals’ perception of fairness of outcome distributions (e.g., distribution of grades) based on three principles of **equity**, **equality**, and **need** (Adams, 1965; Deutsch, 1975). Principle of equity suggests that students would come up with perception of in/justice if they think they do not receive what they deserve by comparing their contributions (e.g., effort) to the outcome they receive (e.g., grade). Whereas, according to the principles of equality and need students would perceive in/justice if the outcomes they receive is not equally distributed (e.g., teacher attention) or is not distributed based on prioritizing the needs.

**Procedural justice** refers to individuals’ perception of the fairness of procedures for outcome distributions. This dimensions encompasses principles of **consistency** (i.e., consistent treatment across time and place), **accuracy** (i.e., accurate application of procedures), **bias-suppression** (neutrality in decision-making process), **correctability** (i.e., correcting wrong decisions), **representativeness of voice** (i.e., provision of voice), **ethicality** (i.e., decision-making based on sound moral standards), **transparency** (i.e., transparent application of procedures), and **reasonableness** (i.e., reasonable application of procedures) to evaluate the fairness of the
procedures according to which outcomes are distributed (Kazemi & Tornblom, 2008; Leventhal, 1980; Rasooli et al., 2019; Thibaut & Walker, 1975).

*Interactional justice* refers to how individuals are treated interpersonally and how they are provided with information (Bies & Moag, 1986; Greenberg, 1993; Rasooli et al., 2019). Interactional justice has two facets: *interpersonal justice* and *informational justice*. Interpersonal justice encompasses principles of *propriety, respect, politeness, dignity, and caring* to evaluate the fairness of interactions among individuals and informational justice considers *adequacy, truthfulness, honesty, and timeliness* to evaluate whether information was communicated among individuals fairly.

Drawing on these three dimensions, justice is defined as students’ perception of a teacher’s or peer’s adherence to principles of justice (e.g., equity). For example, a student might perceive a teacher’s grade for an assignment to be just if it is comparable to the effort that the student put into writing the assignment (equity principle). However, the student might perceive a teacher’s grade to be unjust if the teacher has not provided transparent grading criteria (transparency principle).

*Fairness* then can be defined as students’ overall evaluation of a teacher’s conduct based on the combined interaction of distributive, procedural, and interactional principles as well as related key variables (e.g., student personality). Students’ perception of fairness of an event (e.g., grading event) is shaped by applying various distributive, procedural, and interactional justice principles to an event. For example, consider two students that perceive their grades as unfair. Despite the surface similarity in perceiving their grades as unfair, these students have different underpinning structure of attitudes toward why they believed their grades were unfair. These underpinning structures are shaped by underlying justice principles (e.g., equality, equity, transparency) which these two students apply to come up with a perception of fairness in relation
to their grades. Through this application, students connect their justice principles to shape overall fairness perceptions.

**Classroom Assessment:** In this study, classroom assessment is conceptualized as a phenomenon that is interpreted and enacted within a particular social, cultural, and historical context. Within this context, teachers and students are agents that interpret and make sense of assessment influenced by their past experiences of assessment culture, the contemporary socio-cultural milieu of assessment, learning, curriculum, teaching, and policy as well as the particular demography (e.g., socio-economic, rural or urban) of a school and classroom. Within this complex context, assessment includes teachers and students cooperating in choosing, gathering, examining, and using evidence of student learning on an ongoing basis for various reasons. These reasons include (a) determining the learning objectives, (b) monitoring student progress toward the learning objectives, (c) providing feedback to teachers, students, and parents to decide next steps for learning, and (d) assigning grades (Andrade, 2013; McMillan, 2013). In this context, teachers and students can employ diverse types of assessment (i.e., teacher, peer, and self-led assessments) primarily to provoke content learning and self-regulation skills and to gauge learning objectives summatively (Brookhart, 2004; Cowie & Bell, 1999; Shepard, 2001). Overall, this conceptualization of classroom assessment is congruent with sociological and psychological theories of fairness in that they both conceptualize assessment and fairness as phenomena that should be interpreted within a complex socio-cultural context.

### 1.4 Delimitations of the Study

This study is delimited to focus on exploring first-year undergraduate students’ perception of fairness in classroom assessment in relation to their secondary schools in Ontario, Canada.
1.5 Structure of Dissertation

This dissertation is organized into five chapters: (1) introduction, (2) literature review, (3) methodology, (4) findings, and (5) discussion. This first chapter has outlined the relevant background in relation to the purpose of this study and its research questions as well as the significance for conducting this dissertation. The second chapter reviews the key conceptions of fairness in educational and classroom assessment and enriches these conceptions by using a broad review of the key conceptions of fairness and justice in philosophy, sociology, psychology, economics, and education. This review presents the conceptual investigation into the existing fairness conceptualizations and the directions for future theory-driven research.

The third chapter presents the mixed methods sequential research methodology that were used to answer the research questions. Specifically, it delineates the procedures used for constructing Classroom Assessment Fairness Inventory. The fourth chapter presents the pertinent themes derived from interviews with students for Phase I and the results from Classroom Assessment Fairness Inventory in Phase II. The final chapter discusses students’ perceptions of fairness in relation to the findings of this study and prior literature. It also acknowledges the limitations and directions for future research.
Chapter 2

Literature Review

Do not jump into conclusion with little knowledge/Think and examine from head to toe
Research and fear from shallowness/Look deeply into words, whatever they are (Ferdowsi)

This chapter reviews multidisciplinary literature on fairness and justice as a basis for conceptualizing fairness in classroom assessment. This chapter is divided into four major sections. The first section reviews the key conceptions of assessment and fairness as the central concepts in this study. To enrich the existing conceptions of fairness in assessment literature, the second section reviews justice theory and research in social sciences and humanities, followed by the third section that summarizes conceptions of justice in education literature. Finally, the fourth section amalgamates the discussions in prior sections to propose conceptual bases that can be potentially used to conceptualize fairness in classroom assessment and guides the reader with the conceptual perspective that this study selects to investigate fairness in classroom assessment. Combined, these sections will provide a panoramic view of justice research as a basis for conceptualizing fairness in classroom assessment.

2.1 Fairness in Assessment

2.1.1 Key Conceptions of Assessment

To appreciate the key conceptions of fairness in assessment, it seems logical to first review the key conceptions of assessment and then examine how fairness has been conceptualized in assessment. Assessment can generally be understood in two key contexts: large-scale assessment and classroom assessment. Large-scale assessment is an assessment that is
administered to a large number of students. Large-scale assessment has been dominantly conceptualized through psychometrics perspective. Three philosophies underlie conceptualization of psychometric theory: positivism, pragmatism, and realism (Borsboom, 2005; Maul et al., 2016; Michell, 1999). While these three philosophies might overlap in shaping ontological and epistemological bases of psychometrics, they have distinct delineations on what constitutes psychometrics.

Inspired by empiricism, positivism only counts phenomena that are observable and concrete based on human senses, excluding meta-physical, abstract, and conceptual phenomena. Within positivism, measurement is a process or activity to understand psychological attributes based on a set of procedures that represent the relationship between actual observations and numerical outcomes. Measurement can be defined based on securing a set of procedures to represent a psychological attribute (i.e., operationalism) or a set of procedures to measure an attribute that have been defined in terms of a domain of behavior (behavioral domain theory). Validity is also defined based on correlations between test scores and other outcomes (Cronbach & Meehl, 1955). For example, measurement based on operationalism aims to understand reading comprehension through constructing several items and administering them to participants. In operationalism, the task of measurement is completed through following a set of procedures to represent for example reading comprehension ability. In this view, there is no account of a psychological attribute that is abstract and real in the mind that guides the responses to the test item; rather, the test items are believed to measure the reading comprehension as empirically observed. Measurement based on behavioral domain theory aims to conducts the same procedure but samples the items from a domain to which we want to generalize the findings. Validity of the measurement procedure is then established through correlating the scores of tests measuring the same construct or linking the test scores with performance scores in the test domain.
Within pragmatism, the measurement procedure is understood based on its usefulness for particular uses and outcomes. The usefulness is determined based on goals, motivations, and values of an individual or a group. In pragmatism, the process of measurement is less about understanding a psychological attribute itself and more about the uses that a measurement can be put to work. Validity in this context is about the adequacy and appropriateness (i.e., usefulness) of interpretations and actions based on test scores (Kane, 2006; Messick, 1989).

Within realism, measurement is a process to understand attributes that exist independent from measurement procedure. In this perspective, measurement and testing process reflects the real psychological attribute, for example reading comprehension ability, that causes the variation in test scores. Test items are also believed to be the observable indicators that reflect the psychological attribute (Maul et al., 2016). Validity in this view is considered as the property of the test and assumes (1) existence of the construct and (2) that the construct causes the test response behavior. Therefore, a test is considered valid if it measures what it intends to measure (Borsboom, Mellenbergh, & Van Heerden, 2004).

Building on these three philosophical bases of psychometrics, large-scale assessment can be conceptualized as (a) positivistic: an epistemology to understand learning through a set of operations (e.g., items) and a sample of domain behavior (e.g. math); (b) pragmatistic: a measurement defined based on its consequences and uses; and (c) realistic: measurement is believed to act as a mirror reflecting the ‘objective’ psychological attributes that exist in the individual mind. Validity can also be understood as (a) a correlation between test scores and other test scores (Cronbach & Meehl, 1955); (b) an adequacy and appropriacy of the inferences and actions based on test scores (Kane, 2006; Messick, 1989); and (c) and whether a test measures what it claims to measure (Borsboom et al., 2004).

Classroom assessment has been conceptualized through psychometrics, constructivism, and socio-cultural theory (Elwood & Murphy, 2015; Moss et al., 2005). The key philosophies of
positivism, pragmatism, and realism may underpin the summative (i.e., graded) assessments in classrooms. These philosophies largely consider test-takers as autonomous individuals and disregards the social, cultural, historical, economic, and family contexts as well as school and learning culture within which a student lives and takes the test (Elwood & Murphy, 2015; Gipps, 1994; Moss et al., 2005). Constructivism considers assessment and learning as co-constructed and co-interpreted between teachers and students, with teachers and students actively making meaning of student learning in relation to accumulative evidence. In constructivism, teachers and students use formative assessment to support students to self-regulate meta-cognitively and use feedback to enhance learning. Peer-assessment and self-assessment are also other forms of assessment that support co-constructing the assessment with students. However, the summative assessment resembles to the psychometric conception of classroom assessment in that each student's knowledge is assessed individually to show what they have achieved (Elwood & Murphy, 2015).

Socio-cultural theory highlights the broader social, cultural, economic, and historical contexts within which an assessment happens, and a teacher and student live. Socio-cultural theory largely argues that we need to ‘look in to their [students’] histories and not into their heads’ (Elwood as cited in Elwood & Murphy, 2015, p. 187). This conception of classroom assessment argues that students, as individuals in the society, bring with themselves their various backgrounds (e.g., poverty status, family issues) to the classroom that impact their educational and assessment life. This conception challenges the single knowledge structure that is usually the object of assessment in psychometrics and constructivist conceptions and highlights questions such as whose knowledge is assessed and how assessment can be equitable with diverse learners. This conception also challenges the system-level structures in education and assessment that limit teachers’ agency in acting socially just. Elwood and Murphy (2015) argue that formative assessment needs to extend itself to incorporate more of socio-cultural theory and contend that
summative assessment acts as a structural barrier in liberating teachers’ efforts to assess and recognize multiple ways of knowing.

Combined, psychometrics, constructivism, and socio-cultural theory are the key conceptions that form the foundation for large-scale and classroom assessments. In this study, assessment is understood from a socio-cultural lens as its relevance to fairness inquiry in assessment context is already established (Elwood & Murphy, 2015; Moss et al., 2005).

2.1.2 Key Conceptions of Fairness in Assessment

Four key perspectives of fairness in educational assessment can be considered: (a) psychometric perspective; (b) legal perspective; (c) philosophical perspectives; and (d) social psychological perspective. These perspectives have been inspired by the epistemological bases within psychometrics and psychology, philosophy and sociology, and legal theory. Despite that these perspectives are discussed distinctly; they share overlaps. The key studies are discussed within each perspective.

2.1.2.1 Psychometric Perspective

Fairness has dominantly been conceptualized based on psychometric theory. This is unsurprising as the epistemological roots of educational assessment can be traced back to psychometrics, which has underpinned the assessment practices in educational contexts (Elwood & Murphy, 2015; Gipps, 1994; Haertel & Herman, 2005; McArthur, 1987; Moss et al., 2005; Newton & Shaw, 2014). Within this epistemological understanding of educational assessment, fairness is examined by applying the principles of psychometrics to eliminate factors irrelevant to the construct or student learning. Particularly, fairness has been conceptualized as ensuring equality in the assessment process so that everyone gets the result they deserve according to the extent they have learned (Kane, 2010; Moss et al., 2005). Therefore, fairness in this perspective is ensured through four general views: (a) fairness in treatment during the testing process; (b)
fairness as lack of measurement bias; (c) fairness in access to the construct as measured; and (d) fairness as validity of individual test score interpretations for the intended use (Standards for Educational and Psychological Testing, 2014). Fairness in treatment during the testing highlights standardizing tests, administration, and scoring procedures as well as eliminating construct-irrelevant sources due to influences such as disability, cultural and linguistic background, race, ethnicity, and socio-economic background. Fairness as lack of measurement bias focuses on eliminating biased items through Differential Item Functioning and Differential Test Functioning techniques. Fairness in access to the construct as measured examines factors such as such as disability, cultural and linguistic background, race, ethnicity, and socio-economic background and ensures that these factors would not influence test performance. Fairness as validity of individual test score interpretations for the intended use provides guidance on test score interpretation and use based on the intended construct of the test. These four views contribute to the comparability of the scores across all test-takers by eliminating the irrelevant factors to the construct of the assessment.

As noted by the Standards (2014), the last three views consider fairness as a psychometric quality, indicating the epistemological legacy of psychometrics in conceptualizing fairness in educational assessment. In the fairness studies in educational assessment, the principles of psychometrics are used to exclude factors that are irrelevant to the construct (or ‘learning’) as measured by an assessment tool. This psychometric view can be observed in the many chapters of two handbooks of Fairness in Educational Assessment and Measurement (Dorans & Cook, 2016) and Fairness Issues in Educational Assessment (Karami, 2018) as well as many studies that have addressed the issue of test bias (Zumbo, 2007) and accommodation practices (Abedi, Hofstetter, & Lord, 2004; Sireci, Scarpati, & Li, 2005).

In classroom assessment, scholars have also attempted to conceptualize fairness based on the psychometric perspective (e.g., Camilli, 2013; Herman & Cook, 2019). Herman and Cook
argue that psychometric large-scale and classroom assessment are similar in that they fundamentally share the premise that fairness is a key foundation for examining the quality of assessment but acknowledge that classroom assessment can entail a broader conceptualization of fairness. Aligned with the Standards, Herman and Cook consider fairness as a validity issue and discuss four core issues associated with fairness of classroom assessment: fairness in treatment during assessment, fairness as reducing measurement bias, fairness as access to the construct being measured, and fairness as an opportunity to learn. They also emphasize the accommodated practices for students with disability and English language learners as additional areas for research on fairness in classroom assessment contexts.

Overall, the psychometric perspective plays a key role and is the dominant discourse in conceptualizing fairness in educational assessment.

2.1.2.2 Legal Perspective

Cumming (2008) provides a few conceptions of fairness that have been used in courts in the US, England, and Australia to judge the fairness and equity of large-scale assessments. These conceptions include opportunity to learn, alternative assessments, test validity, exclusion from an assessment and its impact. Opportunity to learn has been conceptualized as whether a student has been presented with the curriculum under assessment, whether a student has been given prior notice with sufficient time about changes to assessment, and whether adequate resources have been given to students to enable them to show what they know. Alternative assessments consider whether students have received various assessment options to demonstrate their learning. Test validity considers whether the uses that a test has been put to are fair. Finally, exclusion from an assessment and its impact considers the fairness in cases, where a student is deprived of taking an assessment due to expulsion or suspension and the outcomes of not taking the assessment for the student. Overall, the legal perspective provides precedents for conceptions of fairness that
conceptualize fairness before (i.e., opportunity to learn), during (i.e., test design), and after assessment administration (i.e., consequences and uses of assessment).

2.1.2.3 Philosophical Perspectives

Critical theory as well as theories in moral and political philosophy have been used to conceptualize fairness in educational assessment.

*Critical Theory*

Inspired by the scholarship on critical theory (e.g., Foucault, 1979), fairness has been conceptualized to critically examine the power of tests as tools for discrimination and control. This view characterizes fairness as including access, curriculum, and assessment cycles and poses critical questions such as who gets taught and by whom, whose knowledge is taught and assessed and whether the values of marginalized and disempowered classes and diverse cultures attending the assessment are represented (Gipps, 1995; Gipps & Stobart, 2009; Klenowski, 2009, 2014; Shohamy, 1998). Hossein Karami (2013) rightly points out that the difference in conceptualizing assessment fairness based on critical and psychometric theories is their starting premises. Critical theory starts from the premise that tests are tools for social control, while psychometric theory starts from the premise that tests are used and interpreted with good intentions. Karami brings a quotation from Spolsky (1995, p. 1) to support his argument as for the use of tests for social control in history.

From its beginnings, testing has been exploited also as a method of control and power – as a way to select, to motivate, to punish. The so-called objective test, by virtue of its claim of scientific backing for its impartiality, and especially when it operates under the aegis and with the efficiency of big business, is even more brutally effective in exercising this authority (Spolsky, 1995, p. 1).

To support his argument as for good intention of tests as viewed by psychometrics, Karami (2013) brings an excerpt from Messick (1980, p. 1020).
One of the key questions to be posed whenever a test is suggested for a specific purpose is ‘Should it be used for that purpose?’ Answers to that question require an evaluation of the potential consequences of the testing in terms of social values, but that is no trivial enterprise. There is no guarantee that at any point in time we will identify all of the critical possibilities, especially those unintended side effects that are distal to the manifest testing aims (emphasis added).

Building on this excerpt and particularly the italicized sentence, he argues that psychometric theory, as articulated by Messick (1980), does not appear to be concerned with the use of tests with the intentional abuse and for the purpose of social control. In psychometric theory, the starting point is to assume that tests are built with positive intentions that might have unintended side effects. Having said that, it is interesting to note that the push for using objective tests in education for a considerable part was historically rooted in an attempt to correct for the biases in teacher assessments (see findings in Bonefeld, Dickhauser, & Karst, 2020; Rasooli et al., 2018 review related to gender and racial bias). In fact, Binet’s development of the IQ test stemmed in part from his recognition that teacher judgements are shaped by irrelevant factors (social class, gender, etc.). In this respect, psychometric practice is often sold as an antidote to the inability of humans to discount irrelevant information. Although psychometrics historically is rooted in a dark view of human trustworthiness, it at the same time assumes a very optimistic view of the intentions of the test users (Personal communication with Borsboom).

Despite that several scholars include critical theory in conceptualizing fairness in assessment implicitly or explicitly (e.g., Gipps & Stobart, 2009; Karami, 2013), they still discuss fairness in relation to and under validity argument, showing the influence of psychometrics theory in their conception of assessment fairness even if they have adopted a more sociologically driven theory such as critical theory.
Given that the critical understanding of fairness is well-established in large-scale assessment contexts, it is also significant for classroom assessment to gradually embrace this view of assessment. The major implication of this perspective for classroom assessment appears to include conversations with pre-service teachers to critically examine the uses and consequences of high-stakes tests on students, teachers, schools, and learning process and efforts to minimize the negative consequences in collaboration with their students, colleagues, parents, and principals. This view of fairness is significant to broaden our assessment conceptualizations to include discussing negative power of assessments along with positive power that is usually discussed in relation to formative assessment. Through this discussion, teachers can be empowered with fairness reflections that aim to minimize the negative impacts of standardized assessments.

**Political and Moral Philosophy**

The third approach to conceptualizing assessment fairness is influenced by moral philosophy. This approach generally takes the normative view, where fairness of an action is judged based on its alignment with prescriptive theories of justice (Frankena, 1973). Influenced by the work of moral and political philosophers such as John Rawls and Amartya Sen, Kunnam (2018) conceptualizes fairness and justice as two normative principles to establish fairness in assessment contexts. Principle of fairness relates to individual test-takers and principle of justice relates to institutions. Principle of fairness includes four subprinciples: equal access and opportunity to learn, consistency and meaningfulness of the assessment, bias-free assessments, and standardization of administration and procedures of assessment. Principle of justice includes two subprinciples: assessment institution ought to produce beneficial consequences to the test-takers, and assessment institutions ought to promote justice through public reasoning. For the justice principle to prevail, fairness principle ought to be met first. Further, all the fairness and justice subprinciples should be individually met.
In addition to Kunnan’s (2018) work, there are other scholars that have attempted to conceptualize fairness and ethics for assessment, especially for the large-scale assessment contexts. Interested reader can refer to Kunnan’s book (Chapter 3) for a nice and concise review of these few other approaches. Having said that, Kunnan’s framework is the most recent, significant, and systematic study using moral and political perspectives.

The moral and political perspective to fairness has also been taken up in the classroom assessment literature. Inspired by Aristotelian virtue ethics as well as practical and professional ethics, classroom assessment studies have argued that a teacher is a moral individual whose assessment actions are inherently driven by the moral principles (Green, Johnson, Kim, & Pope, 2007; Tierney, 2014). Green et al., (2007) used practical and professional ethics lens and has conceptualized ethics in assessment as including two principles of do no harm and avoid score pollution (i.e., assessment results should represent only achievement factors). Inspired by Aristotelian notion of phronesis, Tierney (2014) interviewed teachers to leverage their practical moral experience as a window to conceptualize fairness in classroom assessment. Despite that both studies have used normative theories, they have based their conceptualizations and analyses of their data on inductive interpretation, without exercising judgment. However, these two pioneer studies are very useful in setting the stage to broaden our epistemological bases to conceptualize fairness in classroom assessment contexts. As fairness in classroom assessment literature is burgeoning, there is a need for empirical studies beginning to investigate how teachers, students, and parents understand fairness using relevant and useful theories.

Overall, a few studies have also attempted to conceptualize assessment fairness through philosophical perspectives including critical and political and moral theories.

2.1.2.4 Social Psychology Perspective

The social psychology theory of justice has recently been used to conceptualize fairness for classroom assessment contexts. This theory encompasses the justice principles that students
and teachers use to arrive at conceptions of fairness in relation to classroom assessment, instruction, and interpersonal relationships. This theory also includes the emotional, cognitive, and behavioral consequences of fairness (Rasooli et al., 2018, 2019, 2019). This theory typically characterizes fairness through three dimensions: distributive justice, procedural justice, and interactional justice. Students evaluate distributive justice based on whether outcomes were distributed equitably, equally, or need-based. Distributions are perceived to be equitable when outcomes are distributed differentially based on comparable contributions (Adams, 1965; Rasooli et al., 2019). Distributions are also perceived to be fair if outcomes are provided equally among students or/and are divided based on ones who need them (Deutsch, 1975; Rasooli et al., 2019).

Students evaluate procedural justice based on whether classroom procedures are consistent, reasonable, bias-free, transparent, representative of student voice, correctable, and ethical (Leventhal, 1980). Classroom procedures are perceived fair when a teacher (a) keeps promises and perform classroom procedures consistently; (b) has logical and meaningful expectations from students; (c) behaves neutrally and is devoid of sexual, religious, and ethnic biases; (d) implements classroom procedures transparently and include students’ voices in these procedures; and (e) apologizes and corrects the wrong decisions made about a student and acts according to the ethical norms accepted within an academic community (Leventhal, 1980; Thibaut & Walker, 1975).

Students and teacher evaluate interactional justice if interpersonal relationships are based on respect and care and information are provided with adequate, truthful, justified, and timely (Bies & Moag, 1986; Rasooli et al., 2019). Drawing on these three dimensions, the conceptual framework of social psychology theory of justice within classroom assessment contexts distinguishes between justice and fairness, arguing that justice refers to principles that students use to arrive at justice perceptions, while fairness is a global evaluation based on the interplay of all these distributive, procedural, and interactional justice principles. For example, students may
perceive injustice if a teacher provided feedback disrespectfully (justice principle of respect).

However, students’ global perceptions of fairness of feedback might be formulated based on several justice principles including consistency, transparency, adequacy, equity, and respectfulness. Students’ global perceptions of fairness subsequently impact students’ cognitive, emotional, and behavioral outcomes within classroom context and beyond (Rasooli et al., 2019).

In this way, this theory accounts for the multidimensional view of fairness that conditions conceptions of fairness for teachers and students within the dynamic socio-cultural environment of classroom assessment. This study leverages social psychology theory as a basis to interpret students’ perceptions of fairness in assessment as it lends itself usefully to investigating individuals’ perceived fairness.

2.2 Justice Theory and Research in Social Sciences and Humanities

The research on justice has been promoted in several social science fields including sociology, psychology, economics, and political science and philosophy (Sabbagh & Schmitt, 2016). “Justice research has thus historically developed within a multidisciplinary framework” (P. 2). Justice is not a singular concept; rather, justice literature is informed by a range of perspectives including ‘philosophy, sociology, psychology, and economics’ (p. 2); ‘different models, and especially considerations of justices in terms of ‘distributions, procedures, retribution, or restoration’ (p. 2). In the following sections, key conceptions of justice are described as related to each of the following perspectives: philosophy, sociology, psychology, and economics.

2.2.1 Justice in Philosophy

Justice has been extensively discussed in philosophy in the works of Ancient Persian and Greek philosophers such as Aveicna, Plato, and Aristotle to contemporary philosophers such as Kant, Rawls, John Stuart Mill, Mulla Sadra, and Bentham. These philosophers were informed by moral and political philosophy and attempted to advance various conceptions of justice based on
normative justice (i.e., justice conceptions that prescribe what act is fair and just) (Frankena, 1973). Despite long-standing discussions around justice in philosophy, research on justice in political philosophy had lost significant attention; however, the discussions were reinvigorated after the influential work of Rawls in 1971 (Meyer & Sanklecha, 2016). Rawls rekindled the interest in philosophy of justice in the second half of 20th century and has thus become the most well-known political philosopher in justice in the current era. Here, I would summarize his justice ideas briefly.

Rawls argues that society is a venue for cooperative interaction among free and equal individuals, where everyone seeks their conception of good life. Rawls (2001) provided principles of justice that aim to govern how a society should be organized and maintained so that everyone can follow their conception of good life under equal and free conditions. Rawls conceptualized his justice theory within a liberal democratic society and proposed principles that a liberal democracy should follow to be considered as just. Rawls believed that the conception and adoption of principles of justice need to be justified for those “who disagree with us or to ourselves when we are of two minds” (Rawls, 1971, p. 580). He argued that the function of principles of justice is to “assign … basic rights and duties and … determin[e] … the proper distribution of the benefits and burdens of social cooperation” (Rawls, 1999, p. 5). Based on this definition, Rawls proposed his two principles of justice (2001, pp. 42-43):

1. Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all; and
2. Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least-advantaged members of society.
In other words, Rawls proposed his two principles of justice based on distributing equal rights to basic liberties to everyone, distributing equality of opportunity to everyone, and distributing primary goods to the greatest benefit of worst-off.

After Rawls, other scholars contributed to the normative conceptualization of justice in philosophy. Wolff (2013) argued that two major traditions can be recognized in the conceptualizations of justice: one follows the work of John Stuart Mill and another from Hegel (Meyer & Sanklecha, 2016). In the Millian tradition, well-being and liberty are the key concepts (see Broome, 2004), while in the Hegelian tradition, recognition is considered as the core concept (see Honneth & Fraser, 2003). There were also other philosophical traditions including feminist theory that attempted to critique Rawls’ theory of justice and provide a reconceptualization. In this respect, Young (1990) argues that conceptualizing justice through distributive justice is limiting because it does not entirely address oppression and domination in the society due to issues such as marginalization of groups identified as disadvantaged as well as the social order and structure of the society that has led to injustices, which need more transformative responses than redistribution of resources. Another feminist scholar (Held, 1993) proposed similar critiques in content and suggested ‘care ethics’, that is, moral and just actions are driven by caring interpersonal relationships (Meyer & Sanklecha, 2016). Combined, philosophical traditions to conceptualize justice is greatly informed by Rawls’ theory of justice and has also built on other conceptions such as feminist, Millian, and Hegelian traditions to provide additional theories of justice.

2.2.2 Justice in Sociology

Sociologists are different from philosophers in their approach toward the study of justice. Sociologists do not deal with normative justice (Liebig & Sauer, 2016). Rather, sociologists tend to describe what individuals conceptualize as just in particular societal conditions as well as explain why individuals have such conceptions of justice in specific societal conditions and time.
periods. Since the 1950s, social and behavioral sciences have focused on empirical research that has explored justice as interpreted and enacted by individuals, groups, and institutions in the society. Specifically, the sociology of justice examines three questions: “1. Why is justice regarded as a desirable state in almost all societies? 2. What ideas or conceptions relate to justice? 3. How do conceptions of justice shape the individual and life in society, and vice versa?” (Liebig & Sauer, 2016, p. 37). These questions were supported by early research in sociology, especially the work of Samuel Stouffer on relative deprivation theory as well as the work of Homans (1961) on equity theory. Both these theories contend that individuals compare their situation with the situation of similar others, nurture a sense of entitlement through this comparison, and perceive dis/content if their situation mis/aligns with their expectations.

After this early work, justice as a social phenomenon has had four foci: (a) societal conditions: for example, a course of actions by an institute that makes collective understanding of in/justice, like in a social condition of revolutions and its impact on justice conceptions of its members; (b) institutional design of a society: institutions such as tax institute that are at play to regulate distributive justice; (c) social conditionality: the social conditions that people live at a particular and shared time and place that shape similar justice conceptions; and (d) social consequences of in/justice that relate back to society (Liebig & Sauer, 2016). This emphasis on social contexts and conditions is the key distinction in the departure for conducting research on sociology of justice.

To respond to these focus areas, four kinds of sociological justice research can be considered: institutional analysis, discourse analysis, attitude research, and behavioral consequence research. Institutional analysis research focuses on the distribution policies and practices for distributing goods within an institution. Institutional analysis can focus on identifying the blueprint for conceptions of justice according to which goods are distributed in an institute as well as how this distribution can form the justice conceptions of individuals that are
somehow associated with the institute. Further, institutional analysis examines how distributions at an institute can predict the social consequences of such distributions (Liebig & Sauer, 2016). An example can be examining the policies and practices of resource distributions within Queen’s University to identify the blueprint for Queen’s justice conceptions as well as how this conception influences the justice conceptions of Queen’s staff and students. Discourse analysis can focus on analyzing the verbal as well as written documents and texts within diverse societal spheres and levels to interpret the implicit and explicit discourses for distributing resources. Discourse analysis focuses on under what societal conditions particular discourses of justice arise as well as how justice is conceptualized in response to the societal conditions (Liebig & Sauer, 2016).

Attitude research on justice has got more attention after 1980s as sociological research was more interested in grand macro theories such as Marx theory during 1960s to 1980s (Liebig & Sauer, 2016). Attitude justice research focuses on why individuals choose principles of equality, equity, need, or entitlement (entitled in relation to a status/position) in a particular social context for distribution of goods. Finally, behavioral consequence research focuses on how justice conceptions influence outcomes in societal conditions (Liebig & Sauer, 2016).

Overall, sociology of justice aims to understand the role of societal conditions in contributing to shape individuals’, groups’, and institutions’ conceptions of justice in a particular spatio-temporal context and how these conceptions, in turn, influence societal conditions.

2.2.3 Justice in Psychology

The study of sociology of justice is different from psychology of justice (Liebig & Sauer, 2016). Sociology of justice assumes that individuals’ conceptions of justice are informed by social conditions prevalent within a particular social and temporal context. Individuals’ conceptions of justice are also assumed to have consequences that influence the social conditions and the structure of social contexts. For example, individuals’ social standings in the society can impact how they conceptualize justice, and their conceptualizations might subsequently influence
the social structures as well as their social standings. In contrast to sociology that seeks to understand how social mechanisms in particular social contexts shape conceptions of justice, psychology seeks to identify psychological mechanisms based on which individuals perceive fairness and justice (Liebig & Sauer, 2016).

Research on psychology of justice has been undertaken for 50 years (Gollwitzer & van Prooijen, 2016). Research on psychology of justice focuses on three levels: individual, interpersonal, and intergroup. Individual level examines what motivates people to act according to justice principles. In other words, why, under what conditions and circumstances, an individual cares about justice. Interpersonal level examines how individuals perceive justice in relation to distributive, procedural, interactional, and retributive justice. Distributive justice considers fairness of outcome distributions; procedural justice considers fairness of procedures used for distribution of outcomes; interactional justice considers fairness of interpersonal relationships and information adequacy, justification, and truthfulness; retributive justice considers how individuals respond to violations of justice principles. Finally, intergroup level examines justice-related conflicts across groups, institutions, and cultures (Gollwitzer & Van Prooijen, 2016).

Within individual level, psychological justice research has focused on three issues: justice-related motive, personality traits, and moral self. Research on justice-related motive has examined whether justice is a basic human need or is a proxy representing other basic needs including managing uncertainty, maximizing self-interest, and securing social acceptance. Research on personality traits examines how individuals differ from each other in their justice attitudes and behavioral dispositions. Research on the moral self examines how individuals maintain a positive moral self-concept even if their actions can sometimes be at odds with their justice and moral principles (Gollwitzer & Van Prooijen, 2016).

Research on justice-related motive has shown that individuals show positive and negative reactions in response to in/justice. These responses indicate that justice is an important motive for
individuals to act upon. Some scholars have argued that self-interest is the basis for humans to be interested in justice-related actions (see Montada & Maes, 2016). Some other scholars have shown that humans can sacrifice their self-interest in some cases to respond to injustice and restore justice (e.g., Lerner, 1980). Melvin Lerner (1980) has proposed justice motives as an inherent need for humans, as individuals need to believe that they are living in a just world, where everyone gets what they deserve (Gollwitzer & Van Prooijen, 2016).

