

University Assessment Practices through a Lens of Feminist Pedagogy

by

Veta Chitnev

M. Ed., Moscow State Regional University, 1987

M.A., Simon Fraser University, 2014

A Dissertation Submitted in Partial Fulfillment  
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Curriculum and Instruction

© Veta Chitnev, 2019

University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

**Supervisory Committee**

University Assessment Practices through a Lens of Feminist Pedagogy

by

Veta Chitnev

M.Ed., Moscow State Regional University, 1987

M.A., Simon Fraser University, 2014

**Supervisory Committee**

Dr. Kathy Sanford, Department of Curriculum and Instruction  
**Supervisor**

Dr. Todd Milford, Department of Curriculum and Instruction  
**Departmental Member**

Dr. Tim Hopper, School of Exercise Science, Physical and Health Education  
**Outside Member**

### **Abstract**

Assessment is crucial for students' learning. The theory and practice of assessment, however, has received little attention in the literature on critical feminist pedagogy, due to difficulties in reconciling the notion of *feminist pedagogy* with that of *assessment*. This study aims to address that issue by exploring forms of assessment in higher education language learning that align with feminist principles. This research employs both case study and autoethnographic approaches. Data were collected from interviews with six university instructors and from the researcher's own teaching journal. All participants were full-time faculty at a department of modern languages and literature at a large Canadian research university. This study addresses three questions: 1) What forms of assessment do instructors in the Department of Modern Languages and Literatures use that comply with principles of critical feminist pedagogy? 2) What tensions exist for instructors related to assessment in higher education? 3) How can these tensions be addressed using a feminist reconceptualization of assessment? The study revealed that implementing graded summative assessment, especially when teaching large classes, creates pressure and discontent for university instructors. It was also found that formative teacher's feedback, peer and self-assessment, complete/incomplete grading, and diagnostic assessment reduce competition between students and serve to mitigate the power imbalance between students and teachers. Suggestions are provided to address the discontent and pressures reported by the participants and to reconceptualize assessment practices to bring them into alignment with feminist pedagogy.

*Keywords:* assessment, feminist pedagogy, university, higher education, evaluation, grading, policies

## Table of Contents

Supervisory Committee .....	ii
Abstract .....	iii
Table of Contents .....	v
List of Tables .....	ix
List of Figures .....	x
Acknowledgments .....	xi
Dedication .....	xiii
Epigraph .....	1
CHAPTER 1: INTRODUCTION .....	2
My Educational and Cultural Background .....	4
My Political Position .....	6
My Pedagogical Position .....	8
My Connection with the Topic .....	13
Positioning Myself as a Researcher .....	16
Summary .....	19
CHAPTER 2: LITERATURE REVIEW .....	20
Critical Feminist Pedagogy .....	20
The Term .....	20
The Roots .....	20
Connection with Feminist Theory .....	22
Definitions .....	24
Main Themes .....	25
Power as a Central Theme .....	28
Feminist Classroom Practices .....	32
Assessment .....	34
The Concept of Assessment .....	34
Summative Assessment .....	36
Less Common Types of Assessment .....	40
Purposes of Assessment .....	44
Assessment <i>for</i> , <i>as</i> , and <i>of</i> Learning .....	46
Learning Theories and Assessment Practices .....	50
What is Considered to be Effective Assessment? .....	53
Feminist Assessment .....	56
The Incongruity between Notions of Feminism and Assessment .....	56
Feminist Assessment from the 1980s to the 2010s .....	57
Features and Principles of Feminist Assessment .....	60
Feminist Assessment and Grading .....	62
Forms of Feminist Assessment .....	65
Conclusion .....	69

CHAPTER: 3 METHODOLOGY .....	71
The Rationale for the Methodology .....	71
The Rationale for a Case Study Approach.....	73
The Rationale for Autoethnography .....	75
Methods of the Research.....	76
Procedure .....	80
Participants.....	80
Gathering Data .....	81
Summary .....	86
CHAPTER 4: THE CATEGORIES.....	88
Developing a Structured Category Process.....	88
The Categories in relation to the Research Questions: Findings .....	91
Summary.....	94
CHAPTER 5: PARTICIPANTS' REPORTED ASSESSMENT PRACTICES.....	96
The Overview.....	96
Summative Assessment .....	96
Assessment Modifications .....	96
Participation .....	97
Important Aspects of Summative Assessment.....	100
Criteria for Summative Assessment.....	101
Summative Assessment as the Best Motivator for Learning.....	102
Summative Assessment through a Feminist Lens .....	102
Formative Assessment .....	106
Feedback on Written Assignments .....	106
Ongoing Feedback during Class Activities.....	107
Formative Assessment through a Feminist Lens .....	108
Complete/Incomplete Grading.....	109
Home Assignment.....	110
Presentations and Group Work .....	110
Complete/Incomplete Grading through a Feminist Lens.....	111
Peer and Self-Assessment.....	111
Peer Assessment as Peer Evaluation.....	112
Students Directed Seminars .....	113
Self-Assessment.....	115
Peer/Self-Assessment through a Feminist Lens.....	115
Diagnostic Assessment .....	116
Graded Diagnostic Assessment.....	116
Ungraded Diagnostic Assessment.....	117
Diagnostic Assessment through a Feminist Lens .....	118
Essential Aspects of Assessment .....	119
Focusing on Learning .....	119
Small Classes .....	120
Contextualized and Individualized Assessment.....	120
Caring.....	122
Assessing Teaching.....	123