Research on justice as a personality trait has focused on how individuals differ in their justice perceptions as well as the reactions they show in response to in/justice. Three traits have been noted as significant. First, social value orientation describes how individuals who are more prosocial tend to vary in their perception of justice based on whether others have received the just treatment, while individuals who are more pro-self tend to vary in their perception of justice based on whether they vs. others received the just treatment. Justice sensitivity (i.e., the extent individuals are sensitive toward violation of justice) and belief-in-a-just world (i.e., the extent individuals have weak to strong beliefs in a just world) are additional traits that describe variations in an individual’s perception of fairness and justice as well as subsequent reactions (Gollwitzer & Van Prooijen, 2016).

Research on the moral self focuses on justifications and rationalizations that individuals bring to give credit to the im/moral actions they have conducted. Three constructs have been proposed to describe the rationalizations that individuals propose to justify the misalignment between their moral principles and their actual immoral actions: moral hypocrisy, moral disengagement, and moral self-regulation (Gollwitzer & Van Prooijen, 2016). Moral hypocrisy (Batson, Thompson, Seuferling, Whitney, & Strongman, 1999) and moral disengagement (Bandura, 1999) have been proposed to describe how individuals cannot live up to their conceptions of morality and deviate from that in practice (moral hypocrisy) and how they attempt to justify their immoral reactions to preserve their self-evaluation (e.g., taking someone’s bike
without asking them is just borrowing it). The third construct, moral self-regulation considers how individuals can leverage their past moral good deeds to act as a basis to evaluate themselves as being moral and unbiased as well as using the past good deeds as a bank account where they can use to weigh their past good deeds with the current bad deeds (Effron & Monin, 2010; Monin & Miller, 2001).

Interpersonal level of justice considers three issues: how individuals assess justice; what individuals do to enforce justice; and how individuals restore justice. Individuals assess justice based on distributive, procedural, interactional, and retributive justice principles. For enforcing justice, let’s examine punishment and retributive justice to elucidate the psychology of justice research in an interpersonal level. Research has shown that individuals’ main motivation behind punishment is more to restore justice that has been undermined as a result of an unjust practice (i.e., just desert consideration) than to prevent future occurrence of the unjust act (i.e., utilitarian motivation). Through punishment, the interpersonal relationship within the group is maintained and positive consequences can ensue. The principal finding for restoring justice is that people tend more to punish the violator of justice than to forgive. Additionally, individuals’ decisions to punish and its degree are influenced by the degree of interpersonal relationship with the offender (Gollwitzer & Van Prooijen, 2016).

In intergroup justice, scholars investigate what motivates members of a group to act unjust against members of another group as well as how members of a victim group perceive the injustice. This line of research is principally predicated on the scope of justice, that is, issues that individuals conceive as the topic of justice falls within their scope of justice and issues that fall outside their scope of justice is fine to be harmed. For example, individuals might kill insects easily, but they might consider killing dogs as unjust; this means that killing insects might not fall within their scope of justice (Gollwitzer & Van Prooijen, 2016).
Collectively, individual, interpersonal, and intergroup levels constitute foci of justice research in psychology. Individual level considers justice motive, personality trait, and moral self; interpersonal level focuses on justice perception based on individuals’ interaction; and intergroup level focuses on justice enactment and perception in the interaction across members of groups.

2.2.4 Justice in Economics

Justice in economics represent another dominant perspective in justice research; however, as I found the concept of justice in economics as less relevant to the purpose of this review, I consider only valuable to note that like justice conceptions in philosophy, sociology, and psychology, justice is conceptualized in economics based on normative and descriptive lens. Normative economics prescribes what justice ought to be and descriptive lens (i.e., positive economics) describes what justice is. Descriptive lens has been a dominant framework in economics (i.e., the field of behavioral economics) to investigate justice empirically using experimental methods. The experimental methods leverage ultimatum game and dictator game, where a proposer is asked to distribute a resource (e.g., money) between him/herself and a recipient. In this game, the proposer would select what amount of the resource s/he would prefer to offer to the recipient. If the recipient accepts the offer, both proposer and recipient will receive their agreed-upon amounts, but if the recipient repudiates the offer, both parties will receive nothing. It appears that various issues including self-interest, equality, equity, efficiency, need, and the particular context of the allocation impact the individuals’ decisions in resource distributions and receipts (Konow & Schwettmann, 2016). The discussion of equity vs. efficiency is the key issue that I have found in justice in economics as relevant to the educational administration and policy.

Overall, justice theory as understood through diverse perspectives including philosophy, sociology, psychology, and economics has been investigated empirically in diverse societal spheres such as political context (Rothmund, Becker, & Jost, 2016), legal settings (Sunshine &
Tyler, 2003; Thibaut & Walker, 1975), health contexts (Siegrist, 2015), workplace (The Oxford Handbook of Justice in the Workplace, 2015), environmental settings (Clayton, Kals, & Feygina, 2016), diverse cultural contexts (Fischer, 2016), family contexts (Dette-Hagenmeyer & Reichle, 2016), and education contexts (Resh & Sabbagh, 2016). Compared to other spheres, empirical research on justice in education is recent. In the next section on justice in education, I will review various conceptions of justice in education including a review of empirical research on justice in education.

2.3 Justice in Education

Justice has been endorsed as significant focal area in education for three primary reasons. First, education is a key lever to promote and establish fairness, justice, and equality for members of society to ensure that all members from diverse backgrounds are provided with equal opportunities to attain their educational aspirations and goals on an equitable basis, with students identified as disadvantaged getting more attention and accommodation to compensate for their disadvantaged backgrounds (Abedi et al., 2004; Cumming, 2008; North, 2006; OECD, 2015; UNESCO, 1974, 2014). Second, academic institutions are platforms, where students experience and learn about fairness, justice, and good citizenship and how they might contribute these qualities to society in their future lives (Gorard & Smith, 2010; Sabbagh & Resh, 2016). Third, teachers are key agents of justice in education and have been found to put fairness and justice as top priorities in their treatment of students (see Mills, Gale, Parker, Smith, & Cross, 2019). These primary reasons coupled with other reasons including justice motive theory (Lerner, 1980) support the growing movement towards justice in education.

Despite this movement, scholars have reported that justice is not a well-defined concept in general education literature including education policy literature (Gewirtz, 1998; North, 2006; Sturman, 1997), assessment literature (Nisbet & Shaw, 2019; Rasooli, Zandi, et al., 2019), and special education literature (Connor, 2014). Based on my readings so far, I speculate that justice
in education research has been extensively discussed in the education policy (see Gewirtz, 1998; North, 2006; Sturman, 1997) and socially just pedagogies (e.g., Connor, 2014; Ladson-Billings, 1998). The rationale that the education policy literature has provoked extensive discussions of justice might be attributed to its affinity with the field of political philosophy, where justice is a hot topic (see Rawls, 1971).

In the following parts, I will first describe key justice conceptions in the policy education literature and then concisely describe key discussions in socially just pedagogies. Further, as justice in education has predominantly taken a group perspective (e.g., gender or racial equity) to recognize the injustices experienced by groups identified as disadvantaged, I will also describe group perspective literature in short. Finally, I will describe a recent empirical approach toward the study of justice in education and argue how this field has largely been neglected in the justice discourse in education literature. Combined, the review of these key literature bases provides an overall picture of justice research in education and directs my argument toward the necessity of additional empirical research on perceived justice in education and assessment.

2.3.1 Justice Conceptions in Education Policy Literature

In her attempt to map the territory of justice conceptions in the education policy, Gewirtz (1998) has initially argued that “justice has been an undertheorized concept” (p. 469) and very little research has attempted to explicitly define justice in education policy literature (see also North, 2006). To address this gap, scholars have written manuscripts to delineate the conceptions of justice in education (Gewirtz, 1998, 2006; Gewirtz & Cribb, 2002; North, 2006; Sturman, 1997). These scholars have argued that justice has been historically conceptualized based on distributive justice. Rawls (1972, p. 7) defines distributive justice as following: “the subject matter of justice is the basic structure of society, or more exactly, the way in which the major social institutions…distribute fundamental rights and duties and determine the distribution of advantages from social co-operation.” Based on this definition, Gewirtz (1998) argued that in
Rawls’ thinking, justice is equivalent to distributive justice and his conception of justice includes two principles: equality of opportunity (weak liberal definition of justice) and equality of outcome (strong liberal definition of justice). Lynch (as cited in Gewirtz, p. 472) defines equality of opportunity as “unequal results are justified if everyone has an equal opportunity to succeed.” Overall, equality of opportunity includes equality of rights as well as equality of access and equality of participation. Equality of outcome guarantees equality by compensating for the different outcomes of groups identified as disadvantaged through direct intervention.

Gewirtz (1998) and North (2006) drew on the works of Young (1990) and Fraser (1997) to argue for broadening the conception of justice in education that has been mainly understood through distributive justice to include an additional conception called ‘relational justice’ (Gewirtz, 1998). Relational justice is also referred to as ‘cultural justice’ or ‘recognition’ (Fraser, 1997). Relational justice includes, “all aspects of institutional rules and relations insofar as they are subject to potential collective action” (Young, 1990, p. 16). Young’s conception of injustice includes five forms of oppression: exploitation, marginalization, powerlessness, cultural imperialism, and violence. Exploitation refers to the use of the fruits of the labor of a group by another group; marginalization refers to exclusion of individuals or groups from meaningful and useful participation in a society; powerlessness is the lack of status, authority, and sense of self that prevents from respectful listening to a person’s voice; cultural imperialism refers to stereotyping the oppressed group’s culture as well as experiences and interpretations of the social life and imposing the dominant group’s culture, interpretations, and experiences on the oppressed group; and violence refers to random efforts to intimidate, humiliate, damage, and destroy a person (Young, 1990). In Young’s perspective, as noted by Gewirtz (1998), distributive injustice can be the origin or product of these five forms of oppression. However, none of these forms of oppression can be attributed solely to distributive injustice (i.e., the way goods are distributed), rather, the social structures and relations in which distributions occur need also to be considered.
In a similar vein, Fraser (1997) argues that there is a movement from economic justice (or distributive justice), where individuals and groups identified as disadvantaged strive to defend their ‘interests’, put an end to ‘exploitation’, and triumph ‘redistribution’, to cultural justice, where individuals and groups identified as disadvantaged strive to defend their ‘identities’, put an end to ‘cultural domination’, and triumph ‘recognition’ (pp. 2-3). Based on this argument, Fraser argues for a synergy between economic (i.e., politics of redistribution) and cultural justice (i.e., politics of recognition) to appreciate both how economic disadvantage prevents individuals from equal participation in the society as well as how informal and formal cultural rules, institutionalized in the state and economy, are unfairly against some individuals and groups.

Fraser also suggests affirmative and transformative strategies to remedy the injustice transpired in both economic (i.e., distributive) and cultural (i.e., relational or recognition) justice:

By affirmative remedies for injustice I mean remedies aimed at correcting inequitable outcomes of social arrangements without disturbing the underlying framework that generates them. By transformative remedies, in contrast, I mean remedies aimed at correcting inequitable outcomes precisely by restructuring the underlying generative framework (Fraser 1997, p. 23).

Building on these arguments by Young (1990) and Fraser (1997), some education scholars including (Gewirtz, 1998, 2006; Gewirtz & Cribb, 2002; North, 2006) have highlighted a multidimensional conceptualization of justice in education with attention to both distributive and relational justice. Predicated on this multi-dimensional conceptualization, Gewirtz (2006) puts forward a framework for justice research in education. She argues for the contextualized interpretation and enactment of justice rather than abstract thinking. In this contextual interpretation, a researcher needs to consider (a) other norms that compete with justice considerations such as morality or care, (b) contextual constraints and barriers such as legal, policy, and economic constraints; and (c) consideration of various hierarchical and nested levels.
such as district, school, classroom as well as stakeholders’ various perspectives including teachers, students, principals, and parents. This framework for justice research responds to and aligns well with the call of other educational researchers that emphasize the need to investigate taken-for-granted injustice occurring within education contexts at the individual levels of identity development as well as interpersonal relationships between students, peers, and teachers (Britzman, 2003; Ellsworth, 1997; Kumashiro, 2004; North, 2006).

2.3.2 Socially Just Pedagogies

Extensive research has been conducted on socially just pedagogies (e.g., critical pedagogy, authentic pedagogy, productive pedagogy, creative pedagogy, and transformative pedagogy) to teach and promote justice concepts by building various critical and transformative perspectives with students to reflect on social and learning outcome. Socially just pedagogies aim to critically review the normative and structural schooling practices and argue for ideologies, beliefs, and practices that promote teachers’ and students’ agency to reflect critically on current oppressive forms within education and society. These pedagogies therefore aim to develop learning and justice outcomes by nurturing students’ critical capacity to reflect on and challenge unjust educational and social outcomes and actively contribute within and beyond education sphere to transform these unjust practices (Resh & Sabbagh, 2016). The pedagogies are particularly used for students identified as disadvantaged to support and empower them to leverage their learning as a basis to transform academic and social outcomes for themselves and similar others (Resh & Sabbagh, 2016).

For example, Mills et al. (2019) have used interviews and stimulated recall techniques to explore teachers’ activist and justice dispositions in refugee schools in Australia. Building on Fraser’s and Bourdieu’s philosophies, they found that teachers showed activisms within and beyond school to perform socially just in relation to refugee students. Specifically, teachers showed affirmative activism, where they showed response to primary needs of students for food
and clothing within schools and brokered with technical and further education sectors beyond school to provide skill-learning opportunities for these refugee students. Teachers also showed transformative activism within and beyond their schools through enacting democratic practices within classrooms including giving voice and modelling to students to encourage them to become socially just activists in their communities.

2.3.3 Group Perspective to Justice in Education

Educational research has mainly taken group perspective to examine justice in education. In his book on justice in education, Sturman (1997) has discussed justice issues in relation to groups that have dominantly been identified as disadvantaged in research as well as policy statements due to markers such as, but not limited to, socio-economic background, gender, race, and culture. Justice ethos in education contends that equality of educational opportunity and access for everyone coupled with equality of outcomes through compensation for groups with previous and current disadvantaged background will eliminate or reduce learning gaps. The gaps are provoked by systemic social and economic structures that result in discrimination against student from disadvantaged background (Resh & Sabbagh, 2016; Sturman, 1997).

To address group inequalities, research in education has first discussed justice issues in relation to groups with ‘socio-economic disadvantage’ in 1970s, with following research including groups with additional markers such as disability, culture, aboriginality and indigeneity in 1980s, and gender in 1990s (Sturman, 1997). Sturman argues that in 1990s, another approach based on intersectionality of these markers to identify factors that puts a student at an educational risk is proposed, but there is still a strong tendency to include group perspective in research as well as policy documents. For detailed discussion of each of markers based on group perspective, please refer to Sturman’s (1997) book.

2.3.4 Empirical Research on Perceived Justice in Education
Justice in education has been discussed through various perspectives including philosophy, policy, and socially just pedagogies. These perspectives have been leveraged to examine justice-related issues for groups that have been identified as disadvantaged due to markers such as gender, race, culture, and disability (Sturman, 1997). Drawing on these perspectives, justice research in education has dominantly focused on examining the structural resource distributions and recognitions for groups identified as disadvantaged. These discussions include issues such as distribution of access and funding opportunities to education, distribution of and access to educational places and facilities, and choice of curriculum content that recognizes diversity within a society. Coupled with this structural perspective, additional research has focused on school and classroom level justice-related issues to critically examine the pedagogies for groups identified as disadvantaged and to empower these groups to transform their social and educational outcomes (Ladson-Billings, 1998; North, 2006; Sabbagh & Resh, 2016).

Recently, there is an emerging empirical research tradition that aims to examine perceived justice in education. Resh and Sabbagh (2016) conceptualize this empirical research on perceived justice as including five spheres: (a) right to education including resource allocation to realize this right; (b) educational places (student composition; selection to classes, tracks, ability based learning groups); (c) pedagogy; (d) grading; and (e) teacher–student relations (help, respect, attention, care). The first two spheres (i.e., right to education and educational spaces) are issues of justice that focus on structural and macro-level of justice. Pedagogy, grading, and teacher-student relationships are issues that relate to the micro-level of justice.

For individuals and groups to evaluate the justice of distributions, procedures for distributions, and interactions within each of these five spheres, there is a need to answer three questions: (a) what do individuals and groups perceive as just?; (b) what is the perceived difference between experienced distributions, procedures, and interactions and individuals’ and groups’ ideal just conceptions based on the relevant justice principles; and (c) what are the
psychological and social consequences of justice perception? (Resh & Sabbagh, 2016). Overall, individuals and groups arrive at justice evaluations based on a comparison of how they have been treated with how they expected to be treated. Individuals and groups arrive at these evaluations based on the application of justice principles in three dimensions: distributive justice, procedural justice, and interactional justice (See Table 1).

Table 1. Justice dimensions and their principles

<table>
<thead>
<tr>
<th>Dimensions of Justice</th>
<th>Justice rules</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distributive (Adams, 1965; Deutsch, 1975)</td>
<td>Equity</td>
<td>Outcomes are distributed based on contributions</td>
</tr>
<tr>
<td></td>
<td>Equality</td>
<td>Outcomes are distributed equally</td>
</tr>
<tr>
<td></td>
<td>Need</td>
<td>Outcomes are distributed based on needs</td>
</tr>
<tr>
<td>Procedural (Thibaut &amp; Walker, 1975; Leventhal, 1980)</td>
<td>Consistency</td>
<td>Procedures are implemented consistently across people and time</td>
</tr>
<tr>
<td></td>
<td>Voice (or representativeness)</td>
<td>Procedures consider the concerns of all those involved</td>
</tr>
<tr>
<td></td>
<td>Bias suppression</td>
<td>Procedures are neutral and avoid personal bias</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>Procedures are based on accurate information</td>
</tr>
<tr>
<td></td>
<td>Correctability</td>
<td>Procedures are corrected when identified as wrong</td>
</tr>
<tr>
<td></td>
<td>Ethicality</td>
<td>Procedures uphold ethical and moral values</td>
</tr>
<tr>
<td>Interactional (Bies &amp; Moag, 1986; Greenberg, 1993)</td>
<td>Respect</td>
<td>Procedures are implemented sincerely and politely</td>
</tr>
<tr>
<td></td>
<td>Propriety</td>
<td>Procedures are implemented refraining from improper remarks</td>
</tr>
<tr>
<td></td>
<td>Truthfulness</td>
<td>Procedures are explained honestly</td>
</tr>
<tr>
<td></td>
<td>Justification</td>
<td>Procedures are explained adequately</td>
</tr>
</tbody>
</table>

Distributive justice evaluates the distribution of outcomes based on principles of equality, equity, and need. Procedural justice evaluates the procedures for outcome distributions based on principles of consistency, voice, bias suppression, accuracy, correctability, ethicality, respect, and propriety. Interactional justice is evaluated based on classroom interactions (principles of respect...
and propriety) and communication of information within the classroom (e.g., principles of truthfulness and justification). Combined, individuals and groups evaluate justice based on these principles and subsequently respond to in/justice emotionally, behaviorally, and socially.

Drawing on these three dimensions, empirical qualitative and quantitative studies have used various conceptual and methodological frames to investigate students’ and teachers’ perceptions of fairness. Compared to quantitative studies, few qualitative studies can be identified in the literature. A few qualitative studies to date have been conducted within education literature (Bempechat et al., 2013; Buttner, 2004; Chory, Horan, & Houser, 2017; Čiuladienė & Račelytė, 2016; Horan et al., 2010; Houston & Bettencourt, 1999; Israelashvili, 1997; Robbins & Jeffords, 2009), and in assessment literature (Alm & Colnerud, 2015; Lizzio & Wilson, 2008; Murillo & Hidalgo, 2017). Most of these studies adopted organizational and social psychology theories of justice and used interview methods to explore students’ perception of fairness. These studies were conducted in the US context (Buttner, 2004; Chory et al., 2017; Horan et al., 2010; Houston & Bettencourt, 1999; Robbins & Jeffords, 2009), and the remaining ones were in Israel (Israelashvili, 1997), Australia (Lizzio & Wilson, 2008), Russia (Bempechat et al., 2013), Sweden (Alm & Colnerud, 2015), Lithuania (Čiuladienė & Račelytė, 2016), and Spain (Murillo & Hidalgo, 2017). Even among these few qualitative studies, very few collected data from school students, with additional studies recruiting university student participants. Across all these studies, the findings showed that students relate fairness to domains such as assessment and grades (e.g., Chory et al., 2017; Čiuladienė & Račelytė, 2016), punishment (e.g., Bempechat et al., 2013; Chory et al., 2017), make-up/late policies (e.g., Chory et al., 2017; Horan et al., 2010), scheduling/workload (e.g., Chory et al., 2017; Horan et al., 2010), information for exams (e.g., Chory et al., 2017; Horan et al., 2010), feedback (e.g., Chory et al., 2017; Horan et al., 2010), respectful behavior (e.g., Buttner, 2004; Israelashvili, 1997), and professor conduct in interpersonal interactions (Bempechat et al., 2013; Houston & Bettencourt, 1999). A few studies
have also showed the impact of teacher fairness on student affective (e.g., happiness, anger, trust), and behavioral reactions (e.g., dissent to the teacher, inaction) (Horan et al., 2010).

Quantitative studies to date have been largely conducted within education literature (e.g., Kazemi, 2016; Peter & Dalbert, 2010). A recent systematic review study has identified 96 studies that aimed to measure students’ and teachers’ perceptions of fairness in classrooms (Rasooli et al., under review). This study showed that survey-based measurement is the key dominant quantitative method in previous classroom-based research on fairness. These 96 studies were identified to be conducted in various cultural contexts, with the majority of them being carried out in the US. Figure 1 shows the geographic map of the quantitative studies on fairness in classroom-based research. This study showed that 83 studies out of 96 have adopted organizational and social psychology theories of justice, attesting to the significance of this theory in empirical research on perception of justice. This study also systematically evaluated the psychometric properties of the quantitative surveys. This evaluation showed that most studies have used coefficient alpha to build the internal consistency evidence; however, only a few followed best practice procedures and conducted rigorous statistical analyses to support their validity. These findings point to the need to develop a theory of fairness and justice rooted in classroom contexts and build on this theory to measure fairness and justice adequately and appropriately within classrooms.
Figure 1. The geographic map of the quantitative studies on fairness in classroom-based research.

Note. The total is 102, which is more than 96, because a few studies were cross-cultural.

2.4 Toward Conceptual Frameworks for Assessment Fairness

In this section, I aim first to critically review the key conceptions that have been used to address fairness in the assessment literature and propose how these conceptions can be enriched by drawing on the wealth of justice literature in social sciences and humanities and education literature. This discussion helps the reader to appreciate the particular approach (i.e., social psychology) that this study will leverage to contribute to the conceptual and empirical investigations of fairness in classroom assessment.

2.4.1 Critical Review of Key Conceptions of Fairness in Assessment
As reviewed, the epistemology of fairness has been dominated by psychometrics (Nisbet & Shaw, 2019). Psychometrics deals with measurement of psychological attributes. To ensure that the measurement of psychological attributes is consistent, adequate, and appropriate, concepts of validity and reliability have been introduced and extensively conceptualized in educational and psychological testing (See Brennan, 2001; Haertel & Herman, 2005; Messick, 1989; Newton & Shaw, 2014; different editions of Standards for Educational and Psychological Testing). However, fairness has only recently received scholarly attention to its conceptual bases within educational assessment contexts. This attention to fairness as a criterion for assessment quality alongside validity and reliability is a key distinction between educational and psychological testing and assessment. In psychology, fairness does not play a role as a quality criterion in the measurement of psychological attributes. For example, in the measurement of affective disorders and depression among many other psychological attributes, how fairness can be a criterion to ensure the quality of the measurement instruments. However, fairness has been argued to be a criterion to gauge the quality of assessment in educational contexts because educational assessments are often leveraged for decision making about resource distributions and it is imperative to ensure that assessments are providing fair conditions so that everyone can demonstrate their learning and finally receive their due benefits and burdens.

The inattention to fairness in psychological testing vis-à-vis validity and reliability shows that unlike validity and reliability that are psychometric issues (Tierney, 2016), fairness is a moral virtue (Carr, 2018) and a socio-cultural value (Stobart, 2005). This epistemological nature of fairness can also be found in philosophical, legal, and social psychological perspectives that have attempted to push the margins of psychometrics to conceptualize fairness as a socio-cultural issue. The review of the literature of justice in social sciences and humanities and education in the first two sections also attest to this socio-cultural and moral nature of fairness.
Given that assessment is epistemologically dominated by psychometrics, it is expected that fairness in assessment would also be viewed dominantly through psychometrics. As viewed through positivism and realism, if we believe that psychometrics is an epistemology for understanding learning, then its main question is to understand what learning is through the quantification procedure. In this case, validity ensures that the assessment tool accurately measures what it claims to measure (i.e., learning) and reliability ensures that the assessment scores are consistent. How then would fairness be conceptualized in this process of understanding learning? Fairness cannot help us to understand learning or check the quality of this understanding. Fairness is a social issue and cannot be a criterion to gauge the quality of a test.

As viewed from pragmatism, psychometrics is defined based on its utility and outcomes. Validity would then look at the adequacy and appropriacy of inferences and actions based on test scores (Kane, 2006; Messick, 1989). From this perspective, validity not only includes internal test design, but more distinctly and importantly, consequential validity that evaluates the appropriateness of the uses and consequences of a test (Messick, 1989). It is within this epistemology of psychometrics that fairness is paid more attention. Inspired by Messick, McNamara and Ryan (2011) define fairness as a measurement quality of the test and justice as social consequences of a test. It is also within this epistemology that scholars have probed into questions like how fairness is different from validity (A. Davies, 2010; Kane, 2010; Kunnan, 2010; Xi, 2010) and influential remarks and beliefs that fairness is fundamentally a validity issue (Standards for Educational and Psychological Testing, 2014). It is also within this epistemology that scholars have contemplated that validity plays all the roles of fairness and have questioned whether validity should almost encompass everything that is good in assessment and whether the pursuit of fairness in assessment is in vain (Davies, 2010; Karami, 2013; Nisbet, 2017).

Consider Figure 2 as an illustration to describe the relationship between this epistemology of psychometrics (i.e., pragmatics), validity, and fairness. This figure starts from
the need for distribution of sources that is decided based on the measurement of learning and ensuring the quality of the measurement through validity evidence. An example can help elucidate Figure 2. There is a need in a society to have a standardized test as an entrance exam for a university admission because test-takers outnumber the university seats (i.e., distribution of resources). Therefore, a number of knowledge areas are deemed necessary to give admission to test-takers (i.e., learning). A test is constructed and used to measure the knowledge areas because measurement and testing are used as processes to understand and quantify learning (i.e., psychometrics and testing). Validity is then used to infer the appropriateness and adequacy of the inferences and actions based on test scores (i.e., validity inferences). Based on the test results, we then distribute the resources and decide who can receive university seats.

Figure 2. The link between pragmatics epistemology of psychometrics, validity, and fairness

Based on this example, psychometrics and testing are used to respond to a societal need and are evaluated based on the utility of the decisions (Maul et al., 2016). Validity not only examines the adequacy of the test’s internal structure but also the appropriacy of the inferences and actions from test scores to its uses. Fairness would then be defined as the consideration of test-takers and ruling out the construct-irrelevant factors due to social markers such as culture,
gender, race, and disability (Standards for Educational and Psychological Testing, 2014). There are two epistemological misunderstandings here. First, fairness is a social concept and aims to understand whether distribution of resources is based on merit in this context. It does not align with the epistemology of fairness to examine the quality of measurement, tests, and learning. Rather, fairness can examine the legitimacy of assessment as a tool for distribution of resources (e.g., critical perspective). Second, it appears that validity has been conceptualized out of its epistemological home. If psychometrics is a process for understanding learning; validity is a criterion to check the quality of measurement. Enlarging validity to include appropriateness of actions (i.e., distribution of resources) based on test results are out of the purview of measurement and loads validity with a sociological burden that has not originally been conceptualized to deal with. Appropriateness of actions (i.e., distribution of resources) based on test scores require philosophical, sociological, psychological, and economic arguments about fairness that fall out of the scope of current measurement conceptions. While fairness and validity in assessments are fundamentally different in their epistemological and functional bases, the validity of assessment is a necessary but not sufficient basis for fairness of the uses of assessment results.

Consider another example from the area of assessment accommodations to delineate this relationship. A math test is administered to Grade 12 students in Ontario, Canada. As some of the test-takers have been identified as students with disability and English language learners, we need to provide accommodations. Accommodations are given to provide students with disability and English language learners with an equal opportunity to show their learning (in comparison with their peers). Assessment accommodations are changes to assessment content, format, and administration that do not change the intended construct but remove construct-irrelevant barriers that prevent students with disability or English language learners from showing what they know (Standards for Educational and Psychological Testing, 2014). Based on this definition, validity looks at any construct-irrelevant factors that might undermine the adequateness and
appropriateness of inferences (and actions) based on test scores (Messick, 1989). The adequacy and appropriacy of inferences can be analyzed based on comparing the test scores of students with disability and English language learners with typical test-takers. This analysis feeds back to the validity, showing whether there is an error in our understanding of the learning; however, if we use this analysis to examine actions (i.e., distributions of resources or recognitions of a student standing), it seems that we are stepping out of the purview of validity as a quality criterion for measurement process because measurement process is interested in understanding learning rather than distribution of resources. However, we can argue based on the results of this analysis that actions will tend to be unfair in terms of resource distribution and recognition of standing for students with disability and English language learners. These two outcomes are social outcomes and therefore outside of the scope of educational assessment as understood within the epistemology of psychometrics unless we would like to broaden psychometrics from a process to understand learning to include a process to understand the appropriateness of actions and distribution of resources based on test results.

Let’s not forget that tests are used in a society for distribution of resources as well as recognition of an individual’s standing, therefore, when someone says that a test is unfair, they mean the outcomes of that test are unfair because they could not get the result (i.e., grade as well as uses of that grade for various purposes) they deserve. However, we cannot make a test fair because epistemologically, fairness is a social issue and is of no help to gauge the quality of the technicalities of a test design for which validity and reliability have been proposed.

The conceptual analysis of fairness in assessment has not been paid adequate attention until recently (Nisbet, 2017) and the key reason for this inattention can be explained by the efforts of educational assessment scholars to conceptualize a concept that is social in nature and non-psychometric based on the dominant epistemology of psychometrics. This conclusion is also supported by a systematic review of validity studies, where it was found that no study had
reported consequential validity evidence (i.e., validity for test use) (Cizek, Bowen, & Church, 2010; Cizek, Rosenberg, & Koons, 2008). Cizek (2016) argues that this lack of research on consequential validity, which touches upon ethics and fairness, is due to inherent internal contradiction in the definition of validity that aims to integrate validation of test score and justification for test use under the umbrella of validity. Similarly, Borsboom and Wijsen (2016) argued that science (i.e., validity) and ethics (i.e., values) are two arenas that should be kept separate. While Cizek and Borsboom have concluded that ethics and fairness are separate from validity through analysis of validity literature and philosophy of science, I have reached this conclusion through analysis of epistemology of psychometrics and fairness. These conclusions push us to ask a question whether psychometrics is still an appropriate epistemological basis for conceptualizing fairness in assessment, particularly for classroom assessment contexts? This question is very pertinent because inspired by the dominant epistemological basis of educational assessment in psychometrics, classroom assessment scholars have attempted to conceptualize fairness using psychometric principles (Camilli, 2013; Herman & Cook, 2019).

2.4.2 Theoretical Bases to Reconceptualize Fairness in Classroom Assessment

Consistent with the reviewed literature of fairness in assessment, social psychology and moral and political philosophy theories can be used to theorize fairness in assessment contexts. Social psychology probes into underpinning factors that teachers and students use to perceive fairness in assessment as well as the psychological, social, and moral responses that they show in response to un/fairness. Political and moral philosophy theories challenge the so-called objectivity in assessment and call attention to the misuse of assessment as a structural tool for social control (i.e., critical perspectives). These theories also advocate for studying teachers’, students’, and test-takers’ perspectives as a basis to address fairness in assessment (Kunnan, 2018; Tierney, 2014). These perspectives can be enriched and augmented by drawing on justice literature in social sciences and humanities and education literature.
As related to social psychology, we can build on sociological and psychological underpinnings of fairness theory in social sciences to theorize what psychological and social mechanisms are at play in constructing a fairer culture of assessment. Specifically, we can delve deeper into understanding how students and teachers conceptualize fairness in assessment and how this conception interacts with the social context they live in. For example, we can understand how social demography of a school in a rural area can impact teachers’ and students’ conceptions and enactment of fairness in assessment. The social psychology perspective can also be integrated with critical theory to study the biases that may impact classroom and large-scale assessment. Despite that both classroom summative and large-scale assessments value objectivity (i.e., grading only based on achievement), they both seem to be influenced by social and psychological biases. For example, Artiles (2019) shows how assessment process blinds itself to the issues of race, disability, color, and culture and their intersection and how this blinding has resulted in disproportionate identification of students of color as disabled compared with other cultural groups within inclusive education systems. In classroom assessment, while the positive aspects of involving teachers and students in assessment of achievement is well-acknowledged, little emphasis has been placed on acknowledging the cultural, disability, and socio-economic biases that teachers may have in classroom assessments. The 50-year research on teacher expectations can be particularly useful to further examine bias in teacher assessment judgments (Good, Sterzinger, & Lavigne, 2018; Wang, Rubie-Davies, & Meissel, 2018).

As another example, accommodations are highlighted in assessment literature as mechanisms to give equal and fair opportunity to English language learners and students with disability to show their learning (Abedi et al., 2004; M. Davies, Elliott, & Cumming, 2016; Yang, 2019). Accommodations are actually the response of assessment field to justice for inclusive education in an accountability system of education. Despite its justice spirit, accommodations in the assessment literature have largely been discussed in relation to practices that teachers can use
to give English language learners and students with disability equal opportunity. To augment the spirit of justice in accommodation research, it is also imperative to study teachers’ beliefs and potential biases toward English language learners and students with disability as this will impact the ultimate goal of accommodations as to empower these students to show what they know. As Resh and Erhard (2002) reported in their study of Israeli school counselors’ guidance to high school students on track placement, low-ability students dominantly from low socio-economic backgrounds received significant “cooling-out” messages to continue with a less demanding easier educational routes. While this finding is in the context of school counseling, it envisions how social and psychological biases can impact teachers’ attitudes, assessments, and accommodations toward students from disadvantaged backgrounds.

Despite the promising foundation of social psychology to theorize fairness in classroom assessment contexts, it is important to note that social psychology as a psychological field epistemologically adheres to a variant of positivism or of realism and studies its objects of inquiry (i.e., constructs) dominantly through quantitative methodology. This is unsurprising as social psychology is more a psychological programme of inquiry that has historically used measurement procedures to research psychological constructs including justice (See Michell, 1999 for the history of measurement and psychology; See Journal of Personality and Social Psychology; Social Psychology of Education as evidence for positivistic worldview and quantitative research tradition). The review of social psychological approach to the study of justice in education also attests to this quantitative and positivistic psychological research tradition (Rasooli et al., under review). Nonetheless, I argue that conceptualization of fairness in classroom assessment through social psychology does not need to limit itself to this positivistic and quantitative tradition that is the case in the field of psychology. Classroom assessment scholars can leverage qualitative methods in addition to quantitative methods to study fairness. Recently, the use of qualitative approach has also been promoted to theorize justice for workplace contexts (Cropanzano, Fortin,
Classroom assessment scholars can also balance the psychological orientation with the sociological perspective of justice to examine fairness more fully. For example, it is not only productive to study teachers’ and students’ conceptions of fairness in classroom assessment (i.e., psychology) but also how these conceptions are influenced by the culture and socio-economic demography of a school (i.e., sociology). To answer such questions, it is necessary to use not only quantitative but also qualitative perspectives.

In addition to social psychology, assessment literature can also draw on political and moral philosophy perspectives such as virtue ethics theory, critical perspectives, deontology, utilitarianism, and feminist theory to add to the theoretical foundations of exploring fairness in classroom assessment. The previous use of moral and political philosophies to conceptualize fairness in assessment (Kunnan, 2018; Tierney, 2014) has proved useful, which shows the promising route for further conceptualizations of fairness in assessment.

All in all, conceptualizing fairness in classroom assessment by drawing on the theories of justice in social sciences and humanities and education extends the epistemology of classroom assessment to move toward justice. The field of classroom assessment has largely been conceptualized in relation to learning as a key goal of education (Black & Wiliam, 1998; Brookhart, 2003, 2004, 2018; Cowie & Bell, 1999; Elwood & Murphy, 2015; Lorrie Shepard, 2001). However, we have yet to conceptualize classroom assessment from justice perspective as another key goal of education (Gorard & Smith, 2010). The interest in fairness in classroom assessment seems plausible as far as it relates to and contributes to learning. However, it seems that this perspective is limited to address the significance of justice as an end goal for an educational system. The conceptualization of justice not only as a means to learning but also as an end in itself aligns with the socio-cultural theory of assessment that aims to conceptualize assessment as attentive to social, cultural, and economic milieu within which it happens (Elwood
& Murphy, 2015). Therefore, this study leverages fairness as a basis to potentially reconceptualize and broaden epistemology of classroom assessment.

In this study, I will leverage social psychology theory to conceptualize fairness in classroom assessment. Through my analysis of fairness conceptions across disciplines, I assert that classroom assessment can provide a unique context for conceptualizing fairness as it happens in the intersection of classrooms, assessment, instruction, curriculum, policy, standardized assessments, and diversity of teachers, students, and parents. Within this socio-psychological context of classroom assessment, fairness can be conceptualized using social psychology theory of justice. This theory considers how teachers and students conceptualize fairness and how their conceptions influence their practices and perceptions (i.e., distributions, procedures, and interactions) in assessment, instruction, and interpersonal relationships in classrooms. This theory also takes into account the psychological, behavioral, and social outcomes that teachers’ and students’ conceptions of fairness can bring about for students, classrooms, schools, and society at large. This theory can be leveraged to study how teachers conceptualize and enact fairness in assessment practices because teachers’ assessment practices are deeply entangled with what they believe is fair. Without understanding teachers’ fairness beliefs, it is very hard to move toward changing their assessment practices. Further, this theory can be used to educate teacher candidates about steps that they can prioritize to ensure fairness in their assessment practices, especially because teachers are confronted with various contextual factors including classroom size, time, policy, and standardized testing that may conflict with their own fairness beliefs. This theory can be leveraged to appreciate how students conceptualize fairness in assessment, instruction, and interpersonal relationships and how their understanding influence their moral, psychological, and social outcomes in classrooms and schools. This theory can also be used to understand how we can address assessment fairness in its multiple forms in multicultural societies with teachers and students from diverse backgrounds. All these efforts in addition to many others
will provide a richer and suitable framework to conceptualize fairness in classroom assessment contexts and contribute to fairer assessments, classrooms, schools, and societies.

2.5 Summary

Inspired by the recent 21st century social and educational movements towards equity, diversity, and inclusion for disadvantaged groups, educational researchers have persisted in conceptualizing fairness in educational and classroom assessment contexts. These efforts have provoked promising key theoretical foundations and empirical investigations to examine fairness in educational and classroom assessment. This review has aimed to critically review these theoretical foundations and associated empirical studies to examine their potential for addressing the complex and evolving notions of fairness in classroom assessment contexts. This review also built on fairness and justice literature in social sciences and broader educational discourses to provide additional theoretical grounds to rethink fairness in classroom assessment.

Four key perspectives—psychometric perspective, legal perspective, philosophical perspectives, and social psychological perspective—have been used to conceptualize fairness in educational assessment. While the psychometric perspective has been the dominant perspective to conceptualize fairness within classroom assessment, the legal perspective has not been applied into the classroom assessment given the low stakes of assessments in classrooms. The additional moral, political, critical, and social psychological perspectives have been recently used to conceptualize and empirically investigate fairness in classroom assessment contexts. Future research needs to further draw on these additional theories to evaluate their use and comprehensiveness for interpreting fairness within diverse classroom assessment contexts. Of relevance to the focus of this dissertation study, social psychology theory was used to investigate first-year undergraduate students’ perceptions of fairness in classroom assessment in relation to their secondary schools in Ontario, Canada.
Five key conceptions of justice in sociology, psychology, economics, philosophy and education have been identified and reviewed as potential bases to enrich the existing perspectives of fairness in classroom assessment. Justice in sociology considers the social mechanisms and conditions that shape individuals’ perception of fairness and justice in a particular spatio-temporal context. Justice in psychology considers the individual, interpersonal, and intergroup levels according to which individuals arrive at perception of fairness and justice and react to enforce and restore fairness and justice. Justice in economics uses ultimatum and dictator games to investigate the drivers of two individuals’ (i.e., proposer and recipient) decisions in distributing a resource. Specifically, it aims to investigate the underpinning principles that drives the share amount of a resource that a proposer in a game is ready to offer given the knowledge that the proposer and the recipient would none receive any of the resources if the recipient repudiates the offer. Justice conceptions in psychology, sociology, and economics provide principles and conditions that can be potentially used to shed light on interpreting empirical data in classroom assessment fairness.

Justice in philosophy and educational literature builds on the key perspectives of distributive justice and relational justice perspectives. Distributive justice is concerned with providing equal opportunities for everyone to have the potential to receive the resources. If there is a distributive injustice in the society, the resources should be redistributed to the most benefit of the worst-off. The relational justice considers redistribution of resources as an inadequate response specifically if the injustices are caused by the existing underlying structural inequities that need to be transformed to combat oppression, marginalization, exploitation, powerlessness, cultural imperialism, and violence. Building on distributive, socially just pedagogies and group perspectives in education argue for redistribution of resources to the benefit of disadvantaged student groups due to social markers such as disability, socio-economic, cultural and language background, gender, and ethnicity. Based on relational justice, socially just pedagogies and group perspectives in education contend for cultivating critical capacities in teachers and students to
critically examine the current injustices in the society and education and develop their capacities
to transform the underlying structural injustice for a fairer and more just school societies. The
distributive and relational justice perspective enable a classroom assessment researcher to ask
how assessments can be made fairer for disadvantaged students such as students from low socio-
economic backgrounds and refugee backgrounds.

Overall, this review critically reviewed the current conceptions of fairness in educational
and classroom assessment and expanded these conceptions using justice theories in social
sciences. To that end, this review was able to contribute conceptual grounds for future theory-
driven empirical research to advance fair assessment practices in classrooms.
Chapter 3
Methodology

This study leveraged a mixed methods approach to investigate Ontario-based first-year undergraduate students’ perceptions of classroom assessment fairness during their secondary school experiences as well as their associated psychological and behavioral responses to un/fairness experiences. To that end, this study is guided by the following research questions:

1. How do first-year undergraduate students perceive fairness in classroom assessment in relation to their secondary school experiences in Ontario, Canada?
2. What psychological and behavioral responses are provoked by the perceptions of fairness in classroom assessment contexts?

This chapter first provides information about the contextual significance and ethical adherence for this research. It then delineates the two-phases of the mixed methods research design and associated data collection and analysis strategies that were used in each phase.

3.1 Context

This research was conducted in Ontario, Canada. Educational schooling system in Canada is governed under the jurisdiction of provinces, with ten provinces and three territories exercising their independent educational policies. Despite this jurisdictional authority, educational policies including assessment policies share a remarkable similarity across the jurisdictions (Jang & Sinclair, 2018; Volante & Ben Jaafar, 2008). Of relevance to the specific context of this study, the assessment policy of Growing Success: Assessment and Reporting in Ontario Schools (Ministry of Education, 2010) outlines seven fundamental principles that guide the assessment practice in Ontario K-12 schooling system. These principles ensure that teachers use various assessment processes that enjoy validity and reliability, contribute to student learning, and document their achievement. Particularly, it promotes assessment practices that:
• are fair, transparent, and equitable for all students;
• support all students, including those with special education needs, those who are learning the language of instruction (English or French), and those who are First Nation, Métis, or Inuit;
• are carefully planned to relate to the curriculum expectations and learning goals and, as much as possible, to the interests, learning styles and preferences, needs, and experiences of all students;
• are communicated clearly to students and parents at the beginning of the school year or course and at other appropriate points throughout the school year or course;
• are ongoing, varied in nature, and administered over a period of time to provide multiple opportunities for students to demonstrate the full range of their learning;
• provide ongoing descriptive feedback that is clear, specific, meaningful, and timely to support improved learning and achievement;
• develop students’ self-assessment skills to enable them to assess their own learning, set specific goals, and plan next steps for their learning (p. 6).

These principles explicitly or implicitly argue for fair classroom assessment practices that are transparent and equitable for all students including students with disability, English language learners, and students from First Nation, Métis, or Inuit background. These principles also support providing multiple assessment opportunities predicated on student experiences, needs, and preferences and endorse transparent, clear, and timely communication of assessment process, feedback, and results to students and their parents. These principles put a large emphasis on fair classroom assessment practices within Ontario schools in Canada and support the need for conceptualizing fairness for classroom assessment practices in Ontario schools. The results of this research will provide students’ views of fairness to inform the conceptualization and practice of fair assessments in Ontario schools and in relation to this provincial assessment policy context.
3.2 Research Ethics

Ethical clearance for this research was sought from Education Research Ethics Board and General Research Ethics Board (GREB) at Queen’s University prior to beginning this study. The ethical clearance for this study from GREB was provided in Appendix A, followed by two amendments associated with this ethical clearance (Appendix B and C). The combined letters of information and consent forms for study participants in both qualitative and quantitative phases can be found in Appendices D to G. All participants were provided with these letters and were allowed participation upon their consent.

3.3 Mixed Methods Research Design

A sequential mixed methods approach including two phases with the qualitative phase preceding the quantitative phase was used (Creswell & Clark, 2017). The mixed methods approach was selected because it allows to examine classroom assessment fairness topic from multiple dimensions using integrated evidence from qualitative and quantitative approaches (Creswell & Clark, 2017). In phase I, a qualitative interview method was used because it is a recommended method in the fairness literature to gain deeper perspectives into students’ perceived fairness (Bempechat et al., 2013; Horan et al., 2010). The qualitative phase enabled exploring students’ perceived fairness in classroom assessment and associated psychological and behavioral consequences. Further, the results from the qualitative phase provided a foundation for building the construct of fairness in classroom assessment as interpreted and articulated by the first-year undergraduate students reflecting on their secondary school experiences in Ontario, Canada. As research on social psychology interprets fairness as socially constructed based on (shared) subjectivity (Cropanzano & Greenberg, 1997; Folger & Cropanzano, 1998; Resh & Sabbagh, 2016), it was imperative to build the construct of fairness in classroom assessment based on students’ (shared) subjective perceptions of fairness. The qualitative phase also
contributed to address the geographic gaps in fairness of classroom assessment in Canada and adds to the nascent qualitative evidence on fairness in classroom-based research internationally (Tierney, 2013; Rasooli et al., 2018).

In phase II, a scenario-based instrument was constructed. The derived themes from Phase I analysis of students’ qualitative interview data provided the underpinning domains of assessment (e.g., groupwork) and the principles of justice from social psychology theory (e.g., equity) that students used to perceive fairness in classroom assessment in relation to their secondary school experiences. These domains and pertaining social psychology justice principles were leveraged to build an instrument in the quantitative phase to expand the scope of the inquiry into students’ perceptions of fairness in classroom assessment in Ontario, Canada. The quantitative phase provided a unique instrument with initial evidence of psychometric properties that can be potentially used to provide empirical evidence to inform theory, policy, and practice of fairness in classroom assessment in Canada and internationally.

All in all, both phases guided this study by examining how secondary school students in Ontario, Canada conceptualize fairness in classroom assessment. The following Figure 3 presents a schematic overview of Phase I and Phase II data collection and analysis procedures.
3.4 Phase I

Phase I provided evidence to answer how first-year undergraduate students perceived fairness in classroom assessment in relation to secondary schools in Ontario, Canada as well as what psychological and behavioral responses they reported in response to their un/fairness experiences.

3.4.1 Recruitment

Due to Covid-19 pandemic, the recruitment of participants for this study was conducted via online social media platforms. Recruitment scripts (see Appendix J) were distributed in social media networks including Facebook and Twitter, targeting pages for freshmen (or first year) undergraduate students in various universities and colleges in Ontario. These scripts invited first-year undergraduate students for a thirty-minute interview via the Microsoft Teams platform. Interested participants reached out to the researcher via an email. The researcher then shared
consent forms (see Appendix D) and potential dates for an online interview. Upon the participants’ consents, the interviews were scheduled in August 2020. In total, 27 participants responded to the recruitment scripts and were selected from first-year undergraduate students who had completed their secondary schooling in Ontario, Canada to participate in this study.

3.4.2 Participants

59% of participants were female and 41% were male, with the mean age of 17.8 years. 18% of participants identified themselves as first generation immigrants, meaning that they and their parents moved to Canada from another country. 44% of participants identified themselves as second generation immigrants, meaning that their parents moved to Canada but the students themselves were born in Canada. 37% of participants had parents and were themselves born and raised in Canada. Participants also identified their race as black (11%), Caucasian and White (33%), East Asian and Chinese (22%), South Asian (25%), Middle Eastern (3%), and Mixed (3%). Participants reported that 74% attended public school, with 26% attending Catholic school. Most students were not willing to mention the secondary schools they had attended due to confidentiality; however, they reported in which city their schools were located. 12 students did in Toronto, 4 students in Mississauga, 2 in Thornhill and one student from each of the following cities: Whitby, Brampton, Waterloo, Peterborough, Ottawa, London, and Markham. Two students also attended secondary schools within Toronto District School Board. The participants were going to begin or were doing their undergraduate studies in the following universities: 6 at University of Toronto, 6 at Waterloo University, 5 at Queen’s University, 2 at Wilfrid Laurier University, 2 at Western University, one at each Brock University, University of Ottawa, Ryerson University, and University of Pennsylvania. Two participants did not report the university. Three of the participants also reported that they received Individualized Education Program during high school. More detailed information about the student demography in Phase I is presented in Table 2.
Table 2. Overview of phase I student participants’ demography (n=27)

<table>
<thead>
<tr>
<th>Gender</th>
<th>N (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>16 (59.3%)</td>
</tr>
<tr>
<td>Male</td>
<td>11 (40.7%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>N (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 years</td>
<td>11 (40.7%)</td>
</tr>
<tr>
<td>18 years</td>
<td>11 (40.7%)</td>
</tr>
<tr>
<td>19 years</td>
<td>4 (14.8%)</td>
</tr>
<tr>
<td>20 years</td>
<td>1 (3.8%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Immigration Background</th>
<th>N (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Generation (parents and student moved to Canada)</td>
<td>5 (18.5%)</td>
</tr>
<tr>
<td>Second Generation (parents moved to Canada)</td>
<td>12 (44.4%)</td>
</tr>
<tr>
<td>Parents and the student born and raised in Canada</td>
<td>10 (37.1%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Racial Background</th>
<th>N (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>3 (11.1%)</td>
</tr>
<tr>
<td>Caucasian (White)</td>
<td>9 (33.2%)</td>
</tr>
<tr>
<td>East Asian and Chinese</td>
<td>6 (22.2%)</td>
</tr>
<tr>
<td>South Asian</td>
<td>7 (25.9%)</td>
</tr>
<tr>
<td>Middle East</td>
<td>1 (3.8%)</td>
</tr>
<tr>
<td>Mixed</td>
<td>1 (3.8%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>School Background</th>
<th>N (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>20 (74%)</td>
</tr>
<tr>
<td>Catholic</td>
<td>7 (26%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Students with Disability</th>
<th>N (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individualized Education Plan</td>
<td>3 (11%)</td>
</tr>
</tbody>
</table>

3.4.3 Data Collection

Semi-structured interviews were conducted with student participants to explore their perceptions and experiences of fairness in classroom assessment including psychological and behavioral consequences of their fairness perceptions in relation to their secondary school experiences in Ontario, Canada. Interviews were used as they:

- provide opportunities for mutual discovery, understanding, reflection, and explanation via a path that is organic, adaptive, and oftentimes energizing. Interviews elucidate
subjectively lived experiences and viewpoints from the respondents’ perspective…

interviews are as much about rhetorically constructing meaning and mutually creating a story as they are about mining data gems. Meaning is created between participants and their material surroundings rather than being held in the minds of the interviewer or interviewee and swapped back and forth; indeed, interviews are not neutral exchanges of questions and answers, but active embodied processes in which we co-create stories, and come to know others, ourselves, and the world (Tracy, 2020, p. 156).

Interviews provided an opportunity to learn from participants of this study about how they interpreted fairness in classroom assessment contexts. The interview questions were crafted inspired by the previous qualitative studies exploring students’ open-ended responses to their experiences of fairness in classrooms and assessments (Horan et al., 2010; Rasooli et al., 2019). The interview questions aimed to elicit students’ experiences and perceptions of un/fairness and associated psychological and behavioral reactions in relation to their secondary school experiences in an aim to provide pertinent responses to the research questions of this study. Specifically, the interview questions began with asking a few questions about students’ background information and their high schools, followed by open-ended questions on their experiences of un/fairness in classroom assessment. Additional questions were followed to provoke students’ memories of their additional experiences of un/fairness in relation to classroom assessment, teaching, and interactions as well as their feelings and behaviors after these un/fair experiences. Appendix H provides an overview of the potential questions that were asked in the interviews. These questions provided the basis to tap into central areas of concerns pertaining to fairness in students’ perceptions as shown by prior research (Horan et al., 2010; Rasooli et al., 2019). The interview questions were pilot tested with one student to ensure that they sufficiently provoked diverse arrays of un/fairness experiences in relation to assessments in secondary schools. These questions guided the interviews in this study, but with the latitude to bend toward
emerging interpretations that the participants reveal about their un/fairness experiences. In this way, interviews in this study provided an opportunity for the participants and the researcher to enter a mutual discovery of what constitutes fairness in classroom assessment.

Specifically, in the interviews, participants were asked demographic questions about their schools and their teachers and were inquired about their immediate memories and experiences of assessment un/fairness in their secondary schools as well as their psychological, social, and behavioral responses to these un/fair experiences. The participants were asked about both their fairness and unfairness experiences based on their recollection of a fair or an unfair experience(s). Once the participants articulated their immediate memories and experiences, they were provided chance to share additional experiences. If they could not remember any specific experiences, multiple domains of assessment (e.g., grading, feedback, groupwork) were mentioned to provoke whether they had relevant experiences of un/fairness. These questions were continued based on the participants’ responses as well as the interview protocol to elicit their perceptions of fairness in classroom assessment. In total, each interview ranged in length from 20-45 minutes. All the interviews were conducted by the researcher (me) via the Microsoft Teams and were recorded with the participants video on or off based on their choice. After the completion of the interviews, each participant received 10$ gift card as a compensation for their contributions. All the interviews were conducted during August 2020.

3.4.4 Data Analysis

All the data from interviews were transcribed verbatim and were thematically analyzed (Patton, 2015; Tracy, 2020). The researcher read the entire data to consolidate his understanding of students’ fairness accounts in classroom assessment as grounded in the data. In the second reading, the researcher began coding the data to find patterns significant and relevant to the focus of this study. The perusals of the data showed that students’ perceptions of fairness in classroom assessment were related to diverse assessment domains such as grading, feedback, groupwork,
and exam. Therefore, the unit of analysis focused on domains of assessment (e.g., cheating, group work) in which students used various justice principles in social psychology theory (e.g., equity, transparency) to arrive at perceived fairness. Prior qualitative study on students’ perceived fairness in classroom assessment also recommended this analysis approach by identifying assessment domains and then examining what principles of justice in social psychology theory (e.g., equity, transparency) were used by students to evaluate the fairness of these domains (Rasooli et al., 2019). This analysis approach helps better appreciate what principles of justice are more pronounced and relevant to students’ perceived fairness in each domain of assessment, contributing to informing measurement instrument construction and teachers’ assessment fairness practices.

The domain-specific assessment experiences were analyzed to identify the underpinning justice principles as well as contextual factors (e.g., student personal background) that had led students to perceive un/fairness in classroom assessment. The data were iteratively coded to categorize them into the relevant assessment domain (e.g., grading) and the justice principles (e.g., consistency in implementing grading criteria) and contextual factors (e.g., inclusion of student behavior in grades). The assessment domains were labeled ‘themes’ in this study with associated codes, representing justice principles and contextual influences. In total, 7 themes (except for responses to un/fairness theme) and associated 31 codes were derived, with frequency of 257 examples representing these 7 themes. Table 3 present the themes, codes, examples, and frequencies for codes and themes. The codes for justice principles and contextual factors used within associated domains (i.e., themes) were tallied in students’ articulated experiences to arrive at the frequency. Thus, the code frequencies represent the number of each justice principle and contextual influence students used in their articulation of experiences to evaluate fairness. The theme frequencies are an aggregate of all codes used to evaluate fairness in each theme. In this
way, Table 3 also characterizes the relevant codes and their weightings in students’ evaluation of fairness in each theme in the context of this study.

The data were also analyzed to identify the psychological and behavioral responses students showed in relation to un/fair experiences. The researcher read the data to identify students’ responses and coded the specific psychological and behavioral responses they reported to show in reaction to un/fairness experiences. Table 3 presents 6 codes with a frequency of 50 examples that represented students’ responses to un/fair experiences. The frequencies were calculated by tallying students’ psychological and behavioral reactions in the entire data.

Once the thematic analysis was finalized, an additional reader was asked to analyze the 10% of the data to establish confidence in coding, mapping data examples to themes. The reader expressed overall inter-rater agreement with the thematic analysis and reported a few comments that were incorporated to enhance the trustworthiness of data analysis. This iterative process provided a deep common understanding of the data and data analysis procedure.

Table 3. Summary of themes, codes, examples, and frequencies

<table>
<thead>
<tr>
<th>Themes</th>
<th>Codes</th>
<th>Code</th>
<th>Examples</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Perceptions of Fairness</td>
<td>Equity</td>
<td>8</td>
<td><em>Everyone is treated equally and has the right to speak respectfully.</em> S27</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Equality</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Respect</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairness in Groupwork</td>
<td>Group Composition</td>
<td>14</td>
<td><em>It depends really, because students should be given the opportunity to sort of work with people that they think they could do well with because obviously, you know, school is just like it's a path to your own future.</em> S18</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Group Dynamics</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grading</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairness in Exams</td>
<td>Reasonable</td>
<td>6</td>
<td><em>In chemistry, we'll be learning thermodynamics or something. And on the test, it will be something completely different from the content that we've been learning. Really let students know all work.</em> S26</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Scheduling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Student Voice in Exams</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transparency in Exam Information</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic</td>
<td>Count</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching and Exam Content Alignment</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missed Exam Policies</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transparency in Cheating Policy</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistency and Bias Suppression in Decision</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice Opportunity</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assigning Zero</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caring about Student Background</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Projecting</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consequences of Cheating Decisions</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utilitarian View</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity and Caring</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transparency in Criteria</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistency in Implementation</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias Suppression in Grading</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Voice Opportunity to Appeal</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adequate Justifications</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transparent Feedback</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timely Feedback</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adequate and Honest Feedback</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher-student Relationships</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender Equity</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racial Equity</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

You want students to know that it's wrong, but you also don't want to punish them so severely that their opportunities for the future for university or college to be limited over one mistake. S19

In Grade 10 math, Mr. Sam did things by the book. He was extremely fair and gave only marks according to set criteria. S8

Often the problem was more that we would complete assignments, let's say essays or tests, and we would give them back to the teacher to be marked. And we would only receive feedback several months later. So there was no opportunity for us to improve on our skills, because by that point, we would already be finished with that unit. S19

A lot of my bad experiences stem from teachers that didn't give me the same chance as they did for other students because the other students maybe had more
I got shafted by the teacher. You know what I mean? Because it's frustrating. You see the teacher being nicer to other people. And it's also harder to learn. S7

I would say it impacted my attitude towards the class greatly because it made me dislike the teacher and made me less wanted or less willing to participate in class because I didn't think that my opinion would be respected since he kind of showed that he didn't care about my opinion earlier. S17

3.5 Phase II

This phase leveraged quantitative data to probe deeper and more broadly into the principles of justice that students might use to arrive at perceptions of fairness in classroom assessment by engaging a large sample of students. This phase drew on the themes identified in Phase I (i.e., qualitative phase) to develop the Classroom Assessment Fairness Inventory. The themes in the qualitative phase (i.e., fairness in groupwork, exam, cheating, grading, and feedback) and associated justice principles that rendered students’ perceptions of fairness in classroom assessment were used to construct the inventory. In the subsequent subsections, the process for developing the Classroom Assessment Fairness Inventory is detailed.

3.5.1 Classroom Assessment Fairness Inventory Development Process

This study used multiple steps to develop the Classroom Assessment Fairness Inventory. Specifically, it first developed the underpinning theoretical construct and measurement theory for this inventory using (a) a systematic review of previous instruments and literature examining students’ perceived fairness, (b) social psychology theory, (c) and qualitative data from Phase I. It...
then collected evidence to support the validity for this inventory based on the recommendations from Standards for Educational and Psychological Testing (2014). From the five sources that the Standards recommend (i.e., content, response processes, internal structure, relationship to other variables, and consequences), this phase reports evidence based on content, internal structure, and relationship to other variables.

### 3.5.2 Theoretical and Measurement Underpinnings for Classroom Assessment Fairness Inventory

The purpose of Classroom Assessment Fairness Inventory is to measure students’ perceived fairness in classroom assessment. Given that previous literature includes numerous instruments for measuring students’ perceived fairness in classrooms, how can construction of Classroom Assessment Fairness Inventory be warranted for this study? There are multiple reasons that this study developed a new instrument rather than using previously developed scales. First, a recent systematic review study of 96 studies with instruments measuring students’ and teachers’ perceived fairness in classrooms showed weak conceptual and psychometric bases for these instruments and called for more rigorous instruments to inform evidence-based practice (Rasooli et al., under review). While some of these classroom-based instruments measured fairness in assessment (i.e., grading) as part of the classroom fairness construct, there was a need to develop an instrument particularly focusing on the domain of assessment as none of the scales focused exclusively on the assessment context. This need is real because teachers’ and students’ perceived fairness in classroom includes various dimensions that are not practically feasible to incorporate in a comprehensive instrument. Therefore, there is a need for an instrument that particularly focuses on fairness in classroom assessment and attempts to adequately measure the various relevant domains of fair classroom assessment with strong conceptual and psychometric bases.

Second, given the nascent focus on fairness in classroom assessment and a shortage of measurement instruments in this area, Sonnleitner and Kovacs (2020) recently developed a 17-
item Fairness Barometer to examine teachers’ and students’ perceived fairness. This instrument measures teachers’ and students’ perceived fairness in classroom assessment particularly in the areas of grading and exams. While this new instrument enriches the fairness literature in classroom assessment, it suffers from its narrow focus on grading and exam practices while students perceive fairness in classroom assessment in other significant areas that the previous literature and the results from Phase I in this study show such as group work, cheating, and feedback. The Fairness Barometer also measures fairness using ‘faceted justice’ and ‘overall fairness’ measurement approaches (Colquitt & Rodell, 2015). Faceted justice measures justice principles without direct reference to the term ‘fairness’ in the items. For example, Sonnleitner and Kovacs (2020) used items in the 10-point Likert scale such as “I understand my own grades on oral exams” to measure student agreement with the procedural justice principle of transparency. Overall fairness approach measures student global perception of fairness using one or two broad items. For example, Sonnleitner & Kovacs (2020) used an item to measure overall fairness (i.e., How fairly do you think performance is graded in this subject?). These types of items are useful, but they force a student to respond to items in relation to their classroom assessment experiences even if the student had not personally experienced such an un/fairness in classroom assessment.

The Classroom Assessment Fairness Inventory adopted a scenario format to redress these shortcomings. Scenario-based measurement provides a hypothetical context and asks participants to answer to the response option(s) based on their analysis of the hypothetical context. This approach to measuring fairness provides a better foundation to explore the underpinning justice principles that students use to perceive fairness in classroom assessment in an aim to inform fair assessment practices. A few studies have measured students’ perceptions of fairness using scenarios in the literature (Cherry, Ordóñez, & Gilliland, 2003; Duplaga & Astani, 2010; Pepper & Pathak, 2008; Schmidt, Houston, Bettencourt, & Boughton, 2003; Tata, 2005; Thorkildsen,
1989). These studies asked participants to either rate the fairness of a scenario by posing a few items in the Likert scale format as showed in Sonnleitner and Kovacs’ (2020) items or select from one of the proposed response options that is fairest in dealing with a situation such as student cheating (e.g., Duplaga & Astani, 2010). The Classroom Assessment Fairness Inventory augments these previous efforts by including relevant principles of justice that students reported to use to evaluate the fairness of a specific scenario domain (e.g., grading) in classroom assessment. Accordingly, this inventory included five assessment domains (i.e., groupwork, exam, grading, cheating, and feedback) and associated justice principles that students in Phase I reported as key themes to evaluate the fairness in classroom assessment. This approach to examine participants’ perceptions of fairness in relation to a specific domain is also recommended as a way forward in social psychology of morality research. By reviewing 1,278 empirical research in social psychology of morality from 1940s to 2017, Ellemers, Toorn, Paunov, and Leeuwen (2019, p. 354) concluded that previous research has either largely focused on examining participants’ endorsement of general principles without attention to specific domains of application or have explored specific domains of moral relevance without connecting these specific domains to underpinning general principles.

Current insights on moral reasoning mostly pertain to relatively abstract principles (such as “fairness”) that people can subscribe to, as well as individual differences in which moral guidelines they endorse. The concrete implications of these general principles for specific situations remain less considered. Research on moral judgments complements this by addressing people’s situational experiences, for instance, resulting from concrete choices or behaviors displayed by others. However, these more specific judgments are not systematically traced back to the general moral principles that might inform them or the (dis)agreement that may exist about how to prioritize these.
In this way, Classroom Assessment Fairness Inventory moves toward addressing this gap by examining how students perceive fairness using multiple relevant justice principles in each specific domain of classroom assessment. Therefore, the review of previous studies in fairness in classroom assessment signals the need for a new instrument measuring students’ perceived fairness.

Third, the previous systematic reviews also showed that no study has been conducted on students’ perceived fairness within Canadian context (Rasooli et al., 2018, under review). This study therefore conducted Phase I qualitative interviews to not only gain an understanding of fairness as grounded within Canadian context, but to also extract domains and relevant justice principles significant to students’ evaluation of fairness in the Canadian classroom assessment context. It is through this effort that this study was able to come up with additional domains of fairness and associated justice principles in classroom assessment that were significant to Canadian secondary students’ perceptions of fairness. Therefore, the inventory is conceptually supported in its effort to fit for its purpose in measuring secondary school students’ perceived fairness in classroom assessment.

Fourth, previous literature and instruments measuring students’ and teachers’ perceived fairness have relied heavily on social psychology theory. Given that this theory has been used extensively to examine individuals’ perceived fairness across contexts, this theory (which was used to interpret students’ perceptions of assessment fairness in Phase I) was employed to support the conceptual basis for developing Classroom Assessment Fairness Inventory. This theory is described in detail below.

Overall, the previous systematic reviews of fairness literature and instruments in classroom (assessment), social psychology theory, and Phase I interviews supported the theoretical underpinning for developing Classroom Assessment Fairness Inventory.

3.5.3 Initial Item Development
Drawing on the five domains of assessment (i.e., groupwork, exam, grading, cheating, and feedback) that were central in students’ interpretations of fairness in classroom assessment in Phase I as well as the broader literature, The Classroom Assessment Fairness Inventory was developed and contained two parts: Part A and B. Part A included demographic questions pertaining to students’ school background, gender, age, typical grades received during secondary school, ethnicity, and immigration backgrounds. The demographic questions also included a short seven-item survey on students’ Personal Belief in a Just World (Peter & Dalbert, 2010). This survey asks questions about the extent individuals believe their personal life events are just. As previous research has shown that students’ perceptions of fairness is influenced by their belief in a just world (e.g., Alt, 2015), the administration of this survey as part of the Classroom Assessment Fairness Inventory provided not only validity evidence based on relationship of responses between this survey and participants’ broader inventory response, but also provided more nuanced understanding of how students’ personal beliefs in a just world linked with their responses to classroom assessment fairness scenarios. The validity and reliability evidence for the personal belief in a just world have been extensively reported in previous literature (Alt, 2015; Dalbert, 1999; Peter & Dalbert, 2010). Varied Alpha levels were reported for internal consistency of this survey in prior studies ranging from $\alpha = .71$ and $\alpha = .81$ (Alt, 2015; Dalbert & Stoeb, 2006; Donat, Dalbert, & Kamble, 2014). The demographic items and personal belief in a just world survey can be found in Appendix I.

Part B included Classroom Assessment Fairness Inventory. This inventory included five scenarios, each with associated actions. In total, there were 40 actions. Each scenario addressed a different classroom assessment domain (e.g., groupwork, exam, cheating, grading, and feedback) significant to students’ perceptions of fairness. These domains were selected based on results from Phase I as well as the comprehensive scrutiny of previous literature. The associated actions within each scenario represented an underlying principle (e.g., equality, equity, transparency)
based on social psychology framework of fairness in classroom assessment (Rasooli et al., 2018, 2019, 2019). The inclusion of relevant principles in each domain was decided based on its significance to students’ evaluation of fairness in that domain as showed in Phase I and previous literature. For example, equal group grades were perceived as unfair by student participants in Phase I due to inequity in the distribution of grades vis-à-vis students’ individual contributions. In the construction of groupwork scenario, an action item (i.e., Mr. Chu did not give individual grades for each group member based on their contributions and learning) was developed to represent this underpinning principle. The fairness framework is presented in Table 4 and the conceptual representation of each dimension in each scenario is presented in Table 5.

Table 4. Descriptions of dimensions and principles of justice and their representations in Classroom Assessment Fairness Inventory

<table>
<thead>
<tr>
<th>Dimensions of fairness</th>
<th>Principles</th>
<th>Descriptions</th>
<th>Representation of each principle across all items (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Distributive Justice</td>
<td>1.1. Equity</td>
<td>Distributing resources (e.g., grades, feedback) in a way that each student receives what they deserve.</td>
<td>4 (10%)</td>
</tr>
<tr>
<td></td>
<td>1.2. Equality</td>
<td>Distributing resources (e.g., grades, feedback) equally among all students.</td>
<td>2 (5%)</td>
</tr>
<tr>
<td></td>
<td>1.3. Need</td>
<td>Distributing resources (e.g., grades, feedback) based on caring for students’ individual needs.</td>
<td>5 (12.5%)</td>
</tr>
<tr>
<td></td>
<td>1.4. Consequence</td>
<td>Distributing resources (e.g., grades, feedback) based on considering consequences for students.</td>
<td>3 (7.5%)</td>
</tr>
<tr>
<td>2. Procedural Justice</td>
<td>2.1. Consistency</td>
<td>Consistent application of assessment procedures in classrooms across students.</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td></td>
<td>2.3. Bias Suppression</td>
<td>Neutral and bias-free application of assessment procedures in classrooms.</td>
<td>2 (5%)</td>
</tr>
<tr>
<td></td>
<td>2.4. Correctability</td>
<td>Correcting assessment procedures when there is an error in process or practice.</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td></td>
<td>2.5. Ethicality</td>
<td>Aligning assessment procedures with ethical standards and practices.</td>
<td>1 (2.5%)</td>
</tr>
</tbody>
</table>
2.6. Voice
Providing students with voice and an opportunity to communicate their perspective during assessment procedures. 5 (12.5%)

2.7. Transparency
Enacting assessment procedures with clarity. 3 (7.5%)

2.8. Reasonableness
Enacting assessment procedures (e.g., exam scheduling) in a way that shows good and sensible judgements. 1 (2.5%)

3. Interactional Justice
3.1. Respect
Treating students with respect during assessment procedures. 4 (10%)

3.2. Adequate communication
Providing students with adequate information about assessment procedures. 2 (5%)

3.3. Justification
Providing students with logical explanations about assessment procedures or rationale for grading and scoring decisions. 4 (10%)

3.4. Timeliness
Providing students with timely information about assessment procedures. 2 (5%)

Total number of items: 40

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Distributional</th>
<th>Procedural</th>
<th>Interactional</th>
<th>#Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Totals</td>
<td>14</td>
<td>14</td>
<td>12</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 5. Representation of each fairness dimension within and across five scenarios

Students were asked to read each scenario and identify the extent to which they consider each action to be fair (1= Highly unfair; 6= Highly fair). An additional option (7 = ‘I am not sure’) was also provided. The five scenarios, associated actions, and underpinning principles are presented in Tables 6 to 10 below. Please note that participants were not shown the categories of justice principles.
Table 6. Classroom Assessment Fairness Inventory- scenario 1 blueprint

Scenario 1. Groupwork
Mr. Chu highly values student groupwork. Based on his initial assessments, Mr. Chu formed each group with three students from different ability levels: struggling, average, and high performing students. He believed that high performing students contribute to the learning of other group members. Each group worked on their projects and prepared a final presentation. Mr. Chu has left it to groups to discuss how to distribute workload and allowed students to discuss with him if they had issues over group dynamics. As a response to students’ questions about assessment, Mr. Chu provided an overview of the project to students but not a rubric showing how he will assess students’ groupwork. Mr. Chu encouraged group members to work hard as all group members will receive the same grades as a reflection of group performance and cooperation. Several who were not satisfied with their grades appealed, but Mr. Chu did not accept their complaints.

To what extent do you consider each of these actions to be fair based on the above scenario?

<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mr. Chu selected group members based on mixed ability.</td>
<td>Consequence</td>
</tr>
<tr>
<td>2. Mr. Chu did not provide students a choice in selecting their group members.</td>
<td>Voice</td>
</tr>
<tr>
<td>3. Mr. Chu allowed students’ complaints over group dynamics.</td>
<td>Voice</td>
</tr>
<tr>
<td>4. Mr. Chu was not detailed in communicating how he will assess students’ groupwork.</td>
<td>Adequate communication</td>
</tr>
<tr>
<td>5. Mr. Chu gave the same grades to all group members.</td>
<td>Equality</td>
</tr>
<tr>
<td>6. Mr. Chu did not give individual grades for each group member based on their contributions and learning.</td>
<td>Equity</td>
</tr>
<tr>
<td>7. Mr. Chu did not justify his grades to students who appealed.</td>
<td>Justification</td>
</tr>
</tbody>
</table>
Scenario 2. Exam

Mr. Ahmed announced that the class would have an exam the day before winter break (in 5 days). Students preferred moving the exam date because they had many assignments for other subjects that were also due on the same date. Even so, Mr. Ahmed was firm on his decision as moving the exam date back would create more intensive workload later in the year. Mr. Ahmed did not explicitly state what would be on the exam. However, he did include a mix of easy and difficult questions to give all students an opportunity to show their learning. He also provided accommodations (e.g., more time) to students with disabilities and English language learners. In general, Mr. Ahmed is a lenient teacher in grading compared with other teachers in the school who teach the same subject. On the exam, all students complained about two questions that were not covered during the course. Mr. Ahmed harshly responded that students should be able to answer the two questions from what had been taught. For a few students who missed the exam date, Mr. Ahmed decided to give another exam opportunity after winter break.

To what extent do you consider each of these actions to be fair based on the above scenario?

<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mr. Ahmed held firm on the exam date.</td>
<td>Reasonableness</td>
</tr>
<tr>
<td>2. Mr. Ahmed did not explicitly state what would be on the exam.</td>
<td>Adequate communication</td>
</tr>
<tr>
<td>3. Mr. Ahmed included in the exam both easy to difficult questions.</td>
<td>Equity</td>
</tr>
<tr>
<td>4. Students with disabilities and English language learners received accommodations for the exam (e.g., more writing time).</td>
<td>Need</td>
</tr>
<tr>
<td>5. Mr. Ahmed graded his students more leniently than other teachers.</td>
<td>Consistency</td>
</tr>
<tr>
<td>6. Mr. Ahmed did not remove the two questions on the content that were not taught before.</td>
<td>Correctability</td>
</tr>
<tr>
<td>7. Mr. Ahmed did not respond to students’ complaints with a respectful tone.</td>
<td>Respect</td>
</tr>
<tr>
<td>8. Mr. Ahmed gave another exam opportunity to students who missed the exam.</td>
<td>Need (Equality, Consistency)</td>
</tr>
</tbody>
</table>
Table 8. Classroom Assessment Fairness Inventory- scenario 3 blueprint

**Scenario 3. Cheating**

Ms. Johnston is very strict when she catches a student cheating. However, she did not tell students her policy on cheating at the beginning of the year. One student was caught cheating on an exam and Ms. Johnston decided to give the student a grade of zero. Ms. Johnston did not give the student an opportunity to explain the reasons for cheating before making her decisions. She explained to the class that cheating is unfair to other students and asked the student to leave the classroom. The exam constituted 20% of students’ final grade. After the exam, she met with the student and explained that cheating is ethically wrong, is unfair in relation to the classmates, and she would punish anyone who cheats. The student provided a reason for their behavior and apologized.

**To what extent do you consider each of these actions to be fair based on the above scenario?**

<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ms. Johnston was not transparent about her cheating policy at the beginning of the course.</td>
<td>Transparency</td>
</tr>
<tr>
<td>2. Ms. Johnston did not give the student an opportunity to explain the reasons for cheating before making her decision to give zero.</td>
<td>Voice</td>
</tr>
<tr>
<td>3. Ms. Johnston gave a zero because cheating is unfair to other students’ efforts.</td>
<td>Equity</td>
</tr>
<tr>
<td>4. Ms. Johnston gave a zero because cheating is ethically wrong.</td>
<td>Ethicality</td>
</tr>
<tr>
<td>5. Ms. Johnston gave a zero to signal that anyone who cheats should be punished for this action.</td>
<td>Consequence</td>
</tr>
<tr>
<td>6. Ms. Johnston asked the student to leave the classroom in front of other students.</td>
<td>Respect</td>
</tr>
<tr>
<td>7. Ms. Johnston explained her cheating decision to the student.</td>
<td>Justification</td>
</tr>
<tr>
<td>8. Ms. Johnston did not forgive the student’s cheating this time.</td>
<td>Need</td>
</tr>
</tbody>
</table>
Scenario 4. Grading

Ms. Mendes had students from diverse backgrounds in her classroom. She treated all her students respectfully during classroom teaching, assessment, and interactions. Ms. Mendes informed students that she would give grades based on student achievement. 70% of students’ grades were from multiple tests during the course plus 30% for students’ individual essays. Ms. Mendes communicated test results in one week after handing the test in. Due to busy schedule, she would sometimes allow students to appeal their grades if there was enough time in class. She would fully explain her grading for students who spoke up looking for their grade adjustments. At the end of the course, Ms. Mendes adjusted the grades of failing students with at-risk backgrounds to support their success. She also increased marks for a few students to ensure admission into their desired universities. However, she lowered the grades of a few disruptive students who interrupted the classroom learning.

To what extent do you consider each of these actions to be fair based on the above scenario?

<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ms. Mendes treated students respectfully during classroom assessment.</td>
<td>Respect</td>
</tr>
<tr>
<td>2. Ms. Mendes largely considered student achievement of learning objectives in her grading.</td>
<td>Equality</td>
</tr>
<tr>
<td>3. Ms. Mendes detailed her grading criteria, with test scores making up 70% of a student’s grade.</td>
<td>Transparency</td>
</tr>
<tr>
<td>4. Ms. Mendes communicated test results in one week after handing the test in.</td>
<td>Timeliness</td>
</tr>
<tr>
<td>5. Ms. Mendes would sometimes allow students to discuss their grades if there was enough time in class.</td>
<td>Voice</td>
</tr>
<tr>
<td>6. Ms. Mendes gave adequate justification for students who spoke up for their grades.</td>
<td>Justification</td>
</tr>
<tr>
<td>7. Ms. Mendes adjusted the grades of failing students with at-risk backgrounds.</td>
<td>Need</td>
</tr>
<tr>
<td>8. Ms. Mendes considered students’ future university admissions to adjust grades.</td>
<td>Consequence</td>
</tr>
<tr>
<td>9. Ms. Mendes considered student misbehavior (e.g., disruptions) in her grading.</td>
<td>Bias suppression</td>
</tr>
</tbody>
</table>
Table 10. Classroom Assessment Fairness Inventory - scenario 5 blueprint

<table>
<thead>
<tr>
<th>Scenario 5. Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Dembe has asked students to write an essay about their science lab project. He has shared with students a clear rubric that he will use to assess students’ essays. Four days after the essay submission deadline, Mr. Dembe got back to students with his feedback and expressed that students can contact him for further discussion of his feedback. Students noticed that Mr. Dembe gave more feedback to students with good quality essays as well as his favorite students than students who had handed in essays that were of low quality. Mr. Dembe <em>harshly</em> explained that he had given variable feedback on the essays based on the amount of effort he deemed each student had put in completing essays.</td>
</tr>
</tbody>
</table>

To what extent do you consider each of these actions to be fair based on the above scenario?

<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mr. Dembe provided a clear rubric for assessing students’ essays.</td>
<td>Transparency</td>
</tr>
<tr>
<td>2. Mr. Dembe provided feedback after four days of essay submissions.</td>
<td>Timeliness</td>
</tr>
<tr>
<td>3. Mr. Dembe gave students a chance to further discuss his feedback.</td>
<td>Voice</td>
</tr>
<tr>
<td>4. Mr. Dembe provided feedback based on the amount of effort each student has put in the work.</td>
<td>Equity</td>
</tr>
<tr>
<td>5. Mr. Dembe did not provide more feedback to students who had weaker performance.</td>
<td>Need</td>
</tr>
<tr>
<td>6. Mr. Dembe did not treat students respectfully in his feedback procedure.</td>
<td>Respect</td>
</tr>
<tr>
<td>7. Mr. Dembe gave more feedback to his favorite students.</td>
<td>Bias suppression</td>
</tr>
<tr>
<td>8. Mr. Dembe explained that his feedback procedure was based on the amount of effort each student has put in.</td>
<td>Justification</td>
</tr>
</tbody>
</table>

3.5.4 Validity Evidence
The validity for the Classroom Assessment Fairness Inventory was investigated by collecting evidence for test content, internal structure, and relationship to other variables using expert-panel review, pilot testing process, and psychometric analyses (Standards for Educational and Psychological Testing, 2014). Combined, these evidence sources contributed to evaluate the validity of Classroom Assessment Fairness Inventory.

3.5.4.1 Evidence of Validity Based on Test Content: Expert-Panel and Pilot Testing Review

Expert panel and pilot testing reviews were used to collect evidence for the validity on the Classroom Assessment Fairness Inventory. Nineteen experts were identified based on their previous contributions to the fairness and classroom assessment literature. Experts were identified based on their publications on educational assessment journals focusing on fairness and classroom assessment. Experts were selected to have a fair diverse representation in terms of (a) experienced leader in assessment vs. early career researcher; (b) more psychometrically focused vs. more assessment education focused; and (c) local researcher vs. international researcher. Combined, these criteria enabled to have diverse perspectives commenting on different parts of the inventory. These experts were emailed to invite for reviewing the inventory, with attachments detailing the study purpose and consent forms. 10 experts from North America, Europe, Australia, and Middle East consented to review the manuscript.

The experts were provided with a review template including demographic information as well as the blueprint of Classroom Assessment Fairness Inventory. They were asked to provide their feedback about the revisions and additions to the demographic questions. They were also provided with concise information about the purpose of Classroom Assessment Fairness Inventory, and a blueprint outline of five scenarios, action items in each scenario, and associated justice principles and their definitions. An alignment methodology was used to guide the expert review process (DeLuca & Bellara, 2013). Experts were asked to (a) suggest revisions to the scenarios, (b) suggest revisions to the actions, (c) examine the alignment of scenario actions with
the underpinning justice principles (1= not aligned to, 5= strongly aligned), (d) examine whether the actions in each scenario adequately represented the various justice principles relevant to the particular scenario (1= completely unrepresentative, 5= completely representative), and (e) examine in overall whether all five scenarios represented the construct of fairness in classroom assessment adequately (1= completely unrepresentative, 5= completely representative). The expert ratings of alignment across justice principles and scenario actions are presented in Table 11. The expert ratings of representativeness showed high agreement, indicating that experts considered the actions in each and across scenario(s) to represent fairness in classroom assessment adequately.

The demographic questions, scenarios, and actions were revised according to expert ratings, expert comments, and inter-rater agreement. Expert ratings less than 3 (indicating weaker alignment of scenario actions) signaled closer inspection of expert comments to revise the scenario actions. Experts’ comments were used to revise wording of scenarios and actions, add demographic items, and re-examine the scenario actions and justice principles to ensure alignment. Intraclass correlation coefficient was also calculated for inter-rater agreement for each scenario as well as the overall inventory (McGraw & Wong, 1996). The Intraclass correlation coefficient using two-way mixed model with measures of absolute agreement showed moderate average agreement across experts for scenario 1 (.62, p < .01), scenario 2 (.45, p < .06), scenario 3 (.61, p < .01), scenario 4 (.45, p < .07), and scenario 5 (.51, p < .03). The overall inter-rater agreement also showed moderate average value (.56, p < .001). The scrutiny of moderate inter-rater agreement was also used as another layer to revise the inventory.

In addition, invitations were sent to 10 graduate students with teaching experience in K-12 teaching contexts to pilot test and review the content representativeness of Classroom Assessment Fairness Inventory scenarios. All ten graduate students consented to complete the
review. The demographic questions, scenarios, and actions were further modified to simplify the readability of scenarios for inclusivity issues.

Table 11. The expert ratings of alignment across justice principles and scenario actions

<table>
<thead>
<tr>
<th>Scenario actions</th>
<th>Number of ratings indicating weaker alignment (1-3)</th>
<th>Number of ratings indicating stronger alignment (4-5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sc1-1</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Sc1-2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Sc1-3</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Sc1-4</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Sc1-5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Sc1-6</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Sc1-7</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Sc2-1</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Sc2-2</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Sc2-3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Sc2-4</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Sc2-5</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Sc2-6</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Sc2-7</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Sc2-8</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Sc3-1</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Sc3-2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Sc3-3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sc3-4</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Sc3-5</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Sc3-6</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Sc3-7</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Sc3-8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sc4-1</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Sc4-2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Sc4-3</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Sc4-4</td>
<td>0</td>
<td>9</td>
</tr>
</tbody>
</table>
3.5.4.2 Evidence of Validity Based on Internal Structure and Relationship to Other Variables

Empirical data from student participants were collected to provide evidence for the validity of Classroom Assessment Fairness Inventory based on internal structure and relationship to other variables.

3.5.5 Recruitment

Due to Covid-19 pandemic, the recruitment of participants for Classroom Assessment Fairness Inventory administration was conducted via online social media platforms. Recruitment scripts (Appendix K) including Qualtrics-based links were distributed in social media networks including Facebook and Twitter, inviting first-year undergraduate students for completing the inventory. First-year undergraduate students who had completed at least one of their grade levels at a secondary school in Ontario, Canada represented the target sample for this study (i.e., selection criteria). This sample was targeted on social media pages, representing the first-year undergraduate students in various universities and colleges in Canada. First-year undergraduate...
students were selected because they had recent perceptions and experiences of fairness in classroom assessment in secondary schools in Ontario. Participants who were willing to enter a prize draw to win 1 of 50 10$ gift card were asked to provide their emails at another survey linked to the end of inventory. The recruitment of student participants for Phase II began in November 2020 and ended at the end of January 2021.

3.5.6 Participants

400 participants from first-year undergraduate students clicked on the link to participate in this study. 217 participants completed the entire inventory (showing 54% complete response rate). Most participants were female (68%) and were 18 years old (60%). Most participants reported being born in Canada (62%), with both parents born outside of Canada (46%). Most participants were from Caucasian racial background (47%), and had studied in public school (72%), with average grades between 90-100% (48%). The details of participants’ demographic information are presented in Table 11.

Table 12. Overview of student participants’ demography (n=217)

<table>
<thead>
<tr>
<th>Gender</th>
<th>N (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>147 (68.2%)</td>
</tr>
<tr>
<td>Male</td>
<td>65 (30%)</td>
</tr>
<tr>
<td>I do not identify with gender binary</td>
<td>4 (1.8%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>N (Percentage)</td>
</tr>
<tr>
<td>15 years</td>
<td>2 (.9%)</td>
</tr>
<tr>
<td>17 years</td>
<td>5 (2.3%)</td>
</tr>
<tr>
<td>18 years</td>
<td>131 (60.4%)</td>
</tr>
<tr>
<td>19 years</td>
<td>42 (19.4%)</td>
</tr>
<tr>
<td>20 years</td>
<td>12 (5.5%)</td>
</tr>
<tr>
<td>21</td>
<td>4 (1.8)</td>
</tr>
<tr>
<td>Above 21</td>
<td>21 (9.7%)</td>
</tr>
<tr>
<td>Student Immigration Background</td>
<td>N (Percentage)</td>
</tr>
<tr>
<td>Student was born in Canada</td>
<td>135 (62.2%)</td>
</tr>
<tr>
<td>Student was born outside of Canada</td>
<td>59 (27.2%)</td>
</tr>
<tr>
<td>Missing data on student’s birthplace</td>
<td>23 (10.6%)</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>Parent Immigration Background</strong></td>
<td>N (Percentage)</td>
</tr>
<tr>
<td>Both parents/guardians were both born in Canada</td>
<td>79 (36.4%)</td>
</tr>
<tr>
<td>Both parents/guardians were both born outside of Canada</td>
<td>101 (46.5%)</td>
</tr>
<tr>
<td>One of parents was born outside of Canada</td>
<td>12 (5.5%)</td>
</tr>
<tr>
<td>I don’t know about parent birthplace</td>
<td>2 (.9%)</td>
</tr>
<tr>
<td>Missing data on parents’ birthplace</td>
<td>23 (10.6%)</td>
</tr>
<tr>
<td><strong>Racial Background</strong></td>
<td>N (Percentage)</td>
</tr>
<tr>
<td>Indigenous (First Nations, Metis, or Inuit)</td>
<td>4 (1.8%)</td>
</tr>
<tr>
<td>Persian/Arab/West Asian</td>
<td>10 (4.6%)</td>
</tr>
<tr>
<td>Black</td>
<td>15 (6.9%)</td>
</tr>
<tr>
<td>East Asian</td>
<td>32 (14.7%)</td>
</tr>
<tr>
<td>Latin American</td>
<td>4 (1.8%)</td>
</tr>
<tr>
<td>South Asian</td>
<td>26 (12%)</td>
</tr>
<tr>
<td>South East Asian</td>
<td>14 (6.5%)</td>
</tr>
<tr>
<td>Caucasian (White)</td>
<td>103 (47.5%)</td>
</tr>
<tr>
<td>Mixed</td>
<td>7 (3.2%)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>2 (.9%)</td>
</tr>
<tr>
<td><strong>School Background</strong></td>
<td>N (Percentage)</td>
</tr>
<tr>
<td>Public</td>
<td>157 (72.4%)</td>
</tr>
<tr>
<td>Catholic</td>
<td>40 (18.4%)</td>
</tr>
<tr>
<td>Private</td>
<td>20 (9.2%)</td>
</tr>
<tr>
<td><strong>Average Grade in High School</strong></td>
<td>N (Percentage)</td>
</tr>
<tr>
<td>90-100%</td>
<td>106 (48.8%)</td>
</tr>
<tr>
<td>80-89%</td>
<td>85 (39.2%)</td>
</tr>
<tr>
<td>70-79%</td>
<td>23 (10.6%)</td>
</tr>
<tr>
<td>60-69%</td>
<td>3 (1.4%)</td>
</tr>
</tbody>
</table>

### 3.5.7 Data Collection

The Classroom Assessment Fairness Inventory was distributed via social media platforms such as Facebook and Twitter using a Qualtrics Software link. The Qualtrics software link enabled access to first-year undergraduate students across Ontario. In the Facebook and Twitter posts, potential participants were asked to click on the Qualtrics link to complete the survey that could take 20 minutes of their time. The initial page included the purpose of the study as well as
the consent form (see Appendix G) for students to sign prior to initiating the survey completion. Once the participants consented to complete the survey, they were given demographic questions, followed by Personal Belief in a Just World survey. They were then provided with the Classroom Assessment Fairness Inventory scenarios. To complete this inventory, participants were asked to reflect on their secondary school experiences. At the end of this survey, participants were given the opportunity to click on a link that would take them to another survey, where they could insert their emails to win 1 of the 50 $10 gift cards. Accordingly, 50 participants were drawn from the sample of respondents and compensated for their participation in this study.

3.5.8 Data Analysis

The data analysis followed two steps. First, an initial exploratory factor analysis was conducted to determine the nature and number of factors as a basis to conduct confirmatory factor analysis. Given that Classroom Assessment Fairness Inventory was constructed based on an underpinning conceptual model, the use of confirmatory factor analysis was deemed appropriate to test the model empirically. However, exploratory factor analysis was conducted prior to confirmatory factor analysis to provide initial configurations about the empirical models fitting the data. The result for Kaiser-Meyer-Olkin value was .79, and Bartlett’s Test of Sphericity was also significant (p = .001), supporting the use of exploratory factor analysis. Exploratory factor analysis was run using Maximum Likelihood with direct oblimin rotation to interpret factor solutions. Maximum Likelihood with direct oblimin rotation was selected because (a) students’ responses to Classroom Assessment Fairness Inventory were intercorrelated and assumed to be driven by students’ perceptions of fairness in classroom assessment. The Eigen values, scree plot, and Monte Carlo PCA for Parallel Analysis were used to determine the possible empirically fitting models with the data. Confirmatory factor analysis was followed to test and interpret the best fitting model with the data in light of the conceptual model. Several goodness-of-fit indices were used to compare across models and identify the best fitting model. These indices included
the root mean square error of approximation (RMSEA < .08), comparative fit index (CFI > .95), Tucker-Lewis index (TLI > .95), and chi-square test of model fit (Hu & Bentler, 1999; Kline, 2016).

Second, structural equation modeling was used to examine the relationships across the factors of Classroom Assessment Fairness Inventory and Dalbert’s (1999) personal belief in a just world survey. Several goodness-of-fit indices including RMSEA < .08, CFI > .95, TLI > .95, and chi-square test of model fit were used to test the best model fitting with the data (Hu & Bentler, 1999; Kline, 2016). Reliability analysis was also conducted using Cronbach’s alpha to examine the internal consistency of the personal belief in a just world survey. All the analyses were conducted in SPSS, AMOS, and Mplus. Combined, the analyses provided evidence based on internal structure and relationships to other variables to support the validity of Classroom Assessment Fairness Inventory. The analyses also provided evidence to appreciate students’ perceptions of fairness in classroom assessment in relation to secondary schools in Ontario, Canada.
Chapter 4

Findings

This chapter outlines the findings from Phases I and II in an aim to address how first-year undergraduate students retrospectively perceived fairness in classroom assessment in their secondary experience in Ontario, Canada. It also addresses what psychological and behavioral consequences students reported in response to perceptions of un/fairness. Specifically, this chapter first presents qualitative evidence stemming from interviews in Phase I and then outlines quantitative evidence resulting from the Classroom Assessment Fairness Inventory in Phase II to understand students’ perceived fairness in classroom assessment.

4.1 Findings for Phase I

In the interviews, participants articulated what fairness in classroom assessment generally meant to them as well as how they perceived fairness in various domains of classroom assessment. They also expressed their psychological and behavioral responses to un/fairness in assessment. Combined, students’ perceptions of fairness in classroom assessment were summarized in eight themes: (1) overall perceptions of fairness; (2) fairness in groupwork; (3) fairness in exams; (4) fairness in cheating; (5) fairness in grading; (6) fairness in feedback; (7) socio-emotional environment; and (8) responses to perceptions of un/fairness. Overall, 307 codes underpinned these themes. A general overview of themes and associated codes can be found in Table 13.

Table 13. A general overview of themes and associated codes

<table>
<thead>
<tr>
<th>Themes</th>
<th>Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Perceptions of Fairness</td>
<td>Equity</td>
</tr>
<tr>
<td></td>
<td>Equality</td>
</tr>
<tr>
<td></td>
<td>Respect</td>
</tr>
<tr>
<td>Fairness in Groupwork</td>
<td>Group Composition</td>
</tr>
</tbody>
</table>

94
<table>
<thead>
<tr>
<th>Group Dynamics</th>
<th>Grading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fairness in Exams</td>
<td>Reasonable Scheduling</td>
</tr>
<tr>
<td></td>
<td>Student Voice in Exams</td>
</tr>
<tr>
<td></td>
<td>Transparency in Exam Information</td>
</tr>
<tr>
<td></td>
<td>Teaching and Exam Content Alignment</td>
</tr>
<tr>
<td></td>
<td>Missed Exam Policies</td>
</tr>
<tr>
<td>Fairness in Cheating</td>
<td>Transparency in Cheating Policy</td>
</tr>
<tr>
<td></td>
<td>Consistency and Bias Suppression in Decision</td>
</tr>
<tr>
<td></td>
<td>Voice Opportunity</td>
</tr>
<tr>
<td></td>
<td>Assigning Zero</td>
</tr>
<tr>
<td></td>
<td>Caring about Student Background</td>
</tr>
<tr>
<td></td>
<td>Projecting Consequences of Cheating Decisions</td>
</tr>
<tr>
<td></td>
<td>Utilitarian View</td>
</tr>
<tr>
<td>Fairness in Grading</td>
<td>Equity and Caring</td>
</tr>
<tr>
<td></td>
<td>Transparency in Criteria</td>
</tr>
<tr>
<td></td>
<td>Consistency in Implementation</td>
</tr>
<tr>
<td></td>
<td>Bias Suppression in Grading</td>
</tr>
<tr>
<td></td>
<td>A Voice Opportunity to Appeal</td>
</tr>
<tr>
<td></td>
<td>Adequate Justifications</td>
</tr>
<tr>
<td>Fairness in Feedback</td>
<td>Transparent Feedback</td>
</tr>
<tr>
<td></td>
<td>Timely Feedback</td>
</tr>
<tr>
<td></td>
<td>Adequate and Honest Feedback</td>
</tr>
<tr>
<td>Socio-emotional Environment</td>
<td>Teacher-student Relationships</td>
</tr>
<tr>
<td></td>
<td>Gender Equity</td>
</tr>
<tr>
<td></td>
<td>Racial Equity</td>
</tr>
<tr>
<td></td>
<td>Peer Relationships</td>
</tr>
<tr>
<td>Responses to Un/fairness</td>
<td>Learning and Motivation Outcomes</td>
</tr>
<tr>
<td></td>
<td>Emotional Responses</td>
</tr>
<tr>
<td></td>
<td>Dropout from the Course</td>
</tr>
<tr>
<td></td>
<td>Talk with Principal, Teacher, and Others</td>
</tr>
<tr>
<td></td>
<td>Rationalizations for Perseverance</td>
</tr>
<tr>
<td></td>
<td>Inaction</td>
</tr>
</tbody>
</table>

### 4.1.1 Theme 1: Overall Perceptions of Fairness

Students considered three overarching principles and expressed 24 examples associated with these principles to interpret what fairness in classroom assessment meant to them in general.

These three principles included (a) *equality*, (b) *equity*, and (c) *respect*. Several students
highlighted *equality* in ten examples. They stated that equal opportunity to learn and demonstrate learning contributes to fair assessments. For example, a student articulated that “for me, fairness in classroom assessment means equal opportunities and access to resources” (S20), with another student expressing that “everyone is given the same opportunities and are assessed with the same level of equality.” (S19) Students also interpreted overall fairness in terms of equality in treatment regardless of student background, race, and gender. “To me, fairness in classroom assessment means, getting treatment from people without discrimination taking place and favoritism is avoided.” (S18) Students also perceived unfairness in their observation of unequal treatment targeting peers, “unfairness and bias in an academic environment can be very subtle and a lot of times happens to people of color particularly those who are weak in the English language or have a hard time fitting in.” (S10)

Some students interpreted overall fairness in classroom assessment in relation to *equity* in eight examples. Students argued for equity in the spirit of caring to compensate for the disadvantaged personal histories of some students. For example, a student stated “To me, fairness in classroom assessment means considering the individual context behind each person. Though their history should not be the deciding factor, it should help play a role in making things more equitable.” (S6) While students showed advocacy for equity in caring about students’ personal history, they also emphasized the significance of academic challenges to support students’ learning. A student drew on his personal life to interpret the meaning of fairness in classroom assessment in this respect.

Unfortunately, the circumstance of my living situation growing up was not very good and many people would consider this unfair… classroom assessment fairness means recognizing that a student may have a difficult lot in life and addressing/adjusting the expectations for students on an individualized basis based on their level of need. Sometimes, students need to be challenged so that they can rise to the occasion and see
that hard work pays off, but other times students need to be given more lenience and
offered grace where they may have failed. These two things in conjunction develop a
more complete understanding of fairness in assessment. (S25)

Some students reported respect as another key overall principle of fairness in classroom
assessment in six examples. They stated that everyone should be treated respectfully: “Everyone
is treated equally and has the right to speak respectfully.” (S27) While respect was emphasized to
include all classroom relationships, a student argued for reciprocal form of respect, where if a
student is disrespectful to a peer, the peer can behave differently. “Fairness can be changed
depending on how you are treated by someone. If there is a student that is being disrespectful you
have the right to change the way you act with them in order to keep your dignity.” (S27)

Overall, students used equality, equity, and respect principles to interpret the overall
meaning of fairness in classroom assessment. Students in this sample provided more examples of
equality, followed by equity, and respect. These three overall conceptions were at a complex
interplay with each other in shaping students’ overall perceptions of fairness in classroom
assessment. While the endorsement of one of these conceptions was mainly provoked at a
particular situational experience (e.g., equality and exclusion of any discriminatory treatment),
students have referenced at times more than one of these conceptions in perceiving fairness in
classroom assessment (e.g., everyone should be respected).

4.1.2 Theme 2: Fairness in Groupwork

Students stated 60 examples associated with groupwork, making it a significant domain
in shaping their perception of fairness in classroom assessment. Students perceived fairness of
groupwork in group composition, group dynamics, and grading. Across these three codes, it
seems that students perceived inequity when the groupwork cooperation did not produce
outcomes (i.e., grades) that were comparable to student (individual) contributions. In group
composition that surfaced in 14 examples, most students argued for their control and voice over
selecting group members. They argued that this control over group member selection ensured managing successful cooperation that led to fairer outcomes. In this respect, a student articulated “it depends really, because students should be given the opportunity to sort of work with people that they think they could do well with because obviously, you know, school is just like it's a path to your own future.” (S18) For this reason, some students were against the teacher selection of group members. For example, a student suggested that while a teacher might consider varied ability grouping as fair, the success of cooperation may not be guaranteed by ability alone.

I felt it was a little unfair the way the teacher made groups because we will have to do a couple individual assessments and then she put people, like one person in the group that has like the 90 and then another person in the group that has 80, and a couple people that have like 50, 60s, lower marks in the class. Sometimes we're just not compatible to work with those people. But she just put us like all together because she figured like everyone can help each other out. So I guess it was sort of fair from looking at it from her eyes. But while we're in the group, I felt like it was unfair because we didn't have the freedom. We didn't have the choice to choose who we wanted to work with. (S11)

While some students were against teacher selection, some others were in favor because teacher selection would avoid conflicts with school friends. The comments of these students showed that they gave more priority to the value of friendship that weakened the value of grades in comparison. A student stated that a teacher selection of groupwork members was fair because you can build on each other's strengths. And I think it's harder to say no to your friend if they asked you to do something. Because you're all just going to say yes, because you want them to like you more and stuff. So I like when teachers chose partners because I thought it was more fair because everybody was getting the same chance. (S14)

Other students considered teacher selection as fair because it prevented high ability students grouping together at the expense of excluding others: “when the teacher chooses groups, it's
always gonna be varied ability. Like you would never get all the high students pack together in one group. But if the teacher says, choose your own group, all the high people get together”.

(S16) To avoid these conflicts, a few students opted for random assignment of group members, “I would say that randomize grouping is probably best if the teacher doesn't let the students pick their group for themselves.” (S7)

While student, teacher, and random selection were discussed by some students as strategies for creating fair assessment, a student commented how inclusion in a friendship structure within school may push a student to do all the work in the group to receive a valued status as a friend in compensation.

In many cases, you would have your group of friends and then one person who would be seen as like the nerd. And you have a group of friends who are very popular would say, have a lot of influence in the class. And they would choose this one nerd to incorporate into their group, and then get them to do all the work. So, if you were singled out as someone who was very intelligent, but didn't have many friends, then often the people who had quite a few friends would sort of use you. And then of course, after the project was finished, they wouldn't remain. (S19)

This excerpt signals that the fairness of groupwork not only includes the group member selection strategies but also is shadowed by the broader inclusion structure within the school.

Some students also discussed fairness in group work in relation to group dynamics in 22 examples and grades in 24 examples. As students valued voice and control over group composition, they also highlighted voice and control to appeal for group dynamics and grades. In cases where a teacher decided to give an equal grade to group members, several students valued the opportunity and voice to appeal to the teacher when a student was not contributing to the groupwork. These students reported that equal grades in groupwork increases free-riding, where
one or two group member(s), often high ability members, complete the entire work while other members, often low ability members, are not putting into the work.

The best example I’d have for this is often in group work, is that often one person does more of the work almost the whole project. And then they sometimes get frustrated. And once or twice I saw the person who ended up doing more work felt that it was so unfair that they actually told the teacher that the other person had done no work in order to try to get their mark boosted. (S5)

In these cases, students appreciated if the teachers would provide students with voice to appeal for group dynamics and incorporate strategies that account for each members’ contributions transparently.

Um, something that's helped me a lot is when the teacher does place you in a group, with different ability levels, and they're like, Oh, if any of your group members had any problems whatsoever, you can tell me and I'll try to work with that student. (S16)

But in high school, if you say if you go to the teacher and say, like, oh, like this member is not doing any work, the teacher doesn't really care. It's up to like the students to figure it out. So I feel like in high school oftentimes I have been in groups where like, one person just wasn't doing the work and that was sort of unfair. (S15)

Although some students suggested voice, control, and transparency to ensure fairness in the cases of equal contributions for equal grades, most students considered individual grades as the fairest approach in grading group work. In this sense, students believed that each student would receive equitable grading proportional to their contributions. “I think for group work the way it should be done is individual assessment, like, have people work together and submit their own things. Instead of having the chance that one person can do all the work.” (S13) A few students also reported that fairness in groupwork would be better served if no grades were allocated, leading to less worrisome cooperation.
I feel like if it was a written report and mock grades given that the kids that are in like the higher end of the spectrum, the more I thought they'd be more cooperative to collaborate with other members in their class, because they know that this is basically just focusing on working with each other and creating an end product as opposed to excel all the requirements of a rubric to get that final grade. (S11)

Overall, students perceived fairness in groupwork as a significant domain in assessment and considered group composition, dynamics, and grading as key components in evaluating fairness of groupwork. While having different perceptions of fairness based on the varied possibilities in group composition, dynamics, and grading, students endorsed the underpinning role of fairness in managing successful cooperation that led to more deserved outcomes for group members.

4.1.3 Theme 3: Fairness in Exam

Students reported 30 experiences that centered on fairness in exams. Codes that supported this theme related to (a) enacting reasonable scheduling for exams and quizzes, (b) considering student voice in exam scheduling, (c) enhancing transparency in exam information, (d) aligning teaching content and exam content, and (e) enacting missed exam policies. In six examples, students reported that they perceived unfairness when teachers did not provide reasonable schedules for exams and quizzes. A student stated “sometimes some teachers gave us a test last minute, like the test was gonna be Friday or something. They didn’t tell us like weeks ahead of time.” (S21)

Another issue that provoked some students’ perceived fairness in scheduling of an exam pertained to whether the teacher listened to students’ voice in their decisions. This issue was discussed in eight examples. A student articulated how a teachers’ approach in resolving conflicts with students over exam scheduling was perceived fair when the students’ voice was considered.
If we would say, oh teacher, you know, on this day that you assigned us this we have two other tests and an assignment. They would then say, okay, is there a way we can move it, how does this work. In this way, they could make it fair for us. (S27)

Several students also referred to transparency in exam information in eight examples and alignment of exam content with teaching content in five examples to contribute to their perceived fairness in assessment. For example, a student viewed a lack of transparency in communicating what to expect as the source of his unfair perception. “In chemistry, we'll be learning thermodynamics or something. And on the test, it will be something completely different from the content that we've been learning. Really let students know all work.” (S26)

With respect to aligning exam content with test content, a student reported an experience, where their teacher failed to fully teach the content of a standardized test that was administered to all students with the same subject in the school.

In my school all exams are standardized. My chemistry teacher was the more lenient teacher, but what happened was he was really slow at teaching. So we were really behind. When it came down to the final exam, there were a lot of topics that he didn't teach us, but they were still present on the exam, which I thought was unfair because the other classes had learned in depth with their teachers. (S15)

A few students also reported three experiences of unfairness in relation to missed exam policies, where a teacher followed a rigid policy to enact punishments to students that missed an exam. A student shared an experience where a teacher did not show flexibility in enacting the missed exam policy and gave a zero to the student as a result of her absence on the exam date.

One thing that came to the top of my mind, like when we first started discussing fairness and unfairness, back in this was Grade 9, I believe I had a very strict math teacher. She had a policy that if you missed a test or an assignment for whatever reason, unless you followed a very specific protocol, you would get a zero. So one day I had a math test in
the morning, that I know I was gonna be late that I was so scared of this teacher, you
know, giving zero. So I just decided that I would say I was sick and coming right at the
next day. So I didn't follow the necessary protocol. I came in and asked to write the test.
She told me no and then I got a zero. This dropped my math mark about 5% overall by
the end of the year. (S12)

Combined, students perceived fairness in exams as a key domain of fairness in assessment and
uttered various considerations associated with perceiving fairness in this domain. Specifically,
they highlighted reasonable scheduling of exams with attention to student voice, transparency and
alignment in exam content with teaching content, and flexibility in missed exam policies as issues
contributing to perceptions of fairness.

4.1.4 Theme 4: Fairness in Cheating

Students reported 33 experiences of cheating as another key dimension of fairness in
classroom assessment. Students’ perceptions of fairness in cheating were underpinned by multiple
considerations including (a) observing transparency in the cheating policy, (b) upholding
consistency and bias-suppression in cheating decisions, (c) providing an opportunity for students
to explain cheating behavior, (d) giving zero and deducting scores as a result of cheating to
restore equity, (e) caring about a student’s personal background, (f) projecting consequences of
cheating decisions, and (h) using a utilitarian perspective to deter cheating acts.

A few students shared four examples focusing on transparency in the cheating policy in a
sense that when a teacher explicitly communicates the punishments for cheating at the beginning
of the course, it can increase fairness in the cheating scenarios. For example, a student stated how
a lack of such transparency generates unfairness. “I think the biggest problem that students have
is that they just don't know what the consequences are. It becomes unfair when the students aren't
aware of the consequences.” (S19) A few students also recollected four unfair experiences where
a teacher did not implement *consistent cheating decisions* across students due to perceived favoritism toward a student.

There was a math test of some sort. And throughout the year, there was a bit of cheating just going on in that class in general. But when the teacher caught most students, he would just say that name and tell them to look back on paper, you know, like stop glancing, didn't do anything, maybe move them at most. But then when it came to students who didn't really pay attention, or who rarely showed up, he took away the test and gave them a zero without like warning them. So he was favoriting, the students who were paying attention and trying and then the ones that weren't. (S5)

A few students shared four experiences of fairness when a teacher provided an opportunity for students to explain their cheating behavior. A student remembered a fair experience where a teacher allowed students to explain their cheating behavior.

I can remember an example where two students handed in the same essay to different teachers. But the teachers, you know, obviously caught on. And I remember the process was very fair, they brought both men for meetings and had conversations about what happened. They both received zero. And, you know, that was fair, and they went over what is acceptable, what's not in our code of conduct. They handled it fairly with process and students were given an opportunity to explain the situation. (S27)

Most students considered in 15 examples *giving a grade of zero* as the fairest act in instances of cheating. However, very few also considered deducting marks for the first observations of cheating followed by zero if repeated. A student reported giving zero as a fair strategy “when they catch students’ cheating, give them a straight up zero.” (S21) Another student favored deducting marks followed by a reduction to zero as fair, “I would probably start taking marks off. Like maybe like 5% first times and then 10%. And then if repeated I just like give zero on that
Most students rationalized giving zero as an equitable strategy because cheating is unfair to other students who had put effort in the work.

Yeah, I just think, like, the fact that somebody could go, say a whole school year without paying attention once and doing better than somebody who say is hyper focused and spending so much time doing everything in the class. And they're just able to do better than the other person because they're cheating. (S13)

In addition, a few students shared two experiences that considered giving zero to be beneficial to deter other students from cheating. A student stated, “it acts as a deterrent to prevent others from cheating, which makes things a lot more fair. Not to mention that if they do cheat, they should face some consequences.” (S8) While most students considered giving zero as fair, some articulated a few conditions including considering student personal life (two examples) and consequences of cheating decision (two examples) that need to be taken into account in punishing cheating. With respect to considering students’ personal life challenges, a student uttered the fairness of giving zero due to the public scrutiny of classmates in terms of teachers’ actions against cheating, but also argued for giving another opportunity to complete the work.

For example, say someone has never cheated before. They're only cheating because, something's going on at home or something. I think if that's the case, then you have to consider some factors. I think what's fair is you give them zero, because you have other people thinking that they can cheat and get away with it, right? So if you have to do that, but you do have to look at the other factors give another opportunity. (S27)

With respect to consequences of cheating decisions, students argued that while punishing students for cheating was fair, it also needed to take account of the impact of the punishment on a student’s future. A student stated, “you want students to know that it’s wrong, but you also don't want to punish them so severely that their opportunities for the future for university or college to be limited over one mistake.” (S19)
A few students also shared experiences, where they deemed teacher inaction in the face of cheating as unfair. “On the other hand, there are teachers who have caught students cheating that didn't deal with it. I guess that would be an unfair consequence.” (S12) Several students also suggested strategies such as providing multiple test formats and collecting student phones in the classroom as steps to avoid cheating. “The teacher prepared multiple formats of the test, so that they found that effective to reduce the cheating within the classroom.” (S10)

Altogether, students considered multiple principles to ensure fairness in the outcome and process for teacher decisions in punishing student cheating as well as several conditions that would help contextualize these decisions in a fairness light.

4.1.5 Theme 5: Fairness in Grading

Students recounted 45 experiences of fairness in assessment associated with grading. These experiences pertained to the codes that could ensure fairness in grading outcomes and procedures. These codes centered on (a) equity and caring in grading decisions, (b) transparency in communicating assessment criteria, (c) consistency in implementing grading criteria, (d) bias suppression in grading, (e) opportunity to appeal grades, and (f) adequate justifications for grading outcomes.

A few students leveraged equity and caring to perceive fairness of grade distributions in terms of grade adjustments for late submissions and student personal circumstances in four examples. In late submission cases, a student argued that grade reduction for punishment is unfair as it does not contribute to student learning. However, this punishment might be fair if it aims to enhance student accountability.

If the purpose of assignments is learning, rather than accountability, then you tend to say don't reduce grades because that is not fair. It may counteract in a way if student perceive it unfair, they are not gonna learn more. But on the other side, if it is about accountability,
you think that the students should feel responsible and then learn how to submit things and if they don't face the consequences, this is just for the future. (S23)

A few students also considered caring for students’ special circumstances in grade decisions as a fair. For example, a student recollected an experience where a friend did have a personal family challenge that a teacher considered in grading.

She was capable of having good grades, because her average before was really good. It was just in that one week or two that a lot of events happened in her life that were stressful. I don't think it would have been fair that she lost her entire university acceptance just because she had a bad two weeks. (S22)

Several students also perceived fairness in 9 examples, where a teacher communicated assessment expectations and criteria explicitly and followed these criteria in grading student work. A student stated “in Grade 10 math, Mr. Sam did things by the book. He was extremely fair and gave only marks according to set criteria.” (S8)

Many students shared experiences (=14) of fairness in relation to consistency in implementing assessment criteria to evaluate student work. Several of these experiences focused on consistency in applying assessment criteria across students within a classroom. For example, a student perceived unfairness since a teacher showed favoritism in grading student work.

My teacher would favor for some students, like even if they got the answers wrong, she would still mark it right. But then for other students, if they got the answers wrong, she would just mark it wrong. She did have a lot of favoritism towards marking. (S15)

Some other experiences focused on consistency in grading across teachers of the same subject in a school or across multiple schools. A student shared an experience where such inconsistency in grading could unfairly impact students’ future opportunities.

Each teacher obviously grades differently. So different teachers who run the same class grade differently. So sometimes we'll end up with situations where you could submit the
same assignment to two different teachers and we'll get two completely different marks. Everyone knows that this teachers is hard marker in comparison to other teachers that teach the same subject. And so especially when you're like 11th or 12th Grade, when you're trying to get high marks to go to university or something like that. Sometimes just knowing you got assigned to the teacher that everyone believes is the extremely hard worker, could be a really a lot like a downer on your mentality as well as like, you’re gonna get a B when you could get A in another person's class. (S4)

To redress this inconsistency, a student suggested that “I feel like maybe one thing that could happen was teachers in a department within a school collaborated on maybe marking, creating exams and tests. And it would be more fair across the different classes and teachers.” While most of students’ experiences were relevant to teachers’ grading, a few students reported experiences of inconsistency in peer grading as a result of friendship effect. A student reported,

there is one class where the teacher lets a peer mark like tests. I suppose that would be an instance where a peer was kind of being unfair maybe towards other peers because there was a bias and how they liked their friends versus other people in the class. (S12)

Another issue that undermined consistency and was relevant to a teacher’s bias in grading was concerned with including sources such as student behavior and race in the grade decisions. Four examples were shared in this respect. For example, a student reported unfair cases where a teacher reduced a group of students’ grades due to their misbehaviors.

Personally speaking, I know some of those people, they're very bright students. They just put less effort in that they don't show other people how bright they are. But like in their work, you can see how bright they are, but they still get less marks, I think it's because of the teachers’ impression on their behaviors. (S10)
Another student shared an unfair experience where a teacher considered students’ races in assigning grades. “One year in Grade 10, I had a teacher who of a certain ethnicity and she would favor students of that ethnicity, and they ended the class with higher marks.” (S15)

Several students also articulated 14 experiences in relation to how a teacher provided an opportunity for students to express their grading concerns and how teachers provided adequate justifications for grading outcomes. Once students received grades that they deemed inequitable, they would approach the teacher to hear justifications about why they received such grades. Teachers’ provision of this opportunity for students to contest a grading decision and their following approach to justify grades contributed to students’ perceived fairness. A student articulated how having an opportunity to voice grading concerns produced fair perceptions.

If there's something wrong, that's been marked wrong. I always feel like I have the opportunity to be like, hey, I don't maybe agree with this decision. Can we like go over this question and see what I've done wrong? Or see like, how, like, what happened, right? So I've always felt like there's a lot of fairness when it comes to maybe contesting a decision that a teacher made in Canada. (S24)

Ten experiences surfaced in students’ comments relevant to teachers’ adequate justifications of grades. A student reported an unfairness experience in which a teachers’ inadequate justification created an emotional distress for the student to approach the teacher.

Sometimes, like, I would go to the teacher and ask, what could I have done better? Like, why did I get this mark? At times they wouldn't really have an answer. They would just be like, that's the mark that you got, like, there isn't really anything that is wrong with it. (S25)

Overall, students considered equity and caring as the bases for their evaluation of fairness in grade distributions. They also reported multiple procedural principles such as transparency, consistency, bias suppression, voice, and justification to perceive fairness in grading procedures.
4.1.6 Theme 6: Fairness in Feedback

Students discussed fairness in feedback in 26 examples as a key domain of their perception of fairness in assessment. Specifically, students’ comments focused on three codes to perceive fairness in feedback including (a) transparent feedback, (b) timely feedback, and (c) adequate and honest feedback. Students discussed 10 experiences in relation to transparency of feedback, where a teacher did not provide a clear feedback to students. A student recounted a memory in which a teacher did not present a clear feedback into how a student could receive a A+ in the course.

I still remember my English teacher when I asked her, you know, how can I perform better on my essay, you know, how can I show a better understanding of the text? And I still remember, she told if we wanted to reach that mark, then we would have to show amazing understanding of the book, and we'd have to bring something completely new to the table. But the rubric would just say show a strong understanding of the book, but she would expect a lot more out of students. This kind of feedback was not really helpful because we weren't given a clear, you know, like a blueprint, to try and get to where we want. (S26)

Timeliness in feedback delivery was an important criterion for some students to evaluate fairness in feedback. These students shared seven experiences, where a teacher provided feedback much later than when the feedback could have been useful for students to act and learn upon. A student articulated,

often the problem was more that we would complete assignments, let's say essays or tests, and we would give them back to the teacher to be marked. And we would only receive feedback several months later. So there was no opportunity for us to improve on our skills, because by that point, we would already be finished with that unit. (S3)
Several students also uttered nine experiences with attention to whether the amount of feedback was adequate, or the feedback was honest with respect to student performance. A student recollected an experience when a teacher provided an inadequate amount of feedback to students who needed more feedback.

In Grade 11 English, I just submitted a draft to a teacher, and she would give us feedback on it. It appeared that she would give more feedback to the people who she deemed had put in more effort. So if someone had seemed like they just threw it together quickly, she'd only write like a line or two, even if it was like a worse quality that needed more feedback as opposed to someone who had like a 90. Those students were saying can we have more feedback? And she said that she had done appropriate amount of feedback based on their appropriate amount of effort. (S5)

Another student discussed that teacher honesty with feedback would help students really appreciate what they can do and need to improve.

I just think more input from teachers, like, if you if you do more stuff in like groups as the class and have the teacher be able to actually see what you're doing. And go around, say, like, check the work while you're doing it, and actually give feedback and not just do stuff that'll make you feel better if it's not gonna turn out for the better. So I think teachers just need to be more honest with the feedback. (S13)

Combined, students considered procedural and informational principles such as transparency, timeliness, honesty, and adequacy of feedback to perceive fairness.

4.1.7 Theme 7: Socio-emotional Environment

Students discussed the key role socio-emotional environment plays in their perception of fairness in classroom assessment in 39 experiences. Students articulated multiple codes that contributed to a fair environment: (a) teacher-student relationships, (b) gender equity, (c) racial equity, and (d) peer relationships. Positive teacher-student relationships embodied through
respect and rapport were articulated in 20 experiences that contributed to a fair socio-emotional environment for student assessment and learning. A student recollected a fair experience where a teacher’s respectful behavior created a constructive socio-emotional environment.

I had a teacher who taught us business. I always thought the way he interacted with us was very fair. At the beginning of each class, he would take into consideration how our day went like he would allocate 10 to 15 minutes off because we were doing each individual person, right, because he wanted to see who his students were. He was so polite and respectful towards everyone. (S12)

Students also recounted unfair experiences in teacher-student relationships that undermined positive socio-emotional environment. Unlike the above example, these experiences revealed that a classroom environment could be perceived differently across students; it could be perceived as fair to a student but unfair to another depending on the nature of teacher-student relationships.

Assessment not only plays a significant role in creating this un/fair socio-emotional environment for students but is also subsequently impacted by the created socio-emotional environment. Several students articulated experiences where a teacher’s assessment practice did shape their expectation of a student during the course, which subsequently impacted the socio-emotional relationships across different groups of students. For instance, a student articulated,

another thing that happens often is that the students who are already succeeding, have a good relationship with the teacher. And so they’re more likely to ask questions. To go in after class, and have a conversation with the teacher, whereas the students who are not really succeeding, they're not doing very well. There's a bit of like, they feel a bit intimidated, let's say by the teacher. And so they're less likely to ask for help when they need it. Because there's they feel scared to admit maybe that they're not succeeding, or they feel embarrassed by the grade that they got on the last assessment. And so they're too nervous to approach the teacher. (S11)
Being aware of teachers’ expectations, several students would then consciously devise a behavioral approach that would culminate in positive teacher-student relationships and assessment results. For example, a student stated,

one thing I'd say just so you can get kind of like an image of me as a student, I would have probably been the student that was favored for, and I was always kind of like, my friends would make fun of me too, because I'd always kind of suck up to the teacher to get better grades. It's almost like a spiral like you make your first impressions while at the start of the year. And then you're able to kind of build on that for the rest of the year and typically get going well. (S2)

Being unaware of such teacher expectation effects and the psychology behind it, there were students who may seemed to be the victim of unfair outcomes in this cyclical relationship between a teacher and a student. In this respect, a student expressed,

a lot of my bad experiences stem from teachers that didn't give me the same chance as they did for other students just because the other students maybe had more social skills or friendlier, but maybe I was more quiet. I got shafted by the teacher. You know what I mean? Because it's frustrating. You see the teacher being nicer to other people. And it's also harder to learn. (S7)

Another motivator for being a victim in this cyclical relationship might result from students’ misbehavior. A student shared from a third-party lens this perception of unfairness in the cyclical relationships between a misbehaved classmate and a teacher.

It is a cycle because like if the student starts out misbehaving, then the teacher doesn't really like him. And then the teacher can start to mistreat him. And then if the teacher is mistreating him, then the student begins to dislike the class and begins to become rebellious and like misbehave again. So it just goes back and forth. (S25)
In addition to assessment results and behavior, students considered gender and race as additional signals for unfair teacher-student relationships. For example, several students discussed seven experiences, where a male teacher had favoritism toward a female student. A female student recollected,

I have definitely seen some gender bias. I had one teacher who was heavily biased towards girls. It was kind of weird, but like, if a girl had a question about one of the questions, if she asked the teacher, he would just give her the answer. Um, but if a guy did it, no help. I had this teacher two academic years, like the entire year, he would just be like, super aloof to males in the class, and not really be helpful to them. But then girls, he would be like, super. (S17)

Several students also reported eight experiences of unfairness due to teachers’ consideration of student race in their relationships and grades. A student stated,

I always felt like my teachers, not all of them, but they had favorites in the class, and like, well, that they've talked to more and people that they would maybe know outside of school, and they would be like, give better grades to them. Even though everybody was doing the same amount of work. A lot of teachers like, obviously they wouldn't say this out loud, but like you could tell that they were like racist to so like, they would like give preference to people who are like them. (S14)

While students considered teacher-student relationships to constitute a fair socio-emotional environment, they also highlighted peer relationships and its impact on student assessment and learning. Students shared four experiences of unfairness where a group of students would attempt to alienate a student that subsequently impacted their assessment performance. The next two examples represent different forms of this alienation.

It's not the type of extreme bullying that you would expect, but just ways that certain groups of students would alienate one other student, and then that would cause them to
like sort of retreat into a bubble outside of the classroom, but also in the classroom. So it definitely makes an impact on their, like participation and their ability to learn in the classroom. Because if you're focused on like, if you're sitting in the classroom, and there's a group of students on the other end of the room, who keep whispering and looking in your direction, it's hard to focus on the learning and actually improve yourself. (S19)

There are a lot of times that people who are new or like, aren’t smart about stuff will get targeted and harassed and stuff by the other people just because they don't understand something as much as the other people. So like, say if somebody's more confused about something that someone else thinking I've had people to me have feel they have such a superiority complex and be on such a high horse because I'd be confused about something. And they try and make me feel lesser, because they are doing slightly better (S13, student with IEP).

Overall, students articulated teacher-student relationships, gender and racial equity, and peer relationships as codes that influenced fairness of socio-emotional environment in classroom assessment.

4.1.8 Theme 8: Responses to Perceptions of Un/fairness

Sharing 50 experiences, students discussed multiple responses that they showed in reaction to perceived fairness in classroom assessment. These responses were summarized into following codes: (a) learning and motivation outcomes, (b) emotional responses, (c) dropout from the course, (d) talk with school leadership, teachers, parents, and friends; (e) rationalization for perseverance, and (f) inaction. Students reported in 10 examples that teachers’ fair actions positively contributed to their motivation and learning outcomes. A student noted a fair experience: “definitely affects my learning positively. Because if you can build a relationship with your teacher where there's a sense of fairness and equality you're more open to receiving
information if that makes sense.” (S24) As a negative response to an unfair experience, a student suggested that

I would say it impacted my attitude towards the class greatly because it made me dislike the teacher and made me less wanted or less willing to participate in class because I didn't think that my opinion would be respected since he kind of showed that he didn't care about my opinion earlier. (S17)

Students also showed positive and negative emotions in response to teacher un/fairness in classroom assessment. Overall, students reported 14 emotional-related experiences. In some of these experiences, students felt happy, satisfied, hopeful, and confident when they perceived fairness. “when things are fair and open and transparent, I'd like I feel like positive about it. Like I feel like happy about it”. (S24) In some others, students felt upset, frustrated, and angry in response to perceived unfairness. A student recollected how a teacher’s unfairness provoked a particular image of her.

I felt she kind of made me feel like stupid. As of my grades weren't reflecting that I was smart, so she made me feel like, not smart. And I felt like I was super annoyed. I felt discouraged. I didn't feel motivated. (S14)

Dropout from the course was another reaction to unfairness perception that was shared in three examples. For instance, a few third-party observant male students selected to drop a course with a teacher who was enacting favoritism toward male students. While this was in favor of their self-interest, their perceptions of fairness, as reported by the following quote, drove their choice to drop out.

The boys really didn't like it. I think that might have been the reason why some of the boys did not continue in the class. So a lot of those people after they had that experience, it might have been related to other things, but a lot of them just stopped taking French
altogether because they knew, like there's a chance they would get that teacher again, and then it wouldn't be worth it. (S17)

When students perceived unfairness, they started talking with a principal, teacher, friends, and parents as a way to redress unfairness and its outcomes. Ten examples were related to this code. The following examples represent talking with the principal, friends, and parents to redress unfairness.

I just did not agree with how she would mark the work that I'd given to her, because I put a lot of effort into it. I just had a lot of trouble with this teacher. I kind of didn't have the greatest experience with this teacher, my vice principal decided to switch my English teacher to another English teacher, which in my opinion marked me more objectively. So that was a more fair experience in my opinion. (S26)

So I talked with my friends. I did talk with my parents, but I don't think anyone want to complain afterwards. I mean, I just kind of accepted it. (S12)

A few students also articulated rationalizations for perseverance as coping mechanisms with unfairness experiences. Four examples were shared for this code. While these articulations showed that unfair experiences would have longer-term mental burdens on student well-being, they could also be used for learning about perseverance. However, the question was to what extent the unfairness might provoke positive learning outcomes such as perseverance as also highlighted by students in the following quotations. These quotations also represent the impact of social structures on students’ perceived fairness.

I believe in life, you cannot go back in time but move forward with the situation that you have been given, even if it may not be classified as fair [italicized added to emphasize the influence of social structures]. (S23)

Although I may be white, I may be male, I may be born into a first world country with loving parents and a stable household and parents who make good money and you may
believe that I would have no problems with the amount of fairness in my life; I like a lot of people have dealt with unfairness. As someone who has struggled with ADHD and other mental/learning disabilities for the better half of my life I have seen and dealt with a lot of unfairness in my life. I was always seen as the bad kid who had no personal space boundaries, who couldn’t sit still, who always said the wrong thing at the wrong time, who always forgot to raise his hand and was always a bother, especially in the classroom but in my daily life. But against any odds I faced obstacles I had to overcome, I was still able to do incredibly well through elementary to high school, I was able to move and graduate one of the toughest, most competitive and most demanding programs at Niagara college and receive an advanced diploma, and now in university, on my way to getting my bachelor’s degree. I can say that I am proud of where I am and how much I’ve already accomplished despite that. *Power to those who can overcome adversaries and fight against their controlled and uncontrollable obstacles in the hand and still be able to persevere!* [italicized added to emphasize the influence of social structures and the hesitation of being able to persevere!] (S18)

While students reported multiple experiences where they showed their agency in reacting to perceived unfairness, they articulated nine experiences where they selected *inaction*. Students seemed to justify this inaction by reframing cognitively that an unfair outcome did not deserve reaction. When these students weighed the stakes of reaction against the drive to respond to unfairness, they cognitively chose to rationalize their inaction. The following two examples elucidate this process.

So for me back in grade 10, it was like, grade 10 doesn't really matter as much as like grade 11 or 12. Like in grade 11 and 12, I argued harder for marks that I would have in grade 9 and 10. And for me in chemistry, I didn't really need my chemistry mark to get me anywhere. So I didn't really see the point in arguing. (S15)
The pass is pass, 10 years later no one will care about my high school mark. (S1)

Combined, students reported multiple responses and consequences in the face of un/fair perceptions. They specifically alluded to psychological and behavioral responses including learning and motivation outcomes, emotional responses, dropout from the course, talks with school leadership, teachers, parents, and friends, and inaction.

4.2 Findings for Phase II

This section presents the results from the psychometric analyses and validity evidences for internal structure and relationship to other variables including the results from factor analyses and structural equation modeling. Specifically, it first presents the descriptive statistics showing students’ responses to Classroom Assessment Fairness Inventory. Then, the results of exploratory factor analysis are presented, followed by confirmatory factor analysis to identify the best fitting empirical structure with the data and conceptual model. It then outlines the results for structural equation modeling, indicating the relationships between Classroom Assessment Fairness Inventory and Personal Belief in a Just World Survey.

4.2.1 Descriptive Statistics for Classroom Assessment Fairness Inventory

The descriptive statistics for students’ responses to Classroom Assessment Inventory is presented in Tables 13 to 17. The 6 Likert-point scale was collapsed into 4 (in addition to I don’t know option) only in descriptive statistics presentation to facilitate interpreting students’ perceptions of fairness for each action item. In the groupwork scenario, participants largely agreed with group composition based on mixed abilities, and they were marginally aligned with fairness of a teacher’s action in not providing choice for students to select group members. Participants largely agreed with fairness of considering students’ complaints over group dynamics. Participants also largely considered it as unfair that a teacher was undetailed about grading criteria, gave equal grades to all, and did not justify students’ grades.
In the exam scenario, participants tended to agree with fairness of a teacher’s firm decision to continue with exam schedule as planned. While participants considered inexplicit communication of exam content as largely unfair, they tended to perceive including easy-to-difficult items in the exam as fairer as well as giving accommodations to students with disability and English language learners. Participants agreed with the fairness of lenient grading; they, however, perceived a teacher’s decision to keep surprise items and to respond to students’ complaints with a disrespectful tone. Participants also marginally tended to recognize fairness in giving second exam chance to those who missed it.

In the cheating scenario, participants selected unfairness due to untransparent cheating policy and failure to provide opportunity to students to explains their cheating. However, they supported fairness of giving a grade of zero. Asking the student caught in a cheating act to leave the classroom in front of others was received as largely unfair, while a teacher’s effort to explain her cheating decision was regarded as fairer. Participants were divided whether a teacher should forgive a student with a cheating act, with 51.6% agreeing with it, 41% disagreeing, and 7.4% selecting ‘did not know’ option.

In the grading scenario, several actions items were highly rated as fair including (a) the respectful treatment during classroom assessment, (b) consideration of student achievement in students’ grades, (c) detailed communication of assessment criteria and timely communication of assessment results, (d) provision of an opportunity for students to discuss their grades, and (e) provision of teacher justifications for students who appealed grades. 69% of participants supported adjusting at-risk students’ grades, 55.6% supported adjusting grades for university admissions, and 70.9% disagreed with fairness of considering student misbehavior in grading.

In the feedback scenario, transparent and timely communication of feedback as well as provision of opportunity for students to discuss feedback were perceived highly as fair. Participants were divided in half in their dis/agreements with provision of feedback based on
amount of student effort, with most participants perceiving lack of feedback directed to weaker students as unfair. Failure to respect students in the feedback process, provision of feedback to favorites, and a teacher’s rationale to give feedback based on the amount of effort were largely perceived as unfair.

Table 14. Descriptive statistics for students' responses to the first scenario

<table>
<thead>
<tr>
<th>Scenario 1- Groupwork</th>
<th>Unfair</th>
<th>Somewhat Unfair</th>
<th>Somewhat Fair</th>
<th>Fair</th>
<th>I don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>1. Mr. Chu selected group members based on mixed ability</td>
<td>15.2</td>
<td>12.9</td>
<td>25.8</td>
<td>44.2</td>
<td>1.8</td>
</tr>
<tr>
<td>2. Mr. Chu did not provide students a choice in selecting their group members.</td>
<td>13.9</td>
<td>28.7</td>
<td>25</td>
<td>30.1</td>
<td>2.3</td>
</tr>
<tr>
<td>3. Mr. Chu allowed students’ complaints over group dynamics.</td>
<td>10.2</td>
<td>9.7</td>
<td>22.7</td>
<td>51.9</td>
<td>5.6</td>
</tr>
<tr>
<td>4. Mr. Chu was not detailed in communicating how he will assess students’ groupwork.</td>
<td>70.1</td>
<td>15.7</td>
<td>3.2</td>
<td>8.8</td>
<td>2.3</td>
</tr>
<tr>
<td>5. Mr. Chu gave the same grades to all group members.</td>
<td>53.5</td>
<td>21.7</td>
<td>12.4</td>
<td>11.1</td>
<td>1.4</td>
</tr>
<tr>
<td>6. Mr. Chu did not give individual grades for each group member based on their contributions and learning.</td>
<td>63.1</td>
<td>19.8</td>
<td>8.3</td>
<td>7.4</td>
<td>1.4</td>
</tr>
<tr>
<td>7. Mr. Chu did not justify his grades to students who appealed.</td>
<td>80.6</td>
<td>9.3</td>
<td>3.2</td>
<td>5.6</td>
<td>.9</td>
</tr>
</tbody>
</table>

Table 15. Descriptive statistics for students' responses to the second scenario

<table>
<thead>
<tr>
<th>Scenario 2- Exam</th>
<th>Unfair</th>
<th>Somewhat Unfair</th>
<th>Somewhat Fair</th>
<th>Fair</th>
<th>I don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>1. Mr. Ahmed held firm on the exam date.</td>
<td>16.1</td>
<td>16.6</td>
<td>21.7</td>
<td>43.3</td>
<td>2.3</td>
</tr>
<tr>
<td>2. Mr. Ahmed did not explicitly state what would be on the exam.</td>
<td>40.3</td>
<td>28.2</td>
<td>14.4</td>
<td>16.2</td>
<td>.9</td>
</tr>
<tr>
<td>3. Mr. Ahmed included in the exam both easy to difficult questions.</td>
<td>1.9</td>
<td>2.8</td>
<td>13.8</td>
<td>78.8</td>
<td>2.8</td>
</tr>
<tr>
<td>4. Students with disabilities and English language learners received accommodations for the exam (e.g., more writing time).</td>
<td>2.8</td>
<td>5.6</td>
<td>4.6</td>
<td>84.7</td>
<td>2.3</td>
</tr>
<tr>
<td>5. Mr. Ahmed graded his students more leniently than other teachers.</td>
<td>9.3</td>
<td>13.9</td>
<td>29.2</td>
<td>38</td>
<td>9.7</td>
</tr>
</tbody>
</table>
6. Mr. Ahmed did not remove the two questions on the content that were not taught before.

7. Mr. Ahmed did not respond to students’ complaints with a respectful tone.

8. Mr. Ahmed gave another exam opportunity to students who missed the exam.

<table>
<thead>
<tr>
<th>Scenario 3- Cheating</th>
<th>Unfair</th>
<th>Somewhat Unfair</th>
<th>Somewhat Fair</th>
<th>Fair</th>
<th>I don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ms. Johnston was not transparent about her cheating policy at the beginning of the course.</td>
<td>55.8</td>
<td>24.4</td>
<td>8.8</td>
<td>8.8</td>
<td>2.3</td>
</tr>
<tr>
<td>Ms. Johnston did not give the student an opportunity to explain the reasons for cheating before making her decision to give zero.</td>
<td>43.8</td>
<td>22.6</td>
<td>12.9</td>
<td>19.8</td>
<td>.9</td>
</tr>
<tr>
<td>Ms. Johnston gave a zero because cheating is unfair to other students’ efforts.</td>
<td>8.3</td>
<td>10.7</td>
<td>17.7</td>
<td>61.8</td>
<td>1.4</td>
</tr>
<tr>
<td>Ms. Johnston gave a zero because cheating is ethically wrong.</td>
<td>8.8</td>
<td>7.8</td>
<td>18.9</td>
<td>60.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Ms. Johnston gave a zero to signal that anyone who cheats should be punished for this action.</td>
<td>22.1</td>
<td>12</td>
<td>16.6</td>
<td>45.6</td>
<td>3.7</td>
</tr>
<tr>
<td>Ms. Johnston asked the student to leave the classroom in front of other students.</td>
<td>46.6</td>
<td>21.9</td>
<td>10.2</td>
<td>16.8</td>
<td>4.7</td>
</tr>
<tr>
<td>Ms. Johnston explained her cheating decision to the student.</td>
<td>6</td>
<td>2.3</td>
<td>24.1</td>
<td>64.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Ms. Johnston did not forgive the student’s cheating this time.</td>
<td>19.1</td>
<td>21.9</td>
<td>20.9</td>
<td>30.7</td>
<td>7.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 4- Grading</th>
<th>Unfair</th>
<th>Somewhat Unfair</th>
<th>Somewhat Fair</th>
<th>Fair</th>
<th>I don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ms. Mendes treated students respectfully during classroom assessment.</td>
<td>2.8</td>
<td>4.6</td>
<td>7.8</td>
<td>82</td>
<td>2.8</td>
</tr>
<tr>
<td>Ms. Mendes largely considered student achievement of learning objectives in her grading.</td>
<td>2.8</td>
<td>6.5</td>
<td>9.7</td>
<td>79.1</td>
<td>1.9</td>
</tr>
</tbody>
</table>
3. Ms. Mendes detailed her grading criteria, with test scores making up 70% of a student’s grade.

4. Ms. Mendes communicated test results in one week after handing the test in.

5. Ms. Mendes would sometimes allow students to discuss their grades if there was enough time in class.

6. Ms. Mendes gave adequate justification for students who spoke up for their grades.

7. Ms. Mendes adjusted the grades of failing students with at-risk backgrounds.

8. Ms. Mendes considered students’ future university admissions to adjust grades.

9. Ms. Mendes considered student misbehavior (e.g., disruptions) in her grading.

<table>
<thead>
<tr>
<th>Scenario 5- Feedback</th>
<th>Unfair %</th>
<th>Somewhat Unfair %</th>
<th>Somewhat Fair %</th>
<th>Fair %</th>
<th>I don’t know %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mr. Dembe provided a clear rubric for assessing students’ essays.</td>
<td>2.8</td>
<td>4.2</td>
<td>6</td>
<td>84.7</td>
<td>2.3</td>
</tr>
<tr>
<td>2. Mr. Dembe provided feedback after four days of essay submissions.</td>
<td>5.6</td>
<td>3.7</td>
<td>8.4</td>
<td>84.1</td>
<td>1.9</td>
</tr>
<tr>
<td>3. Mr. Dembe gave students a chance to further discuss his feedback.</td>
<td>1.9</td>
<td>3.8</td>
<td>4.2</td>
<td>88.3</td>
<td>1.9</td>
</tr>
<tr>
<td>4. Mr. Dembe provided feedback based on the amount of effort each student has put in the work.</td>
<td>33.5</td>
<td>16.3</td>
<td>15.8</td>
<td>32.6</td>
<td>1.9</td>
</tr>
<tr>
<td>5. Mr. Dembe did not provide more feedback to students who had weaker performance.</td>
<td>67</td>
<td>17.2</td>
<td>7.9</td>
<td>7</td>
<td>.9</td>
</tr>
<tr>
<td>6. Mr. Dembe did not treat students respectfully in his feedback procedure.</td>
<td>75.1</td>
<td>10.3</td>
<td>5.6</td>
<td>7.5</td>
<td>1.4</td>
</tr>
<tr>
<td>7. Mr. Dembe gave more feedback to his favorite students.</td>
<td>84.6</td>
<td>5.6</td>
<td>2.8</td>
<td>6.1</td>
<td>.9</td>
</tr>
</tbody>
</table>

123
8. Mr. Dembe explained that his feedback procedure was based on the amount of effort each student has put in.

4.2.2 Exploratory Factor Analysis

The initial scrutiny of results for exploratory factor analysis showed 10 factors with Eigen values greater than 1. The scree plot is presented in Figure 4. To find the best factor structure pattern, Monte Carlo PCA for Parallel Analysis was examined, supporting models with four to six factors (explaining 46-54% of the variance) with eigenvalues greater than the corresponding criterion values in a random data, generated with a matrix of the same size (40 variables × 217 participants). Each model was then analyzed closely based on the item loadings within and across factors and the meaning of factors in light of the conceptual model. Based on this analysis, two models were deemed to have stronger psychometric foundation and resemblance to the initial conceptual model: the five-factor model and the six-factor model. These models somewhat reflected the coherence of action items within the relevant scenarios as initially conceptualized. Given that the only difference between the five and six factor models was an additional factor that included a few items pertaining to the second scenario (S2-1, S2-2, S2-6, S2-7), the five-factor model was considered as the best meaningful factor solution for the current sample. Eight items (S1-1, S1-2, S1-3, S2-1, S2-3, S2-4, S2-5, S2-8) were removed as they either did not load significantly beyond .3 on any of the five factors or were uninterpretable in their grouping with other action items representing a factor. The five-factor model explained 50.4% of variance in the data. This model is presented in Table 18.
Figure 4. Scree plot for the Classroom Assessment Fairness Inventory

Table 19. Five-factor model using maximum likelihood with direct oblimin rotation

<table>
<thead>
<tr>
<th>Actions</th>
<th>Unfairness in Groupwork and Exam</th>
<th>Fairness in Cheating</th>
<th>Fairness in Grading</th>
<th>Unfairness in Feedback</th>
<th>Fairness in Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1_4. Mr. Chu was not detailed in communicating how he will assess</td>
<td>0.624*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>students’ groupwork.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1_5 Mr. Chu gave the same grades to all group members.</td>
<td>0.650*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1_6 Mr. Chu did not give individual grades for each group member</td>
<td>0.700*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>based on their contributions and learning.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1_7 Mr. Chu did not justify his grades to students who appealed.</td>
<td>0.561*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2_2 Mr. Ahmed did not explicitly state what would be on the exam.</td>
<td>0.381*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2_6 Mr. Ahmed did not remove the two questions on the content that</td>
<td>0.600*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>were not taught before.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2_7 Mr. Ahmed did not respond to students’ complaints with a</td>
<td>0.636*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>respectful tone.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_1 Ms. Johnston was not transparent about her cheating policy at the</td>
<td>0.387*</td>
<td>0.421*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>beginning of the course.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_2 Ms. Johnston did not give the student an opportunity to explain</td>
<td>0.388*</td>
<td>0.543*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the reasons for cheating before making her decision to give zero.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_3 Ms. Johnston gave a zero because cheating is unfair to other</td>
<td>0.764*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>students’ efforts.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_4 Ms. Johnston gave a zero because cheating is ethically wrong.</td>
<td></td>
<td></td>
<td></td>
<td>0.694*</td>
<td></td>
</tr>
<tr>
<td>S3_5 Ms. Johnston gave a zero to signal that anyone who cheats should</td>
<td></td>
<td></td>
<td>0.746*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>be punished for this action.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Statement</td>
<td>Variance Explained</td>
<td>Cronbach Alpha</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_6</td>
<td>Ms. Johnston asked the student to leave the classroom in front of other students.</td>
<td>0.300*</td>
<td>.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_7</td>
<td>Ms. Johnston explained her cheating decision to the student.</td>
<td>0.541*</td>
<td>.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_8</td>
<td>Ms. Johnston did not forgive the student’s cheating this time.</td>
<td>0.583*</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_1</td>
<td>Ms. Mendes treated students respectfully during classroom assessment.</td>
<td>0.703*</td>
<td>.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_2</td>
<td>Ms. Mendes largely considered student achievement of learning objectives in her grading.</td>
<td>0.611*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_3</td>
<td>Ms. Mendes detailed her grading criteria, with test scores making up 70% of a student’s grade.</td>
<td>0.542*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_4</td>
<td>Ms. Mendes communicated test results in one week after handing the test in.</td>
<td>0.616*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_5</td>
<td>Ms. Mendes would sometimes allow students to discuss their grades if there was enough time in class.</td>
<td>0.662*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_6</td>
<td>Ms. Mendes gave adequate justification for students who spoke up for their grades.</td>
<td>0.657*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_7</td>
<td>Ms. Mendes adjusted the grades of failing students with at-risk backgrounds.</td>
<td>0.648* 0.367*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_8</td>
<td>Ms. Mendes considered students’ future university admissions to adjust grades.</td>
<td>0.526* 0.566*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_9</td>
<td>Ms. Mendes considered student misbehavior (e.g., disruptions) in her grading.</td>
<td>0.449* 0.477*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5_1</td>
<td>Mr. Dembe provided a clear rubric for assessing students’ essays.</td>
<td>0.308* 0.532*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5_2</td>
<td>Mr. Dembe provided feedback after four days of essay submissions.</td>
<td>0.745*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5_3</td>
<td>Mr. Dembe gave students a chance to further discuss his feedback.</td>
<td>1.001*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5_4</td>
<td>Mr. Dembe provided feedback based on the amount of effort each student has put in the work.</td>
<td>0.629*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5_5</td>
<td>Mr. Dembe did not provide more feedback to students who had weaker performance.</td>
<td>0.597*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5_6</td>
<td>Mr. Dembe did not treat students respectfully in his feedback procedure.</td>
<td>0.400* 0.457*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5_7</td>
<td>Mr. Dembe gave more feedback to his favorite students.</td>
<td>0.543*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5_8</td>
<td>Mr. Dembe explained that his feedback procedure was based on the amount of effort each student has put in.</td>
<td>0.731*</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Extraction method: Maximum Likelihood. Rotation method: Oblimin with Kaiser Normalization. Factor loadings lower than .3 were suppressed. The actions with $p < .05$ are shown with *.  
126
The first factor included seven items and was termed *unfairness in groupwork and exam*. Lack of detailed communication of assessment process in the group work, assigning the same grade for all group members, lack of grade justification were items from the groupwork scenario in this factor. Lack of transparent communication of exam content, failure to remove surprise items and taking into account students’ complaints with a respectful tone were additional items from the exam scenario loading on this factor. Given that teachers in the groupwork and exam scenarios were violating a justice principle in these seven items, students perceived these items as unfair. The second factor included eight items and was termed *fairness in cheating*. Items focused on untransparent cheating policy, lack of student opportunity to explain cheating reasons, assignment of a grade of zero, student removal from the class in front of other students, teacher explanation of a cheating decision, and failure to forgive a student cheating. This factor included all the items associated with fairness in cheating.

The third factor included nine items and was labeled *fairness in grading* in parallel with its conceptual scenario. Items in this factor contained respectful treatment of student in classroom assessment, consideration of student achievement in grading, detailed communication of grading criteria, timely announcement of grades, provision of an opportunity for students to discuss their grades, a teacher provision of grade justifications, adjustment of grades for at-risk students, for university admissions, and for student misbehavior. There were a few items with strong cross-loading in the third factor. However, the decision to retain the action items in this factor was based on their conceptual focus on measuring fairness in grading. The fourth factor (*unfairness in feedback*) included five items and the fifth factor (*fairness in feedback*) included three items. *Unfairness in feedback* factor represented items where a teacher violated the justice principles in relation to providing feedback comparable to student effort, disregarding weaker students’ needs for additional feedback, disrespecting in feedback process, giving more feedback to favorites. In *fairness in feedback* factor, the teacher adhered to justice principles in relation to offering
transparent and timely feedback as well as giving students a chance to discuss their feedback. The reliability analysis showed high indices for unfairness in groupwork and exam ($\alpha = .79$), fairness in cheating ($\alpha = .83$), fairness in grading ($\alpha = .82$), unfairness in feedback ($\alpha = .83$), and fairness in feedback ($\alpha = .87$). Taken together, the results from exploratory factor analysis signify students’ perceived fairness in classroom assessment in relation to these five factors.

Correlations across factors were analyzed to understand the interrelationship of factors. Positive correlations were observed between unfairness in groupwork and exam and unfairness in feedback ($r = .25, p < .05$) while negative correlations were found between unfairness in groupwork and exam and fairness in grading ($r = -.18, p < .05$) and fairness in feedback ($r = -.20, p < .05$). Fairness in cheating showed positive correlation with fairness in grading ($r = .21, p < .05$) and fairness in feedback ($r = .24, p < .05$). Fairness in grading had negative correlation with unfairness in feedback ($r = -.17, p < .05$) and positive correlation with fairness in feedback ($r = .48, p < .05$). Finally, unfairness in feedback had a negative correlation with fairness in feedback ($r = -.18, p < .05$). Table 19 presents these correlational relationships.

### Table 20. Correlations across factors of the Classroom Assessment Fairness Inventory

<table>
<thead>
<tr>
<th></th>
<th>Unfairness in Groupwork and Exam</th>
<th>Fairness in Cheating</th>
<th>Fairness in Grading</th>
<th>Unfairness in Feedback</th>
<th>Fairness in Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfairness in Groupwork and Exam</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairness in Cheating</td>
<td>0.026</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairness in Grading</td>
<td>-0.182*</td>
<td>0.211*</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfairness in Feedback</td>
<td>0.257*</td>
<td>-0.041</td>
<td>-0.178*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Fairness in Feedback</td>
<td>-0.207*</td>
<td>0.242*</td>
<td>0.482*</td>
<td>-0.189*</td>
<td>1</td>
</tr>
</tbody>
</table>

**4.2.3 Confirmatory Factor Analysis**
A confirmatory factor analysis was conducted to confirm the factor structure derived from the exploratory factor analysis. The five-factor and six-factor models derived from exploratory factor analysis were subjected to confirmatory factor analysis using Maximum Likelihood estimator. The initial results did not show a satisfactory fit for both five- and six-factor models and goodness-of-fit indices including RMSEA < .08, CFI > .95, and TLI > .95. The standardized regression weights for each item were therefore scrutinized to identify items that did not load onto the relevant factors strongly. The items with weaker loading were removed. The additional factor in the six-factor model only retained two action items from Scenario 2 (i.e., S2-6 and S2-7).

Consistent with the exploratory factor analysis results, the five-factor model showed a better fit with the data. The overall goodness-of-fit indices for the five-factor model showed RMSEA < .06 (90% CI [.05-.07]), CFI > .91, and TLI > .89, indicating a satisfactory fit with the data (see Table 20). While the chi-square test of model fit, $\chi^2(179, N = 217) = 217.5, p < .001$, was significant, indicating a poor fit; however, the chi-square test is stringent, requiring a perfect model fit (Brown, 2006).

Table 21. Goodness of fit and comparative indices for confirmatory factor analysis

<table>
<thead>
<tr>
<th>Solution</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA (LO-HI 90)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-factor</td>
<td>363.558</td>
<td>179</td>
<td>.91</td>
<td>.89</td>
<td>.06 (.05-.07)</td>
</tr>
<tr>
<td>6-factor</td>
<td>449.256</td>
<td>215</td>
<td>.90</td>
<td>.87</td>
<td>.07 (.06-.08)</td>
</tr>
</tbody>
</table>

The results of standardized factor loadings for a five-factor model are presented in Table 21. The first factor with three items was labeled *unfairness in groupwork* as it included items referring to violation of justice principles in groupwork assessment. The second factor with five items was labeled *fairness in cheating* as it included items relevant to fairness in the cheating scenario. The third factor with six items was labeled *fairness in grading* in line with its paralleled
conceptual scenario. The fourth factor (*unfairness in feedback*) included four items and the fifth factor (*fairness in feedback*) included three items. These factors were labeled based on either violation of or adherence to justice principles in the feedback scenario. The reliability analysis showed high indices for *unfairness in groupwork* ($\alpha = .79$), *fairness in cheating* ($\alpha = .82$), *fairness in grading* ($\alpha = .85$), *unfairness in feedback* ($\alpha = .84$), and *fairness in feedback* ($\alpha = .87$). The confirmatory factor analysis yielded significant correlations across latent factors of Classroom Assessment Fairness Inventory. *Unfairness in groupwork* showed positive correlation with *unfairness in feedback* ($r = .37, p < .001$); *fairness in cheating* showed positive correlation with *fairness in grading* ($r = .38, p < .001$) and *fairness in feedback* ($r = .32, p < .001$). *Fairness in grading* had negative correlation with *unfairness in feedback* ($r = -.34, p < .001$), and positive correlation with *fairness in feedback* ($r = .65, p < .001$). Finally, *unfairness in feedback* had a negative correlation with *fairness in feedback* ($r = -.36, p < .001$). These correlational relationships offer support for the conceptual separation of each factor.

Table 22. Standardized factor loadings from confirmatory factor analysis

<table>
<thead>
<tr>
<th>Actions</th>
<th>Unfairness in Groupwork</th>
<th>Fairness in Cheating</th>
<th>Fairness in Grading</th>
<th>Unfairness in Feedback</th>
<th>Fairness in Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1_5 Mr. Chu gave the same grades to all group members.</td>
<td>0.544</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1_6 Mr. Chu did not give individual grades for each group member</td>
<td>0.903</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>based on their contributions and learning.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1_7 Mr. Chu did not justify his grades to students who appealed.</td>
<td>0.544</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_3 Ms. Johnston gave a zero because cheating is unfair to other</td>
<td></td>
<td>0.792</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>students’ efforts.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_4 Ms. Johnston gave a zero because cheating is ethically wrong.</td>
<td></td>
<td>0.823</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_5 Ms. Johnston gave a zero to signal that anyone who cheats should</td>
<td></td>
<td>0.772</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>be punished for this action.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_7 Ms. Johnston explained her cheating decision to the student.</td>
<td></td>
<td>0.596</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3_8 Ms. Johnston did not forgive the student’s cheating this time.</td>
<td></td>
<td>0.516</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_1 Ms. Mendes treated students respectfully during classroom</td>
<td></td>
<td></td>
<td></td>
<td>0.808</td>
<td></td>
</tr>
<tr>
<td>assessment.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4_2 Ms. Mendes largely considered student achievement of learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.729</td>
</tr>
<tr>
<td>objectives in her grading.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ms. Mendes detailed her grading criteria, with test scores making up 70% of a student’s grade.

Ms. Mendes communicated test results in one week after handing the test in.

Ms. Mendes would sometimes allow students to discuss their grades if there was enough time in class.

Ms. Mendes gave adequate justification for students who spoke up for their grades.

Mr. Dembe provided a clear rubric for assessing students’ essays.

Mr. Dembe provided feedback after four days of essay submissions.

Mr. Dembe gave students a chance to further discuss his feedback.

Mr. Dembe did not provide more feedback to students who had weaker performance.

Mr. Dembe did not treat students respectfully in his feedback procedure.

Mr. Dembe gave more feedback to his favorite students.

Mr. Dembe explained that his feedback procedure was based on the amount of effort each student has put in.

### Cronbach Alpha

<table>
<thead>
<tr>
<th></th>
<th>.79</th>
<th>.82</th>
<th>.85</th>
<th>.84</th>
<th>.87</th>
</tr>
</thead>
</table>

#### 4.2.4 Structural Equation Modeling

A full structural equation modeling was conducted using Maximum Likelihood estimator to examine the relationships across factors of Classroom Assessment Fairness Inventory and Personal Belief in a Just World survey in order to provide additional validity evidence. Specifically, a direct path was specified from the latent factor of Personal Belief in a Just World survey predicting the five latent factors of Classroom Assessment Fairness Inventory. Three items (items 1, 3, 7) were removed from Personal Belief in a Just World survey given their lower loadings. The reliability analysis showed high internal consistency index for personal belief in a just world survey (\( \alpha = .85 \)).

The chi-square test of model fit, \( \chi^2(270, N = 217) = 644.997, p < .001 \), was significant. While the RMSEA < .080 (90% CI [.072-.088]) showed acceptable fit of the model, the other two indices including CFI > .86, and TLI > .844 showed values less than acceptable fit with the
data. It appears that small sample size of this study partly limited reaching adequate model fit. Figure 5 shows two statistically significant paths; students’ personal belief in a just world yielded significant positive correlations with fairness in grading ($\beta = .32, p < .001$), and with fairness in feedback ($\beta = .31, p < .001$).

Given that SEM model did not fit with the data, a multivariate regression analysis was conducted to regress the classroom assessment fairness factors on the personal belief in a just world. The results showed that students’ personal belief in a just world significantly predicted fairness in grading ($\beta = .262, p < .05$), and fairness in feedback ($\beta = .294, p < .05$) and inversely predicted students’ perceived unfairness in feedback ($\beta = -.19, p < .05$).

4.2.5 Subgroup Analysis

Multiple One-way ANOVAs were conducted to examine whether students’ perceived fairness as represented by the five factors of the Classroom Assessment Fairness Inventory was significantly different across the demographic variables (i.e., school type, gender, age, average grade, ethnicity, student immigration background, and parent immigration background). This analysis was also conducted for Personal Belief in a Just World survey. Across the classroom assessment fairness factors, only age was identified as significant for unfairness in groupwork ($F(6, 210) = 2.85, p < .05$). Post hoc Bonferroni test was conducted to further examine the differences across age groups. Given that there were seven age groups for comparison, Bonferroni test was adjusted to be $p < .007$ to allow for interpreting potential significant differences. While the age group of above 21 showed significant differences at the $p < .05$ from the age groups of 18 ($p = .017$) and 19 ($p = .033$), none of these results were significant according to the adjusted Bonferroni significance level. The disproportional representation of students across the age groups as shown in the demographics table (Table 12) suggested that further analysis was not warranted. None of the demographic variables were also significant in relation to
personal belief in a just world. The disparity in the number of students representing the different levels within each demographic variable also warranted no further analyses.

Figure 5. Standardized SEM results for belief in a just world and classroom assessment fairness
Chapter 5

Discussion

تو مگو همه به جنگند و ز صلح من چه اید / تو یکی نه ای، هزاری، تو چراغ خود برافروز
که یکی جراغ روشن ز هزار مرده بهتر (مولانا)

Say not all are fighting, what use is my lone call for peace / You’re not one, but thousands; light
your beacon /For one burning beacon is better than a thousands dead (Rumi)

Students’ perception of assessment is a growing area of research in classroom assessment
(Brown, 2021; Leighton, 2019; McMillan, 2016, 2019; Struyven, Dochy, & Janssens, 2005);
however, far less attention has been paid to this area of research vis-à-vis teachers’ perceptions of
assessment (Barnes, Fives, Dacey, 2014), particularly within K-12 educational contexts. While
fairness has been recognized as the key component of students’ perceptions of assessment
(Dorman & Knightly, 2006; McMillan 2016), it has been given limited and inconsistent attention
in the literature. On the flip side, examining students’ perception of assessment and fairness is
critical for multiple reasons. Students’ perceptions of assessment including fairness influences the
way they think, feel, and engage with assessment processes as well as their subsequent learning
and psychosocial outcomes. Students’ perceptions not only influence their immediate learning
opportunities and outcomes, but also establish entrenched prolonged attitudes toward assessment,
as shown by research on students’ conceptions of assessment (Brown, 2021) as well as student
teachers’ beliefs upon their induction into teacher education programs (Daniels, Poth, Papile, &
Hutchison, 2014). Students’ perceptions of assessment including fairness also offers a window to
examining their teachers’ assessment practices, providing an additional lens to understand and
mobilize effective assessment practices in the classroom. Despite its multiple implications,
students’ perceptions of fairness of assessments and assessment processes are yet to receive
systematic attention and conceptualization in the literature.
Four major research directions can be identified in the students’ perceptions of assessment literature, with three of them reviewed in a comprehensive synthesis by McMillan (2016). These four research directions include (a) Brown and colleagues’ research on students’ conceptions of assessment (e.g., Brown, 2011, 2021), (b) Alkharusi and colleagues’ (e.g., Alkharusi, 2011) as well as Dorman and colleagues’ (e.g., Dorman & Knightly, 2006) research on students’ perception of classroom assessment environment; (c) Pekrun and colleagues’ research on students’ emotions of assessment (e.g., Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011), and (d) Flores and colleagues’ students’ perceptions of assessment in higher education (e.g., Flores, Veiga Simão, Barros, & Pereira, 2015). Across the four research directions, fairness has not received adequate attention embedded within the theoretical frameworks of these research bases and has only explicitly been measured with a few items. Table 22 presents these research directions, associated instruments, and fairness-relevant dimensions and items.

Table 23. Students’ perceptions of assessment research and fairness conceptualizations and measurement

<table>
<thead>
<tr>
<th>Research Directions</th>
<th>Instrument</th>
<th>Dimension</th>
<th>Fairness Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown and colleagues</td>
<td>Student Conceptions of Assessment</td>
<td>Assessment is Irrelevant</td>
<td>Assessment is unfair to students</td>
</tr>
<tr>
<td>Alkharusi and colleagues</td>
<td>Perceived Classroom Assessment Environment</td>
<td>Performance-oriented Classroom Environment</td>
<td>Two items (e.g., In this class, the assessment results do not fairly reflect the effort put in studying the subject)</td>
</tr>
<tr>
<td>Dorman and colleagues¹</td>
<td>Perceptions of Assessment Task Inventory</td>
<td>Congruence with Planned Learning</td>
<td>Three items (e.g., My assessment in science is a fair indicator of my work.)</td>
</tr>
</tbody>
</table>

¹ Note that most items under student consultation, transparency, and diversity in Dorman and colleagues can be interpreted as relevant to fairness. However, only items with explicit reference to fairness were retained to avoid this researcher-imposed interpretations.
Given the lack of explicit conceptualization of fairness in students’ perceptions of assessment literature, the results of this dissertation study contribute to conceptualizing fairness in assessment based on empirical data from first-year undergraduate students reflecting on their secondary school experiences. The results also contribute to provide a framework for fairness in response to policy demands in Canada for fair and equitable assessments (Ontario Ministry of Education, 2010; *Student Assessment Policy*, Alberta Ministry of Education). For example, while the Growing Success: Assessment, Evaluation, and Reporting in Ontario Schools (Ontario Ministry of Education, 2010) puts fairness as a key fundamental quality in student assessments, it does not provide guidance on what fairness means and how it can be practiced. To address this gap, the empirical results in this study provide a multidimensional conceptualization of fairness in assessment as viewed by students as a basis to inform teacher practice. Further, this study contributes methodologically to the literature by developing the Classroom Assessment Fairness Inventory, which can subsequently be used to collect additional empirical evidence as a basis to enhance fair practices in school-based assessments.

The results from qualitative interviews and the Classroom Assessment Fairness Inventory showed that students in this study perceived fairness in assessment in relation to seven themes: (a) overall perception of fairness, (b) fairness in groupwork, (c) fairness in exams, (d) fairness in cheating, (e) fairness in grading, (f) fairness in feedback, and (g) socioemotional environment. Students also reported psychological and behavioral responses in response to un/fair conditions. Figure 1 illustrates a framework for students’ perception of fairness as shaped by these themes.

---

<table>
<thead>
<tr>
<th>Pekrun and colleagues</th>
<th>Test Emotions Questionnaire and Achievement Emotions Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flores and colleagues</td>
<td>Students’ Perceptions of Fairness Perceptions of Fairness of Learner-centered Methods Two items (e.g., ‘learner-centered assessment methods such as portfolios are fairer than traditional ones’)</td>
</tr>
</tbody>
</table>
Figure 6. A framework for students' perceptions of fairness in classroom assessment

These themes have been identified in prior assessment literature as key contributors to students’ perceptions of fairness; however, prior literature has presented these themes in unrelated and disparate ways. Given that each of these themes centered on a specific domain in assessment (e.g., grading, cheating, feedback), researchers with expertise in each domain have examined fairness as relevant to their specific research area. While this shows the saliency of fairness in each domain, the appreciation of fairness theory across all assessment domains is largely hindered, thus precluding the conceptualization and systematic importance of fairness theory in students’ assessment and school experiences.

The disparate presentation of fairness across various assessment domains might have stemmed from the lack of a unified fairness theory such as social psychology theory that can further connect and shape an explicit research area for fairness in assessment. Therefore, the
discussion of the results in this study is followed in relation to the identified themes (Figure 1) and their connections with social psychology theory and findings of prior research in each domain. In this way, the following discussion aims to bring together this disparate literature and discuss the findings of this research in light of a unifying social psychology theory. This discussion further stimulates theory-driven and domain-interconnected research on fairness in assessment. This research not only informs the discourses within the individual research domains but also provides interconnected perspectives into fairness investigations that can more systematically benefit the fairness theory in assessment. As such, the findings from this study resonates with the premise of Standards for Educational and Psychological Testing (2014) that posit that fairness is a fundamental quality in assessment as fairness is a quality that underpins many domains of assessment practice.

Accordingly, the following discussion responds to the overall purpose of this dissertation research, which was to extend the program of research on investigating students’ perceptions of fairness through the theory of social psychology of fairness. By integrating empirical evidence collected in relation to each of the guiding research questions with the social psychology theory, I offer a unifying theory for fairness in classroom assessment. The guiding research questions for this study were:

1. How do first-year undergraduate students perceive fairness in classroom assessment in relation to their secondary school experiences in Ontario, Canada?
2. What psychological and behavioral responses are provoked by the experiences of fairness in classroom assessment?

5.1 A Unifying Theory of Fairness in Assessment

The framework (Figure 6) represents various domains of assessment in which students perceived fairness in classroom assessment. Across these domains, social psychology is the unifying theory that underpins students’ perceptions of fairness in classroom assessment. Students
used the multiple distributive, procedural, and interactional principles to interpret the overall fairness of assessments in groupwork, exams, cheating, grading, feedback, and socio-emotional environment. Accordingly, students perceived fairness in classroom assessment by evaluating the outcome distributions (e.g., equitable grades, feedback, punishments), procedures for these outcome distributions (e.g., transparency, consistency, and voice in exams, cheating, and groupwork), as well as the interpersonal and informational aspects of these distributions (respect, adequacy, and honesty in the socio-emotional environment of assessment).

While the noted distributive, procedural, and interactional justice principles were prevalent in student perception of assessment fairness, this study also identified larger social and psychological mechanisms as the underpinning factors in students’ perceptions of fairness in assessment that contribute to the theory of social psychology of fairness. With respect to social mechanisms, the examples of the qualitative phase showed that the social inclusion structure of a school influenced a students’ acceptance to burden the groupwork activities at the expense of group-members’ free-riding to gain social recognitions and friendship at school. Additional examples included considering (a) various contextual issues such as considering accountability-versus achievement-based grading in punishing a late work submission, (b) students’ historical and personal background in decisions around cheating and grading, (c) race and gender in assessment processes, and (d) peer bullying in student relationships. The psychological mechanisms included examples such as the influence of teachers’ expectations and students’ un/awareness of such expectations in the assessment processes, and the rationalizations of students for the fairness of their personal life, impacting their life- and assessment-related perceptions and reactions. Overall, these social and psychological mechanisms provide additional layers coupled with distributive, procedural, and interactional justice principles in social psychology of fairness to interpret the underpinning factors and processes that contribute to students’ perception of fairness. Recent research in assessment (Murillo & Hidalgo, 2020) has
showed the influence of socio-economic mechanisms on teachers’ conceptions of fairness in assessment and future research needs to further investigate these social and psychological mechanisms underlying students’ and teachers’ perceptions of fairness in assessment. This conclusion resonates with the conceptual investigations of fairness in this dissertation study that argued for inclusion of sociological and psychological in constructing a fairer culture of assessment. Accordingly, the conceptual and empirical investigations in this dissertation study present and expand the theoretical foundations for social psychology theory to provoke additional theory-driven empirical work in various domains of assessment that can potentially inform fairer practices in classrooms. The following discussion examines the themes of this study, analyzed through social psychology theory, in relation to the disparate literature of fairness in each assessment domain to further support the framework of this study and provide further granular theory-driven research directions in the assessment literature.

5.1.1 Overall Perceptions of Fairness

Participants used three social psychology justice principles of equality, equity, and respect as their overall perceptions of fairness in assessment. Paralleled with the empirical findings in classroom assessment fairness (Rasooli et al., 2019; Tierney, 2014; Murillo & Hidalgo, 2017, 2020), participants in the qualitative phase reported equality in opportunity and treatment as well as equity as caring for students’ personal histories to represent the overall foundations for their perceived fairness in classroom assessment. Participants’ emphases on a balanced look of students’ personal histories and their curriculum achievement in the equity-based interpretation of fairness reflected the long-standing tenet in the socio-cultural theory of assessment that calls for the need to ‘look in to their [students’] histories and not into their heads’ (Elwood as cited in Elwood & Murphy, 2015, p. 187). This interpretation of equity has been highlighted in previous work on fairness in assessment (Gipps & Stobart, 2009); however, little empirical attention to this aspect of socio-cultural theory and its promotion in classroom
assessment practices has been observed. Future empirical research needs to further examine the prevalence of the equity-based interpretations (i.e., balancing the look at student head and/or history) in students’ and teachers’ perception and practices of classroom assessment. While less emphasized vis-à-vis equality and equity in the previous research (Rasooli et al., 2018; Tierney, 2013), students highlighted respect as another key foundation for their overall perceptions of fairness in this study. Students interpreted fairness in respect based on respectful treatment of all as well as reciprocal treatment in cases where an unfairness was observed.

5.1.2 Fairness in Groupwork

Participants used distributive and procedural justice principles to perceive fairness in group composition, dynamics, and grading. In response to the beginning interview question, ‘can you share with us any experiences of fairness or unfairness from your high school?’, most participants shared their experiences of groupwork, signaling further attention into this theme. Participants were highly concerned about the grading outcome of the groupwork, driving their preference to control over group member selection over preference for teacher-based assignment. Most participants argued for desiring to have control over selecting their group members given that they could better distribute the work division and successfully cooperate to deliver groupwork outcomes and receive deserved outcomes. In a systematic review of 83 studies on assessment in groupwork, Forsell, Forslund Frykedal, and Hammar Chiriac (2020) reported that students generally tend to consider groupwork as fair and positive when they are provided with more control and voice in the assessment process.

Arguing for teacher-based assignment as opposed to student-based selection, participants were often concerned with concentration of high-ability students in a group composition, driven by a desire to receive good outcomes (see Webb, Nemer, Chizhik, & Sugrue, 1998) and free-riding of friends. Most students also reported experiences of unfairness due to equal group grades and therefore suggested individual grading as the fairest. The unfairness in groupwork factor in
the Classroom Assessment Fairness Inventory also included three items focusing on providing equal group grades and lack of grade justifications. Respondents largely perceived these items as unfair.

Forsell et al., (2020) reported that in evaluating the quality (i.e., validity, reliability, and fairness) of assessment in groupwork, students considered individual grading as fair while they perceived receiving equal group grades as unfair. Given that four purposes were identified for groupwork assessment in Forsell et al.’s review (i.e., improving group work, promoting learning, giving grades, and simulating real-life contexts), it is significant for future research and practice to further probe whether the current groupwork practices do contribute to these four purposes beneficially. As groupwork were one of the salient themes for student responses in this study, this domain warrants further attention into the fairness of existing assessment practices in groupwork in schools. An additional issue that does not seem to be adequately addressed in the previous groupwork literature in assessment is the impact of social inclusion structure of a school in the formation of groups. A mixed-race student in this study reported a year-long experience of exclusion from student groups that had costed the student to burden the free-riding of group insiders to gain friendship-- a goal that does not seem to be reflected in the identified objectives of prior studies (Forsell et al., 2019). This experience has led this student to change the school next year. Future research on social psychology of fairness needs to extend beyond three noted distributive, procedural, and interactional justice dimensions to account for the larger socio-economic, cultural, and historical structures that shape students’ perception of fairness in various domains of assessment including groupwork.

5.1.3 Fairness in Exams

Participants used multiple procedural justice principles to evaluate fairness in exams. Aligned with the findings of previous literature (Murillo & Hidalgo, 2017; Rasooli et al., 2018, 2019; Sonnleitner & Kovacs, 2020), students considered transparent communication of exam
information, voice and reasonableness in exam scheduling, alignment between exam and teaching content, and flexible policy for missed exam policies as contributing to their perceived fairness. Given the prevalence of these procedural principles in students’ evaluations of fairness in exams across countries as well as K-12 and higher education contexts, teachers’ assessment practices need to be directed toward attending to and enacting these procedures.

While the results of this study in consonant with prior literature have consistently documented the violations of fair exam practices (Rasooli et al., 2019b) and have advised principles and practices to promote fairness in such practices (Gordon & Fay, 2010; Houston & Bettencourt, 1999; Pepper & Pathak, 2008), it appears that these principles are not effectively mobilized into teachers’ assessment practices in classrooms. While fairness is a priority for teachers, it is key to provide fairness learning opportunities in pre- and in-service education programs. Teachers would benefit from dialogues examining how fairness can be upheld in various domains of assessment, particularly in relation to exams.

5.1.4 Fairness in Cheating

Participants in this study used multiple distributive and procedural justice principles as well as a utilitarian perspective to perceive fairness in cheating. These principles and perspectives focused on students’ perceived fairness in cheating decisions and processes, while prior fairness research has largely aimed to demonstrate the impact of students’ perceived unfairness on their tendency to cheating in classrooms (Donat, Dalbert, & Kamble, 2014; Lemons & Seaton, 2011; Murdock, Beauchamp, & Hinton, 2008; Murdock, Miller, & Goetzinger, 2007).

Participants used justice principles such as transparency, voice, and consistency to perceive fairness in implementing cheating decisions. The common response across participants focused on giving zero as a fair punishment for cheating behavior. Participants in both phases considered that zero is fair because cheating (a) is unfair to students who have put effort, (b) is unethical, and (c) the cheating act deserves punishing. A few participants also drew on a
utilitarian perspective to argue for punishing the perpetrator of cheating act to deter other students’ from considering the act as legitimate. While less strongly than participants in this study, 54% of higher education students in Duplaga and Astani’s (2010) study agreed with giving zero as fair for the first occurrence of cheating on an exam.

Despite the approval for punishment of cheating, a few participants in the qualitative phase argued that teachers should care about students’ personal background, the number of cheating behaviors observed in the student, and the consequences of cheating decision on student future. In the quantitative phase, participants however marginally supported a teacher’s avoidance to forgiving students’ cheating (51%) and also largely perceived explaining cheating decisions to students as fair. Previous research also revealed that student and teachers consider student reputation (i.e., well- or mis-behaved), controllability (i.e., the action was within the control of the student), severity of the behavior (e.g., fighting vs. talking with friends), and stability and consistency of the misbehavior in their evaluation of the fairness of procedures and decisions regarding addressing cheating behavior (Bear & Fink, 1991; Reyna & Weiner, 2001). Reyna and Weiner (2001) showed that teachers punished students’ misbehavior (e.g., cheating) when the action was within students’ control and student was consistent in misbehaving. Teachers also demonstrated utilitarian strategies to modify students’ misbehavior and help remove them in the future. While the findings of this study coupled with previous studies demonstrated the complexity of students’ perceptions of fairness in cheating, it appears that punishing (i.e., giving zero) for majority cases is what students agreed on, with consideration of other inputs (e.g., student personal background) impacting this decision.

5.1.5 Fairness in Grading

Participants in this study articulated distributive, procedural, and interactional justice principles in perceiving fairness of grading issues. This is unsurprising as grading is a key theme for students’ perceptions of fairness (Resh & Sabbagh, 2016). Several participants in the
qualitative phase contested reducing grades for the late assignments on the basis that this reduction is irrelevant to student learning while they argued for its relevance if the teacher aimed to teach students to be accountable. This purpose-driven interpretation of fairness in grade adjustments has previously been observed in teachers and is informative to observe this acknowledgement in students’ perspectives (Alm & Colnerud, 2015; Olsen & Buchanan, 2019). Participants also highlighted the role of caring in grade decisions with attention to students’ personal history and background. Despite the significance of caring in students’ and teachers’ interpretation of fair assessments (Rasooli et al., 2019; Tierney, 2014), it still begs additional research to explore how teachers manage the tensions in accounting for students’ personal histories and their immediate performance in relation to the curriculum, especially in the context of achievement-based grading (Brookhart et al., 2016; Randall & Engelhard, 2010).

Aligned with the previous findings (Rasooli et al., 2018, 2019), participants also considered procedural justice principles of transparency in grading criteria, voice to appeal for grades, and grade justifications as key components of their perceived fairness in grading. The fairness in grading factor in the Classroom Assessment Fairness Inventory included six items, where participants largely supported treating students’ respectfully, considering achievement in grading, detailing the grading criteria, communicating the test results timely, providing voice for students to appeal for grades, and justifying the grading decisions. Importantly, participants in the qualitative phase repeatedly stated experiences of unfairness in relation to the consistency principle. Participants interpreted consistency not only in relation to teachers’ ability to apply similar grading criteria across students of the same class, but also relevant to teachers’ ability to grade students’ work in the same subject based on similar criteria across schools. Given that the Canadian pathway to university is largely dependent on the grades from high school in Ontario, students’ concerns about the consistency of implementing grading criteria were signalling the stakes of fairness in Canadian assessment culture (Scott et al., 2014). Future research needs to
empirically examine the fairness of teachers’ assessment decisions within Canadian contexts as it has a deciding influence on students’ university pathways.

5.1.6 Fairness in Feedback

Participants in this study considered multiple informational and procedural justice principles to arrive at perceptions of fairness in feedback. Participants in the qualitative phase reported experiences of un/fairness, where the teacher feedback was (not) adequate, honest, timely, and transparent. The *unfairness in feedback* factor and the *fairness in feedback* factor in the Classroom Assessment Fairness Inventory included seven items, where participants largely perceived inequitable feedback provision based on student effort, disrespectful treatment during feedback, biased feedback provision, and unreasonable justification for inequitable amount of feedback provision as unfair. However, participants largely perceived providing transparent criteria, timely communication of feedback, and voice opportunity to discuss the feedback as fair. These principles were also found in the previous studies to influence students’ perceived fairness in feedback (Flores, Veiga Simão, Barros, & Pereira, 2015; Kerssen-Griep & Witt, 2012; Lizzio & Wilson, 2008; Rasooli, DeLuca, et al., 2019).

With the recent emphases on feedback literacy and students’ agency in interpreting and acting upon feedback (Carless, 2006; Carless & Boud, 2018; Van der Kleij & Lipnevich, 2020), students’ perceived fairness is identified as a key component influencing students’ feedback-driven learning. Despite the recognized significance of fairness in effectiveness of feedback, it is yet to be conceptualized and measured adequately in assessment feedback literature. The cross-cultural findings in K-12 and higher education suggest that students consider the outlined justice principles to perceive fairness in feedback that can be potentially used to measure students’ perceived fairness in classroom assessment.

5.1.7 Socioemotional Environment
Participants in the qualitative phase highlighted distributive and interactional justice principles in socio-emotional environment as a key theme in their perceived fairness in classroom assessment. Previous research has also considered socio-emotional and constructive environment as a key component for fairness in socio-cultural theory of assessment (Moss et al., 2005; Tierney, 2014). In the participants’ perspectives, the socio-emotional environment is influenced by teacher-student relationships, student gender, race, and peer influence. Participants discussed the key role that assessment performs in creating diverse socio-emotional environment (i.e., negative or positive) for various students. Fair assessment performance produces positive cycles of teacher-student relationship while the reverse is the case for lower performances.

Based on the student performance, teachers make expectations about the student and their achievement that influence the subsequent dynamic classroom interactions. Previous research in teacher expectation has demonstrated that teachers make inferences about students’ achievement and these inferences impact students’ summative achievement positively or negatively (Good et al., 2018). Mediating this relationship is the differential explicit or subtle behavior (e.g., assessment, feedback, and interactions) that teachers project onto different students (Wang et al., 2018). The outcomes of this differential interactions subsequently impact students’ psychosocial and achievement outcomes (Wang et al., 2018). The findings of the qualitative phase showed that participants perceived teachers’ differential assessment behavior and classroom interactions as unfair. With previous research showing the negative psychosocial outcomes of perceived unfairness (Kazemi, 2016; Chory, 2007), the results of this study direct additional attention to systematically examine the connections between teachers’ expectations, differential assessment and interaction behavior, student perceived fairness, and psychosocial outcomes. What needs additional attention in both social psychology of fairness and teacher expectation literature is what role students consciously play in devising approaches that influence teachers’ expectations and whether this conscious ability is influenced by students’ advantaged versus disadvantaged
backgrounds. At least, a student participant in this study reported consciously impacting her teacher’s attitudes.

In addition to assessment outcomes, participants in this study reported unfair experiences, where their teachers had considered race and gender in their assessments. Previous research has showed that students most at risk of unfair treatment come from socially disadvantaged backgrounds that are more susceptible to low assessment expectations based on their ethnicity (Peterson, Childs, & Kennedy, 2004), gender (Watson et al., 2015), socio-economic status (Sorhagen, 2013), and disability (Wang et al., 2018). This differential treatment was also reported in peer relationships by a student participant with Individualized Education Plan, where his peers attempted to embarrass him for his responses to assessment questions. While recent research is moving toward analyzing fairness conceptions based on sociological structures such as socio-economic background (Murillo & Hidalgo, 2020), future research in fairness in classroom assessment requires further movement toward examining students’ and teachers’ perspectives based on the social, cultural, historical, and economic structures within which they live, grow, teach, learn, assess, and get assessed.

It is noteworthy that while socioemotional environment marked a key theme in the qualitative phase, it was not included in the Classroom Assessment Fairness Inventory because capturing the social environment and relationships within the confines of a measurement instrument is not possible. However, the inclusion of demographic questions and Personal Belief in a Just World survey aimed to provide additional social variables to account for the complexity of students’ perceptions of fairness in assessment. Specifically, the results of multivariate regression also showed that the personal belief in a just world significantly predicted fairness in grading (β = .262, p < .05), and fairness in feedback (β = .294, p < .05) and inversely predicted students’ perceived unfairness in feedback (β = -.19, p < .05). These results provided partial validity evidence to support the relationship of Classroom Assessment Fairness Inventory and an
external variable. Previous research has demonstrated the positive relationships between students’ perceived fairness in classrooms and personal belief in a just world (Donat et al., 2014; Peter & Dalbert, 2010). Given that personal belief in a just world has often been administered with its counterpart 10-item teacher justice scale (Dalbert & Stoeber, 2006), future research needs to further investigate the validity of Classroom Assessment Fairness Inventory and the outlined teacher justice and personal belief in a just world scales in cross-cultural samples. Additionally, future research is required to take both qualitative and quantitative designs to provide multi-perspectives into socio-emotional environment of fairness perceptions.

5.1.8 Psychological and Behavioral Responses

Finally, participants in the qualitative phase articulated multiple psychological and behavioral responses to un/fairness experiences such as learning and motivation consequences, emotional reactions, drop out, talk with parents, friends, and principals, and inaction. These findings confirm the multiple psychological and behavioral outcomes that students reported in response to un/fairness perceptions in previous empirical studies (Donat, Peter, Dalbert, & Kamble, 2016; Horan et al., 2010; Mameli, Biolcati, Passini, & Mancini, 2018; Rasooli, DeLuca, et al., 2019; Resh & Sabbagh, 2016). Participants’ strong emotional and psychological responses to un/fair experiences call for additional attention into exploring students’ perceived fairness in Canada and elsewhere. The existing findings also need to be used to drive teachers’ education into psychology of fairness in school-based assessment, an area that is yet to be given adequate attention in the field.

With respect to behavioral responses and on the positive side, participants articulated experiences, where they were able to show agency through talking with the teacher, principal, parents to contest the perceived unfair experiences. The students’ agency and power in K-12 education to show behavioral responses against unfairness is energized by the democratic foundations in Canadian society. On the negative side, participants chose not to respond in
several cases, where they perceived their reaction not to worth its stakes. Participants rationalized and cognitively reframed their inactions to show the insignificance of the perceived unfairness to protect their well-being, as was explained in the previous fairness studies (Adams, 1965).

5.2 Limitations of This Research

As with all studies, this study also has several limitations. The sampling of the first-year undergraduate students provided a comprehensive understanding of the diversity and depth of students’ perceived fairness across various grade levels in secondary schooling. The sampling also provided insights into the impact of un/fair experiences on students’ academic and social life. However, this sampling limited the potential of this study to capture the fairness experiences of students who did not choose to enter post-secondary or have dropped out of the secondary school. Future studies can aim to collect data from among students with diverse backgrounds that can provide potential diversity. Further, while previous research has showed that fairness experiences are acute in students’ memories that they dwell for years and students can recollect them after years (Rasooli et al., 2019), psychological research suggests that individuals reinterpret their memories in the course of time. Future studies could collect more direct data from students attending secondary schools.

Due to the pandemic, the quantitative phase also had a small sample size that may have impacted the statistical analyses and results. Specifically, the use of CFA after EFA on the same sample has limited evaluating the structural validity evidence for the inventory across separate samples. Future research needs to administer the inventory with larger sample sizes across cultural contexts to provide additional logical and empirical evidence to evaluate the validity of the inventory. Further, future could collect data from rural schools or schools from low socio-economic background to examine the impact of social context more adequately on students’ perceptions of fairness.
In addition to faceted justice approach to measurement, there is another approach for measuring perceived fairness that is worth including in the future versions of the Classroom Assessment Fairness Inventory. Faceted fairness measures the “global perceptions of appropriateness rather than specific justice rules[principles]” (Colquitt & Rodell, 2015, p. 194) by using the term ‘fairness’ directly in the stem. For example, Tata (2005) measured procedural fairness using the following items (e.g. The grading procedure used for the term paper was fair, The way that the instructor decided the grade was unjust). This method could not produce useful items alone for measuring perceived fairness in classroom assessment scenarios as it leaves us wondering what justice principle(s) a student might use to constitute fairness of assessment. However, it is a nice addition to the faceted justice approach in the Classroom Assessment Fairness Inventory to examine whether students’ evaluation of each item including individual justice principles in a scenario matches with their global perception of fairness of the teacher in the scenario. Future research can include this item and examine the results empirically.

While this study provided validity evidence for the Classroom Assessment Fairness Inventory based on test content, internal structure, and relationship to other variables, future research would need to investigate empirically whether the justice principles align with the action items as conceptualized in scenarios. As this inventory is a new instrument in classroom assessment, future research needs to tweak the scenarios and associated items that did not load significantly in the current study’s five-factor measurement model in an aim to provide a more adequate conceptual understanding of classroom assessment fairness shaped by the initial five scenarios. Specifically, the evidence from confirmatory factor analysis showed that the items in the exam scenario did not contribute to the (five-factor) measurement model for the inventory. However, items 6 and 7 from the exam scenario contributed to an additional factor in the six-factor model, which had an adequate fit according to goodness-of-fit indices (Hu & Bentler, 1999; Kline, 2016). Considering that only these two items loaded onto the sixth factor, they were
deemed conceptually inadequate to represent the exam scenario and were therefore disregarded. While the measurement model showed five factors with relevant items mirroring initial conceptualizations of four scenarios, there were still additional items that were removed from each of the four scenarios to reach to parsimonious well-fit model. The meanings of items in scenario-based instruments depend on the content coverage of scenarios and when items are removed to arrive at a model fit, the conceptual coverage of the scenarios would suffer as the results of this dissertation showed. It seems that there exists a tension between empirical model fit and logical content coverage as building blocks of validity evaluations in tests and measures (Newton & Shaw, 2014). While these tensions have existed and have shaped the validity discourse within the testing context, it seems that it further needs to be closely examined for the scenario-based instruments such as the Classroom Assessment Fairness Inventory (Maul, 2017). Further, think-aloud protocols and neurological evidence with student participants would provide empirical support to understand whether students consider justice principles that were initially conceptualized by the researcher in their responses to the scenarios’ action items.

Finally, despite the appeal of scenario-based measurement of fairness in classroom assessment, it inevitably suffers from the assumption that students in their daily school lives make conscious effortful thinking when they comp up with fairness perceptions. Justice theory and research, for example fairness heuristic theory (Lind, 2001; Van den Bos, 2001) has showed that individuals create a heuristic based on an instant impression and evaluation of available information and then use this heuristic to form perception of fairness. Individuals continue to use this heuristic until they gain more information that is used as a basis to reconsider this heuristic. Scenario-based measurement requires individuals to make conscious efforts, which individuals do if they have adequate information and time available to make a fairness judgment (Folger & Cropanzano, 2001), but is unable to capture the other side of fairness decision-making that is unconscious and heuristic-based.
5.3 Implications

The results from this study offered theoretical, practical, policy-wise implications to promote fairer assessments in classrooms. Theory-wise, by reviewing the key conceptions of fairness in assessment and beyond, it provided a critical synthesis of theories that can be potentially used to examine fairness in classroom assessment. As previous research has suggested that classroom assessment is largely an atheoretical field (Brookhart, 2004) and the existing studies into fairness in classroom assessment has poorly defined fairness and rarely analyzed their results using a theoretical perspective (Rasooli et al., 2019), the critical synthesis of theories reviewed in this dissertation study contributes conceptual bases that can guide interpreting empirical data and conducting theory-driven empirical research on fairness in classroom assessment. Additionally, the results from the empirical Phases I and II also provided a framework that brought together various domains of fairness in classroom assessment using the unifying theory of social psychology that can be further used to develop the theory of fairness in classroom assessment in a more coherent and systematic way.

Practice-wise, the results from this study provide areas that students potentially use to perceive fairness in classroom assessment. Given that students’ perception of fairness in assessment might impact their engagement, achievement, and psychosocial outcomes, teachers can learn from the results of this study to provide fairer experiences in student assessments. International research evidence has consistently identified disparities in students’ assessment interactions and outcomes due to issues such as race, gender, language background, socio-economic background, and disability (Rasooli et al., 2018; Wang et al., 2018). The research evidence also suggests that students have persistently perceived unfairness in classroom assessment, which subsequently affects their learning and socio-emotional outcome (Horan et al., 2010; Kazemi, 2016). These findings support that students’ experiences of unfairness is prevalent
internationally and is also impacted by students’ backgrounds. The results from this study echo the need for using the current evidence base to support fairer practices in classrooms.

Policy-wise, the results from this study support the recent 21st century policy agendas to promote equity and fairness in assessment for all students including students from underrepresented and disadvantaged backgrounds. The results directly contribute to the Growing Success: Assessment, Evaluation, and Reporting in Ontario Schools (Ontario Ministry of Education, 2010) by providing a fairness framework that encompasses domains of assessment that students perceive fairness according to principles of social psychology theory. While this policy centralizes fairness as a core fundamental principle in classroom assessments of students in Ontario, it needs a more explicit framework for fairness that this study has been able to contribute to.

5.4 Conclusion

With recent social and educational movements toward equity, research on fairness in educational assessment is energized with researchers aiming to use various approaches to conceptualize and measure fairness in educational and classroom assessment (Cole & Zieky, 2001; Herman & Cook, 2019; Poe & Elliot, 2019; Randall, 2021; Tierney, 2013; Zwick, 2019). Specifically, researchers in classroom assessment have endeavoured to conceptualize fairness by drawing on teachers’ and students’ perceptions as situated within the historical, cultural, social, and economic contexts of classrooms (Murillo & Hidalgo, 2017, 2020; Tierney, 2013). To contribute to this growing literature, this dissertation provided conceptual and empirical advancements into investigating fairness in classroom assessment. Conceptually, it critically reviewed four key conceptions of fairness in classroom assessment: psychometric perspective, legal perspective, philosophical perspectives, and social psychological perspective. It then leveraged justice theories in philosophy, sociology, psychology, economics, and education to enrich the existing conceptions of fairness in classroom assessment. Accordingly, this conceptual
review provided additional avenues and conceptual foundations for empirical theory-driven research in the field.

Empirically, this dissertation leveraged two-phase mixed methods design to explore Ontario-based Canadian first-year undergraduate students’ perceptions of fairness in classroom assessment during their secondary school experiences in order to contribute to developing social psychology theory of fairness in classroom assessment. The empirical results from Phases I (i.e., qualitative phase) identified additional seven themes, whereby students perceived fairness in school-based assessments and showed psychological and behavioral responses to un/fairness. These qualitative results support that students’ experiences of un/fairness are shaped in multiple assessment domains and have positive and negative impacts on students’ outcomes. The qualitative results also provided the contextually relevant themes that were used to build the Classroom Assessment Fairness Inventory. This inventory contributes to the methodological gaps in the classroom assessment fairness literature and provides initial validity evidence to support evidence-based fair practices in classroom assessments.

Combined, the conceptual and empirical evidence from this study contributed to responding to policy-driven imperatives (Ministry of Education, 2010) in Ontario to provide fairer assessment practices built on students’ voices. Considering the prevalence of students’ experiences of unfairness in assessment coupled with the findings that suggest that teachers are underprepared in assessment and need further professional development in assessment (DeLuca & Klinger, 2010), the results from this study provided initial empirical foundation for promoting fairness in teachers’ assessment practices in Ontario and elsewhere. Overall, the results from this dissertation bridges the geographical gaps in assessment fairness research in Canada and within K-12 contexts internationally and provides evidence that can support policymakers, researchers, teachers, and students to move toward creating fairer conditions for classroom assessments.
References


Alberta Ministry of Education. (n.d.). *Student Assessment Policy*.


International Journal of Behavioral Development, 30, 200-207.
doi:10.1177/0165025406063638


doi:10.1177/0265532209349466

doi:10.1080/13603116.2016.1159256


doi:10.1016/j.stueduc.2017.01.001


doi:10.1080/0969594X.2019.1586643


doi:10.3102/00346543076004507


doi:10.3102/002831219841349


doi:10.1080/17449640902861562

doi:10.1080/0969594960030308

doi:10.1080/0969594960030310


doi:10.5465/annals.2014.0051


Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly, 4*, 223-233.

doi:10.1080/15434300701375832


doi:10.1111/emip.12299
Appendix A  
Ethical Clearance Letter

June 26, 2020

Mr. Aasim Rasooli
Ph.D. Candidate
Faculty of Education
Queen’s University
Dunton McArthur Hall
211 Union Street West
Kingston, ON, K7M 3R7

GKEB Ref #: GEDUC-1807-01, TRAQ #: 6034987
Title: “GEDUC-1807-01 Fairness and Justice in Classroom Assessment”

Dear Mr. Rasooli,

The General Research Ethics Board (GKEB), by means of a delegated board reviewer, has cleared your proposal entitled “GEDUC-1807-01 Fairness and Justice in Classroom Assessment” for ethical compliance with the Tri-Council Guidelines (TCP 2) and Queen’s ethics policies. In accordance with the Tri-Council Guidelines (Article 5.14) and Standard Operating Procedures (105), your project has been cleared for one year. You are reminded of your obligation to submit an annual renewal form prior to the annual renewal due date (access this form at http://www.queensu.ca/gre/annual.html; click on “Events” under “Create New Event” click on “General Research Ethics Board Annual Research Closure Form for Closed Studies”). Please note that when your research project is completed, you need to submit an Annual Renewal/Closure Form in ResearchNet indicating that the project is ‘completed’ so that the file can be closed. This should be submitted at the time of completion; there is no need to wait until the annual renewal due date.

You are reminded of your obligation to advise the GKEB of any adverse event(s) that occur during this one-year period (access this form at http://www.queensu.ca/gre/annual.html; click on “Events” under “Create New Event” click on “General Research Ethics Board Adverse Event Form”). An adverse event includes, but is not limited to, a complaint, a change or unanticipated event that alters the level of risk for the research participants or situation that requires a substantial change in approach to a participant(s). You are also advised that all adverse events must be reported to the GKEB within 48 hours.

You are also reminded that all changes that might affect human participants must be cleared by the GKEB. For example, you may make changes to the level of risk, applicant characteristics, and implementation of new procedures. To submit an amendment form, access the application by at http://www.queensu.ca/gre/annual.html; click on “Events” under “Create New Event” click on “General Research Ethics Board Request for the Amendment of Approved Studies.” Once submitted, these changes will automatically be sent to the Ethics Coordinator, GKEB, at University Research Services for further review and clearance by GKEB or the Chair, GKEB.

Note: Due to COVID-19, human participant research policies, in relation to hospital and non-hospital based research, are being continually updated. Many restrictions are now in place with respect to in-person research. For the most current information on the COVID-19 impact on research, please visit https://www.queensu.ca/vpr/covid-19. For information directly related to GKEB please visit the Research Ethics FAQ.

On behalf of the General Research Ethics Board, I wish you continued success in your research.

Sincerely,

[Signature]

Chair, General Research Ethics Board (GKEB)
Professor Dean A. Trupp, PhD
Depaetments of Psychology, Anesthetics and Urology Queen’s University

c:  Dr. Christopher Delucia, Supervesor
Dr. Fausta Beche, Chair, Unit REB
Kyle Commins-Benedry, Dept. Admin

182
Appendix B
Ethics Amendment Approval Letter

July 16, 2020

Mr. Amir Rasooli
Ph.D. Candidate
Faculty of Education
Queen’s University
Dunsmuir McArthur Hall
311 Union Street West
Kingston, ON, K7M 3N7

Dear Mr. Rasooli:

RE: Amendment for your study entitled CEDUC-100*-20 Fairness and Justice in Classroom Assessment; TRAQ # 6936837

Thank you for submitting your amendment requesting the following changes:

1) To include a $10 gift card to Starbucks or Amazon to each of the 20 participants for Phase 1 interviews, as an appreciation for the time and contribution;

2) Phase 1 amended Recruitment Scripts (v. 2020/07/02);

3) Phase 1 amended Letter of Information/Consent Form (v. 2020/07/02).

Note: Due to COVID-19, human participant research policies, in relation to hospital and non-hospital based research, are being continually updated. Many restrictions are now in place with respect to in-person research. For the most current information on the COVID-19 impact on research, please visit https://www.queensu.ca/vpr/covid-19. For information directly related to GREB please visit the Research Ethics FAQs.

By this letter, you have ethics approval for these changes.

Good luck with your research.

Sincerely,

Chair, General Research Ethics Board (GREB)
Professor Dean A. Tripp, PhD
Departments of Psychology, Anesthesiology & Urology Queen’s University

cc: Dr. Christopher DeLuca, Supervisor
Appendix C
Ethics Amendment Approval Letter

November 2, 2020

Mr. Amir Rasooli
Ph.D. Candidate
Faculty of Education
Queen’s University
Duncan McArthur Hall
511 Union Street West
Kingston, ON, K7L 3N7

Dear Mr. Rasooli:

RE: Amendment for your study entitled: “GEDUC-1007-20 Fairness and Justice in Classroom Assessment; TFAQ 603603”

Thank you for submitting your amendment requesting the following changes:

1) To attach the instrument for use in the second phase of the research, the "classroom assessment fairness inventory."

2) To increase the interviews with first-year undergraduate students from 20 participants to 30 participants, to gain more diverse responses.

3) Classroom Assessment Fairness Inventory v.1 (v. 2020/10/26).

By this letter, you have ethics approval for these changes.

Good luck with your research.

Sincerely,

[Signature]

Chair, General Research Ethics Board (GREB)
Professor Dean A. Trapp, PhD
Department of Psychology, Anesthesiology & Urology Queen’s University

Dr. Christopher DaLuca, Supervisor
Appendix D
Phase I Interview Consent Form

Study Title: Fairness and Justice in Classroom Assessment: Conceptual and Empirical Investigations

Name of Principal Investigator: Amir Rasooli, Faculty of Education, Queen’s University
Name of Supervisor: Dr. Christopher Deluca, Faculty of Education, Queen’s University

What is this study about? The purpose of this research is to explore students' perceptions of fairness and justice in classroom assessment. This research involves interviewing first-year undergraduate students in Ontario about their experiences of classroom assessment fairness.

What is involved to participate in this study? The study will require one 15-30 minute interview. The interview will take place via Microsoft Teams at a time convenient to you. With your consent, your interview will be audio recorded. There are no known physical, economic, or social risks associated with this study. You will receive a $10 gift card to Starbucks or Amazon, as an appreciation for the time and contribution. There is a slight risk that memories of unfair school experiences may provoke negative emotions; however, your fairness experiences may also provoke positive experiences. If you feel upset during or after the interview, please call the Telephone Aid Line Kingston (TALK) at 613-544-1771. The benefits to this study include amplifying your voice in our efforts to make classrooms and assessment fairer for students. You will not be paid for taking part in this study. This research will contribute to the body of research surrounding fairness in assessment and will promote ways to enhance students' fair experiences in classrooms.

Is participation voluntary? Yes. You do not have to answer any questions you do not want to. You can stop your participation at any time by telling the researcher without impact on your academic standing. You may choose to withdraw from the study at any time. You may request to have your data withdrawn from the study up until December 31st 2020 by contacting me at Amir.Rasooli@queensu.ca. If you withdraw, your data will be destroyed.

What will happen to your responses? Your confidentiality will be protected to the extent permitted by applicable laws. I will do this by replacing your name with a pseudonym in all publications and a study ID number in all study records. The study data will be stored on an encrypted hard drive on the Principal Investigator’s password protected computer. The code file that links real names with pseudonyms and study ID numbers will be stored securely and separately from the data on an encrypted USB key. Access to study data is limited to those researchers on the study team, as well as the Queen’s General Research Ethics Board (GREB) may request access to study data to ensure that the researcher(s) have or are meeting their ethical obligations in conducting this research. In accordance with the General Research Ethics Board’s policy, data will be securely/password protected for a minimum of five years. The code file identifying your pseudonym and study ID number will be permanently erased from the encrypted USB key five years after study closure.

The data for this study are collected for my dissertation and I also plan to publish the results of this study in academic journals and present them at conferences. I will not include any personally identifying information from the interviews when presenting my findings. I will never include any real names with quotes. I will do my best to make sure quotes do not identify participants. During the interview, please let me know if you say anything you do not want me to quote. I will also share with you a link to your interview transcript saved in Queen’s OneDrive so you have an opportunity to make revisions and approve your interview transcript prior to conducting further data analysis. After the interview, your contact information (email address) will also be removed and destroyed.

What if you have concerns? If you have any ethics concerns please contact. Any questions about study participation may be directed to the General Research Ethics Board at 1-844-535-2988 (Toll free in North America) or chair@csu.ca. If non-English speaking participants wish to contact the Chair for ethics concerns, translation assistance may be necessary, as the GREB Chair communicates in English only.
If you have any questions about the research, please contact me at amir.rasouli@queensu.ca or contact my supervisor at cdeluc@queensu.ca or 613-533-6000 ext. 77675.

This Letter of Information provides you with the details to help you make an informed choice. All your questions should be answered to your satisfaction before you decide whether or not to participate in this research study. Keep one copy of the Letter of Information for your records.

You have not waived any legal rights by consenting to participate in this study.

If you agree to participate, you are verifying to provide an audio-recorded verbal consent upon the beginning of our interview. By consenting verbally, you are verifying that you have read the Letter of Information and all of your questions have been answered. You also consent to the use of your data gathered from interviews in the communication of this study’s findings in a dissertation, articles, or conferences, with the use of a pseudonym to protect your identity.

☐ Yes, you have my permission to use quotes
☐ No, you do not have my permission to use quotes

If you consent to participate, you will be asked at the end of the interview if you are also interested in participating in a pilot survey study on your conceptions and experiences of fairness in high school. If you like to participate, you can send an email to me within 2-3 weeks after our interview. I will then follow-up and provide you with the consent form for the survey study. Please note that there is no obligation for participation in that survey study.
Appendix E
Phase II Expert Panel Consent Form

Study Title: Fairness and Justice in Classroom Assessment: Conceptual and Empirical Investigations

Name of Principal Investigator: Amir Rasooli, Faculty of Education, Queen’s University
Name of Supervisor: Dr. Christopher Deluca, Faculty of Education, Queen’s University

What is this study about? The purpose of this research is to use your feedback to improve the quality of a recently-developed survey to measure students’ perceptions of fairness and justice with enhanced validity and reliability. You will be asked to review the survey for its validity and content alignment. Procedures for this process will be provided upon agreement to participate in this study.

What is involved to participate in this study? You will be asked to complete the survey review based on your convenient timeline; however, the review may take 45-60 minutes of your time. Please send me your comments and feedback via an email. There are no known physical, economic, or social risks associated with this study. The benefits to this study include building a more rigorous survey for measuring fairness and justice in classroom assessment based on your contributions.

Is participation voluntary? Yes. You do not have to answer any questions you do not want to. You can stop your participation at any time by telling the researcher. You may choose to withdraw from the study at any time. You may choose to withdraw from the study at any time. You may request to have your data withdrawn from the study up until December 31st 2020 by contacting me at amir.rasooli@queensu.ca. If you withdraw, your data will be destroyed.

What will happen to your responses? Your confidentiality will be protected to the extent permitted by applicable laws. I will do this by replacing your name with a pseudonym in all publications and a study ID number in all study records. The study data will be stored on the Principal Investigator’s password protected computer. The code file that links real names with pseudonyms and study ID numbers will be stored securely and separately from the data on an encrypted USB key. Access to study data is limited to those researchers on the study team, as well as the Queen’s General Research Ethics Board (GREB) may request access to study data to ensure that the researcher(s) have or are meeting their ethical obligations in conducting this research. In accordance with the General Research Ethics Board’s policy, data will be securely/password protected for a minimum of five years. The code file identifying your pseudonym and study ID number will be permanently erased from the encrypted USB key five years after study closure.

The data for this study are collected for my dissertation and I plan to publish the results of this study in academic journals and present them at conferences. I will not include any personally identifying information from the data when presenting my findings. I will never include any real names with quotes. I will do my best to make sure quotes do not identify participants. During the interview, please let me know if you say anything you do not want me to quote.

What if you have concerns? If you have any ethics concerns please contact any questions about study participation may be directed to the General Research Ethics Board at 1-844-535-2988 (Toll free in North America) or chair.GREB@queensu.ca. If non-English speaking participants wish to contact the chair for ethics concerns, translation assistance may be necessary, as the REB Chair communicates in English only.

If you have any questions about the research, please contact me at amir.rasooli@queensu.ca or contact my supervisor at cdeluca@queensu.ca or 613-533-6000 ext. 77675.
Appendix F

Phase II Pilot Testing Consent Form

Study Title: Fairness and Justice in Classroom Assessment: Conceptual and Empirical Investigations

Name of Principal Investigator: Amir Rasooli, Faculty of Education, Queen’s University
Name of Supervisor: Dr. Christopher Deluca, Faculty of Education, Queen’s University

What is this study about? The purpose of this research is to use your feedback to improve the quality of a survey to measure students’ perceptions of fairness and justice more adequately. This research asks you to take the survey while narrating aloud your thinking about the survey questions and response options.

What is involved to participate in this study? The study will be completed in a 20-minute meeting. The study will take place over Microsoft Teams at a time to your convenience. You will engage in think aloud narration while responding to the survey (i.e., explaining your thoughts and reactions while completing the survey). This process will be audio recorded with your consent. There are no known physical, economic, or social risks associated with this study. The benefits to this study include building a more rigorous survey for measuring fairness and justice in classroom assessment based on your contributions.

Is participation voluntary? Yes. You do not have to answer any questions you do not want to. You can stop your participation at any time by telling the researcher without impact on your academic standing. You may choose to withdraw from the study at any time. You may choose to withdraw from the study at any time. You may request to have your data withdrawn from the study up until December 31st 2020 by contacting me at amir.rasooli@queensu.ca. If you withdraw, your data will be destroyed.

What will happen to your responses? Your confidentiality will be protected to the extent permitted by applicable laws. I will do this by replacing your name with a pseudonym in all publications and a study ID number in all study records. The study data will be stored on the Principal Investigator’s password protected computer. The code file that links real names with pseudonyms and study ID numbers will be stored securely and separately from the data on an encrypted USB key. Access to study data is limited to those researchers on the study team, as well as the Queen’s General Research Ethics Board (GREB) may request access to study data to ensure that the researcher(s) have or are meeting their ethical obligations in conducting this research. In accordance with the General Research Ethics Board’s policy, data will be securely/password protected for a minimum of five years. The code file identifying your pseudonym and study ID number will be permanently erased from the encrypted USB key five years after study closure.

The data for this study are collected for my dissertation and I plan to publish the results of this study in academic journals and present them at conferences. I will not include any personally identifying information from the interviews when presenting my findings. I will never include any real names with quotes. I will do my best to make sure quotes do not identify participants. During the interview, please let me know if you say anything you do not want me to quote. After the interview, your contact information (email address) will also be removed and destroyed.

What if you have concerns? If you have any ethics concerns please contact. Any questions about study participation may be directed to the General Research Ethics Board at 1-844-535-2988 (Toll free in North America) or chair.GREB@queensu.ca. If non-English speaking participants wish to contact the Chair for ethics concerns, translation assistance may be necessary, as the REB Chair communicates in English only.
If you have any questions about the research, please contact me at amir.rasouli@queensu.ca or contact my supervisor at cdeluca@queensu.ca or 613-533-6000 ext. 77675.
This Letter of Information provides you with the details to help you make an informed choice. All your questions should be answered to your satisfaction before you decide whether or not to participate in this research study. Keep one copy of the Letter of Information for your records.

You have not waived any legal rights by consenting to participate in this study.

If you agree to participate, you are verifying to provide an audio-recorded verbal consent upon the beginning of our interview. By consenting verbally, you are verifying that you have read the Letter of Information and all of your questions have been answered. You also consent to the use of your data gathered from interviews in the communication of this study’s findings in a dissertation, articles, or conferences, with the use of a pseudonym to protect your identity.

☐ Yes, you have my permission to use quotes
☐ No, you do not have my permission to use quotes
Appendix G

Phase II Survey Consent Form

Study Title: Fairness in Classroom Assessment

Name of Principal Investigator: Amir Rasooi, Faculty of Education, Queen’s University
Name of Supervisor: Dr. Christopher DeLuca, Faculty of Education, Queen’s University

What is this study about? The purpose of this research is to investigate how you perceive fairness in classroom assessment based on your responses to assessment scenarios.

What is involved in participating in this study? If you read this letter and decide to participate, you will be invited to complete an online survey, which will take you approximately 20 minutes. There are no known physical, psychological, economic, or social risks associated with this study. You will not be paid for taking part in this study; however, you can choose to enter into a draw to win 1 of 50 $10 gift cards to Starbucks or Amazon, as an appreciation for the time and contribution. The benefits to this study also include amplifying your voice to be heard in our efforts to make classrooms and assessment fairer.

Is participation voluntary? Yes. Although we ask that you answer all questions as frankly as possible, you do not have to answer any questions you do not want to. You may choose to withdraw from the study at any time. You can choose to partially answer or not submit the questionnaire. You may also withdraw at any time without consequences by simply closing your browser window. However, after the questionnaire is submitted, withdrawal will not be possible as the data are not associated with names.

What will happen to your responses? Your responses will be kept confidential. Access to study data is limited to those researchers on the study team as well as the Queen’s General Research Ethics Board (GREB). GREB may request access to study data to ensure that the researcher(s) have or are meeting their ethical obligations in conducting this research. In accordance with the General Research Ethics Board’s policy, the data from this research will be retained for five years. After this time, the data will be destroyed. If data are used for secondary analysis, they will contain no identifying information. All electronic files will be password protected. Should you be interested, you are entitled to a copy of the findings by contacting the researcher through the information provided at the end of the Letter of Information. The data for this study are collected for my dissertation and I plan to publish the results of this study in academic journals and present them at conferences, but such presentations will maintain individual confidentiality. If you choose to enter the draw, your contact information (email address) will be destroyed after communicating with the draw winners.

What if you have concerns? If you have any concerns about study participation, please contact the General Research Ethics Board at 1-844-555-2988 (Toll free in North America) or chair.GREB@queensu.ca. If non-English speaking participants wish to contact the Chair for ethics concerns, translation assistance may be necessary, as the GREB Chair communicates in English only. If you have any questions about the research, please contact me at amir.rasooi@queensu.ca or contact my supervisor at cdeluca@queensu.ca or 613-555-6000 ext. 77675. This Letter of Information provides you with the details to help you make an informed choice. All your questions should be answered to your satisfaction before you decide whether or not to participate in this research study. You have not waived any legal rights by consenting to participate in this study.
Appendix H
Interview Guide

Thank you for your willingness to participate in this interview. As you know, the purpose of this research is to explore your perspectives of fairness in classroom assessment. The interview will take approximately 15-30 minutes. For the purpose of this interview, do you agree to have your voice recorded? After the interview, the audio file for this interview will be assigned a code to maintain the confidentiality of this interview. My supervisor and I will only have access to the interview. If you agree to participate, please read and sign the consent letter.

Before we get started, do you have any questions?
I am now going to start audio recording.

As you read, we are doing research on students’ experiences of fairness and unfairness in their classroom assessments because fairness is a very important issue for students.

1. Have you recently finished secondary school?
2. Did you study in public school?
3. In which province did you complete secondary school?
4. During the past five years, do you remember a time when you experienced fairness or un/fairness in the classroom assessment? If so, please describe one experience.
5. How did this fair or unfair experience impact your attitude and feelings towards your class work, your peers, or your teacher?
6. How can a teacher ensure greater fairness in classroom assessment (for example, in grading, feedback, peer assessment, group work, testing, classroom attendance, cheating)?
7. What other reactions or responses did you have to this experience?
8. Why do you think you reacted in this way?
9. In your opinion, how can experiences like this be avoided (if unfair) or promoted (if fair)?
10. How old were you when this experience occurred?
11. Why do you think it is important for assessments to be fair?
12. How can a teacher ensure greater fairness in teaching and learning generally?
13. How can a teacher ensure greater fairness in classroom interactions?
14. Do you remember a time when a teacher was un/fair to one of your classmates? If yes, can you describe what has happened?
15. Do you remember a time when a peer was un/fair to you or one of your classmates? If yes, can you describe what has happened?
16. What was your response to these situations? And why?
17. How could these experiences be avoided (if unfair)?

Demographic Information:

- How old are you now?
- What stream were you in secondary school?
- How do you identify your gender?
- How do you identify your race/ethnicity?
- Do you identify as the first or second immigration generation?
- Were you eligible for accommodations any time during your high school?
Ending the Interview

As I bring this interview to a close, I would like to know if there were anything you would like to add that you haven’t had a chance to talk about regarding fairness. I will have a survey on fairness around November with also gift card for participating. If you would like to participate in the survey, I will reach out through email with details about that study and see if you are willing to complete the survey. Thank you for participating in this interview. Your participation has been very helpful.
Appendix I
Expert Panel and Pilot Testing Instruction

Note to reviewers:

The purpose of this inventory is to explore high school students’ perceptions of fairness in classroom assessment.

As an expert in educational assessment, you have important insights and understandings that will help us to refine and enhance the inventory. On the subsequent pages, you will find the components of the instrument that students will see (highlighted in yellow) as well as instructions to reviewers (in blue).

There are no correct answers on this inventory. Instead, students’ responses will be examined to describe underpinning principles and actions driving students’ perceptions of fairness in classroom assessment.

There are two parts to the survey:

Part A: Questions about the participants
Part B: Scenario-based questions

Please follow the instructions to reviewers and record your feedback on the inventory in this document. The review should take approximately 45-60 minutes. Once you have completed your review, please save and send this document electronically to Amir Rasooli at amir.rasooli@queensu.ca (Principal Investigator) or Chris DeLuca at cdeluca@queensu.ca (Supervisor). Thank you kindly for your time and expertise! If you have questions, do not hesitate to email Amir for clarification. Reviews will be accepted until November 6th, 2020.

Part A: Questions About the participants

The purpose of Part A is to collect demographic information about the participants as well as their personal belief in a ‘just’ world (i.e., to what extent they believe their life events are generally fair).

For Part A, please use the comment feature in the Word (i.e., found in the ‘track changes’ review function) to provide your feedback on the below questions. Please add questions that you think will add to the richness of the questions in Part A.
Please check (✔) the appropriate choice below.

1. In what cities were your high school located? (e.g. Kingston, Ontario): __________________

2. What kind of high school did you attend?
   - Public
   - Private
   - Catholic
   - Other (please describe)

3. What is your gender?
   - Female
   - Male
   - I do not identify within the gender binary
   - Prefer not to disclose

4. How old are you?
   - Below 14
   - 14
   - 15
   - 16
   - 17
   - 18
   - 19
   - 20
   - Above 20

5. On average, what grades did you typically receive in high school?
   - 90%-100%
   - 80%-89%
   - 70%-79%
   - 60%-69%
   - 50%-59%
   - below 50%

6. How do you identify your ethnicity? (Select all that apply).
   - Indigenous (First Nations, Métis, or Inuk (Inuit))
   - Persian/Arab/West Asian
   - Black
   - Chinese
   - Filipino
   - Japanese
   - Korean
   - Latin American
   - South Asian
   - South East Asian
   - White (Caucasian)
   - Other: ______________
   - Prefer not to say

7. Where were you born?
In Canada
☐ Outside of Canada

8. Where were your parents born?
☐ My parents were both born in Canada.
☐ My parents were both born outside of Canada.
☐ One of my parents was born outside of Canada.

9. Considering your personal life beyond school, please indicate your agreement with the following statements (1 = strongly disagree and 6 = strongly agree).

| Example: If you strongly agree, you should mark “6”. |
|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Items          | Strongly        | Mostly          | Slightly        | Slightly        | Mostly          | Strongly        |
| To what extent do you agree or disagree with the following items? |
| a. Overall, I believe that I deserve what happens to me. |
| b. I am usually treated fairly. |
| c. I believe that I usually get what I deserve. |
| d. Overall, events in my life are fair. |
| e. In my life, unfairness does not happen very often. |
| f. I believe that most of the things that happen in my life are fair. |
| g. I think that important decisions that are made concerning me are usually fair. |

195
Part B: Classroom Assessment Fairness Inventory

The purpose of Part B is for students to rate their perceived level of fairness in various classroom assessment scenarios.

There are five scenarios, each with associated actions. Each scenario addresses a different classroom assessment domain (e.g., groupwork, exam, cheating, grading, and feedback,) significant to students’ perceptions of fairness. These domains were selected based on previous qualitative study on high school students’ perceptions of fairness in classroom assessment in Ontario as well as comprehensive scrutiny of previous literature. The associated actions within each scenario represent an underlying principle (e.g., equality, equity, transparency) based on a framework of fairness in classroom assessment (Rasooli et al., 2018, 2019, 2019). This fairness framework, presented in the next page, was selected because it provides strong conceptual and empirical bases to examine students’ perceptions of fairness. Below you will see a table that defines each fairness principle in the framework.

Students will be asked to read each scenario and identify the extent to which they consider each action to be fair (1= Highly unfair; 6= Highly fair). Students will only see the components that are highlighted in yellow (i.e., they will not see the corresponding fairness principles).

Below, you will be provided with a table including the principles and their definition. As you read each scenario and associated actions, we would like to know:

1. Do you suggest revisions to the scenario? (if so, please record suggestions in the appropriate box)
2. Do the actions align with their underlying principles? (1= not aligned to 5 = strongly aligned)
3. Do you suggest revisions to the actions? (if so, please record suggestions in the appropriate box)

Overall, do you think that scenarios represent classroom assessment fairness adequately? (if so, please record suggestions in the appropriate box).
Table 1. Descriptions of dimensions and principles of fairness and their representations in Classroom Assessment Fairness Inventory

<table>
<thead>
<tr>
<th>Dimensions of fairness</th>
<th>Principles</th>
<th>Descriptions</th>
<th>Representation of each principle across all items (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Distributive Justice</td>
<td>1.1. Equity</td>
<td>Distributing resources (e.g., grades, feedback) in a way that each student receives what they deserve.</td>
<td>4 (10%)</td>
</tr>
<tr>
<td></td>
<td>1.2. Equality</td>
<td>Distributing resources (e.g., grades, feedback) equally among all students.</td>
<td>2 (5%)</td>
</tr>
<tr>
<td></td>
<td>1.3. Need</td>
<td>Distributing resources (e.g., grades, feedback) based on caring for students’ individual needs.</td>
<td>5 (12.5%)</td>
</tr>
<tr>
<td></td>
<td>1.4. Consequence</td>
<td>Distributing resources (e.g., grades, feedback) based on considering consequences for students.</td>
<td>3 (7.5%)</td>
</tr>
<tr>
<td>2. Procedural Justice</td>
<td>2.1. Consistency</td>
<td>Consistent application of assessment procedures in classrooms across students.</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td></td>
<td>2.3. Bias Suppression</td>
<td>Neutral and bias-free application of assessment procedures in classrooms.</td>
<td>2 (5%)</td>
</tr>
<tr>
<td></td>
<td>2.4. Correctability</td>
<td>Correcting assessment procedures when there is an error in process or practice.</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td></td>
<td>2.5. Ethicality</td>
<td>Aligning assessment procedures with ethical standards and practices.</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td></td>
<td>2.6. Voice</td>
<td>Providing students with voice and an opportunity to communicate their perspective during assessment procedures.</td>
<td>5 (12.5%)</td>
</tr>
<tr>
<td></td>
<td>2.7. Transparency</td>
<td>Enacting assessment procedures with clarity.</td>
<td>3 (7.5%)</td>
</tr>
<tr>
<td></td>
<td>2.8. Reasonableness</td>
<td>Enacting assessment procedures (e.g., exam scheduling) in a way that shows good and sensible judgements.</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td>3. Interactional Justice</td>
<td>3.1. Respect</td>
<td>Treating students with respect during assessment procedures.</td>
<td>4 (10%)</td>
</tr>
<tr>
<td></td>
<td>3.2. Adequate communication</td>
<td>Providing students with adequate information about assessment procedures.</td>
<td>2 (5%)</td>
</tr>
</tbody>
</table>
Table 2: Representation of each fairness dimension within and across 5 scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Distributional</th>
<th>Procedural</th>
<th>Interactional</th>
<th>#Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Totals</td>
<td>14</td>
<td>14</td>
<td>12</td>
<td>40</td>
</tr>
</tbody>
</table>

Scenario 1. Groupwork

Mr. Chu highly values student groupwork. Based on his initial assessments, Mr. Chu formed each group with three students from different ability levels: struggling, average, and high performing students. He believed that high performing students contribute to the learning of other group members. Each group worked on their projects and prepared a final presentation. Mr. Chu has left it to groups to discuss how to distribute workload and allowed students to discuss with him if they had issues over group dynamics. As a response to students’ questions about assessment, Mr. Chu provided an overview of the project to students but not a rubric showing how he will assess students’ groupwork. Mr. Chu encouraged group members to work hard as all group members will receive the same grades as a reflection of group performance and cooperation. Several who were not satisfied with their grades appealed, but Mr. Chu did not accept their complaints.

Suggested revisions to Scenario 1:

To what extent do you consider each of these actions to be fair based on the above scenario?
<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
<th>Alignment of actions to principles 1 = not aligned to 5 = strongly aligned</th>
<th>Suggested revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mr. Chu selected group members based on mixed ability.</td>
<td>Consequence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Mr. Chu did not provide students a choice in selecting their group members.</td>
<td>Voice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Mr. Chu allowed students' complaints over group dynamics.</td>
<td>Voice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Mr. Chu was not detailed in communicating how he will assess students’ groupwork.</td>
<td>Adequate communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Mr. Chu gave the same grades to all group members.</td>
<td>Equality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Mr. Chu did not give individual grades for each group member based on their contributions and learning.</td>
<td>Equity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Mr. Chu did not justify his grades to students who appealed.</td>
<td>Justification</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Overall, do the actions represent the various fairness principles relevant to the groupwork scenario? (1= Completely Unrepresentative, 5= Completely Representative)*

**Scenario 2. Exam**

Mr. Ahmed announced that the class would have an exam the day before winter break (in 5 days). While students preferred moving the exam date because they had a pile of assignments for other subjects that were also due on the same date, Mr. Ahmed was firm on his decision as moving the exam date back would create more intensive workload later in the year. Mr. Ahmed did not explicitly state what would be on the exam. However, he did include a mix of easy and difficult questions to give all students an opportunity to show their learning. He also provided accommodations (e.g., more time) to students with disabilities and English language learners. In general, Mr. Ahmed is a lenient teacher in grading compared with other teachers in the school who teach the same subject. He also gives students a chance to appeal their grades. On the exam,
all students complained about two questions that were not covered during the course. Mr. Ahmed *harshly responded* that students should be able to answer the two questions from what had been taught. For a few students who missed the exam date, Mr. Ahmed decided to give another exam opportunity after winter break.

Suggested revisions to Scenario 2:

**To what extent do you consider each of these actions to be fair based on the above scenario?**

<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
<th>Alignment of actions to principles</th>
<th>Suggested revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mr. Ahmed held firm on the exam date.</td>
<td>Reasonableness</td>
<td>1 = not aligned to 5 = strongly aligned</td>
<td></td>
</tr>
<tr>
<td>2. Mr. Ahmed did not communicate more adequately what will be on the exam.</td>
<td>Adequate communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. The exam included both easy to difficult questions.</td>
<td>Equity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Students with disability and English language learners received</td>
<td>Need</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accommodations for the exam (e.g., more writing time).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Mr. Ahmed graded his students more leniently than other teachers.</td>
<td>Consistency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Mr. Ahmed did not remove the two questions on the content that students</td>
<td>Correctability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>claimed they were not taught before.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Mr. Ahmed did not respond to students’ complaints with respectful tone.</td>
<td>Respect</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
8. Mr. Ahmed gave another exam opportunity to students who missed the exam.

<table>
<thead>
<tr>
<th>Need (Equality, Consistency)</th>
</tr>
</thead>
</table>

*Overall, do the actions represent the various fairness principles relevant to the exam scenario? (1= Completely Unrepresentative, 5= Completely Representative)*

**Scenario 3. Cheating**

Ms. Johnston is a very strict when she catches a student cheating. However, she did not tell students her policy on cheating at the beginning of the year. One student was caught cheating on an exam and Ms. Johnston decided to give the student zero. Ms. Johnston did not give the student an opportunity to explain the reasons for cheating before making her decisions. She explained to the class that cheating is unfair to other students and asked the student to leave the classroom. The exam constituted 20% of students’ final grade. After the exam, she met with the student and explained that cheating is ethically wrong, is unfair in relation to the classmates, and she would punish anyone who cheats. The student provided a reason for their behaviour and apologized.

**Suggested revisions to Scenario 3:**

**To what extent do you consider each of these actions to be fair based on the above scenario?**

<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
<th>Alignment of actions to principles</th>
<th>Suggested revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ms. Johnston was not transparent about her cheating policy at the beginning of the course.</td>
<td>Transparency</td>
<td>1 = not aligned to 5 = strongly aligned</td>
<td></td>
</tr>
<tr>
<td>2. Ms. Johnston did not give the student an opportunity to explain the reasons for cheating before making her decision to give zero.</td>
<td>Voice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Ms. Johnston gave a zero because cheating is unfair to other students’ efforts.</td>
<td>Equity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

201
4. Ms. Johnston gave a zero because cheating is ethically wrong. Ethicality

5. Ms. Johnston gave a zero to signal that anyone who cheats should be punished for this action. Consequence

6. Ms. Johnston asked the student to leave the classroom in front of other students. Respect

7. Ms. Johnston explained her cheating decision to the student. Justification

8. Ms. Johnston did not forgive the student’s cheating this time while ensuring that the student would receive a zero next time if caught cheating. Need

**Overall, do the actions represent the various fairness principles relevant to the cheating scenario? (1= Completely Unrepresentative, 5= Completely Representative)**

---

**Scenario 4. Grading**

Ms. Mendes had students from diverse backgrounds in her classroom. She treated all her students respectfully during classroom teaching, assessment, and interactions. Ms. Mendes informed students that she would give grades based on student achievement. 70% of students’ grades were from multiple tests during the course plus 30% for students’ individual essays. Ms. Mendes communicated test results in one week and would sometimes allow students to appeal their grades if the class time is enough. She would fully explain her grading for students who spoke up for adjusting their grades. At the end of the course, Ms. Mendes adjusted the grades of failing students with at-risk backgrounds to support their success. She also increased marks for a few students to ensure admission into their desired universities. However, she lowered the grades of a few disruptive students who interrupted the classroom learning.

**Suggested revisions to Scenario 4:**

**To what extent do you consider each of these actions to be fair based on the above scenario?**
<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
<th>Alignment of actions to principles</th>
<th>Suggested revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ms. Mendes treated students respectfully during classroom assessment.</td>
<td>Respect</td>
<td>1 = not aligned to 5 = strongly aligned</td>
<td></td>
</tr>
<tr>
<td>2. Ms. Mendes largely considered student achievement of learning objectives in her grading.</td>
<td>Equality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Ms. Mendes detailed her grading criteria, with test scores making up 70% of a student’s grade.</td>
<td>Transparency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Ms. Mendes communicated test results in one week.</td>
<td>Timeliness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Ms. Mendes would sometimes allow students to discuss their grades if the class time is enough.</td>
<td>Voice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Ms. Mendes gave adequate justification for students who spoke up for their grades.</td>
<td>Justification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Ms. Mendes adjusted the grades of failing students with at-risk backgrounds.</td>
<td>Need</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Ms. Mendes considered students’ future university admissions to adjust grades.</td>
<td>Consequence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Ms. Mendes considered student misbehavior in her grading.</td>
<td>Bias suppression</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall, do the actions represent the various fairness principles relevant to the grading scenario? (1= Completely Unrepresentative, 5= Completely Representative)

**Scenario 5. Feedback**

Mr. Dembe has asked students to write an essay about their science lab project. He has shared with students a clear rubric that he will use to assess students’ essays. Four days after the essay submission deadline, Mr.
Dembe got back to students with his feedback and expressed that students can contact him for further discussion of his feedback. Students noticed that Mr. Dembe gave more feedback to students with good quality essays as well as his favorite students than students who had handed in essays that were of low quality. Mr. Dembe explained that he had given variable feedback on the essays based on the amount of effort he deemed each student had put in completing essays.

Suggested revisions to Scenario 6:

To what extent do you consider each of these actions to be fair based on the above scenario?

<table>
<thead>
<tr>
<th>Actions</th>
<th>Relevant underlying principle</th>
<th>Alignment of actions to principles</th>
<th>Suggested revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mr. Dembe provided a clear rubric for assessing students’ essays.</td>
<td>Transparency</td>
<td>1 = not aligned to 5 = strongly aligned</td>
<td></td>
</tr>
<tr>
<td>2. Mr. Dembe provided feedback after four days of essay submissions.</td>
<td>Timeliness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Mr. Dembe gave students a chance to further discuss his feedback.</td>
<td>Voice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Mr. Dembe provided feedback based on the amount of effort each student has put in the work.</td>
<td>Equity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Mr. Dembe did not provide more feedback to students who had weaker performance.</td>
<td>Need</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Mr. Dembe did not treat students’ respectfully in his feedback procedure.</td>
<td>Respect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Mr. Dembe gave more feedback to his favorite students.</td>
<td>Bias suppression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Mr. Dembe’s rationale that he provided to students about his feedback procedure was justifiable.</td>
<td>Justification</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Overall, do the actions represent the various fairness principles relevant to the feedback scenario? (1= Completely Unrepresentative, 5= Completely Representative)

Overall, do the scenarios represent adequately the significant domains in classroom assessment fairness? (1= Completely Unrepresentative, 5= Completely Representative). Please also provide any general feedback or suggestions regarding this inventory.

Thank you for your time and insights!
Appendix J
Recruitment Scripts for Phase I

First year undergraduate students who completed your secondary school in Ontario? Consider participating in a 30-minute interview focusing on your experiences of fairness in classroom assessment. You will receive 10$ at the end of the interview. Email amir.rasooli@queensu.ca.
Appendix K
Recruitment Scripts for Phase II

First year undergraduate students who completed your secondary school in Ontario? Consider completing this survey, focusing on your perceptions of fairness in classroom assessment (LINK). At the end of the survey, you will be entered into a draw lottery to win 1 of the 50 $10 gift cards for your participation. If you have any questions, please email amir.rasooli@queensu.ca.