Essential Aspects of Assessment through a Feminist Lens .....	123
Summary .....	125
CHAPTER 6: PRESSURES, TENSION AND DISCONTENT .....	127
Overview .....	127
Students' Values, Attitudes, and Capabilities.....	128
Focusing on Grades over Learning .....	128
Individualistic Values .....	131
Lack of Self-criticism, Gender-sensitivity, and English Language Fluency .....	132
Discontent with Students' Values, Attitudes and Capabilities through a Feminist Lens.....	133
Pressures from Students.....	134
Pressure on Female Instructors .....	135
Pressure for Higher Grades.....	135
Pressure from Students through a Feminist Lens .....	136
Pressure from the Administration.....	137
University Grading Policies.....	137
Conforming to Students' Expectations .....	139
Creating Competition between Instructors .....	140
Large Classes .....	141
Pressure from the Administration through a Feminist Lens .....	144
Discontent with Grading.....	145
Grading is Disliked but Necessary.....	146
Grading is Harmful and should be Avoided .....	148
Grading through a Feminist Lens .....	149
Self-Perceived Limitations in Assessment Practices .....	150
Gender Bias.....	150
Bias on Prior Achievements.....	151
Intolerance of Uncertainty .....	152
Self-Perceived Limitations in Assessment Practices through a Feminist Lens .....	152
Summary .....	154
CHAPTER 7: MY JOURNEY TO BECOME A FEMINIST EDUCATOR .....	157
Introduction.....	157
At the Beginning of my Journey.....	158
Challenges and Pressures.....	160
Can Summative Assessment be Feminist? .....	160
A "Recommended" Class Average Grade .....	170
My First-Year Student Anna.....	172
Challenges of Feminist Assessment in Large Classes .....	175
Internal Barriers for Feminist Assessment.....	178
Promising Assessment Practices for a Feminist Educator .....	179
Formative Assessment .....	180
Diagnostic Assessment .....	183
Complete/incomplete Grading .....	184
Peer and Self-Assessment.....	186
Teacher's Self-Assessment.....	188

Some Outcomes of my Journey .....	190
Were my Attempts Successful? .....	190
What Can Be Done?.....	192
CHAPTER 8: PULLING IT ALL TOGETHER .....	194
Assessment Aligned with Principles of Critical Feminist Pedagogy.....	194
Tensions Related to Assessment.....	196
Institutional Resistance to Alternative Assessment Approaches .....	197
Grades as a Market-Driven Commodity .....	198
Power, Authority and Hierarchy .....	199
Reconceptualization of University Assessment through a Feminist Lens.....	201
Conclusion .....	204
Bibliography .....	206
Appendix A.....	238
Appendix B .....	239
Appendix C .....	242
Appendix D.....	243
Appendix E .....	245
Appendix F.....	246
Appendix G.....	247

**List of Tables**

Table 2. 1 ..... 53

**List of Figures**

*Figure 3.1* Assessment thematic map 1 ..... 246

## Acknowledgments

I owe my deepest gratitude to the six faculty members participating in this study who were willing to contribute their time and to share with me their pedagogical practices and concerns. I hope this work will demonstrate my respect for the experience and views on pedagogy examined in herein.

I would like to express my appreciation and heartfelt thanks to the members of my supervisory committee. I am incredibly grateful to my research supervisor, Dr. Kathy Sanford, for her guidance from the very first year of my study at UVic, for sharing her knowledge about gender and pedagogy and for her continuous encouragement through the whole process of writing this dissertation. Her detailed reading, her generosity in granting me her time whenever I needed it, and her insightful, sensitive, inspiring and empowering feminist feedback turned my work on this dissertation into an exciting learning experience.

My deepest gratitude is owed to Dr. Tim Hopper, my first teacher at UVic, who taught me to be reflective and analytical and who provided insightful recommendations on how to refine my research work. I am greatly indebted to Dr. Todd Milford for his valuable comments that encouraged me to deepen my understanding of the topic of assessment. His expert insights significantly improved this work, particularly at the early stage of my research. Also, a big thank you to my external examiner, Dr. Kaela Jubas, for taking the time out of her busy schedules to be a part of my examining committee.

I could not have completed this work without the love and support of my family. I am particularly grateful to my husband, Alex, for believing in my success and for taking

care of me during this pursuit, and to my daughter Kate for her valuable advice and emotional support.

## **Dedication**

I dedicate this work to my loving daughter, Kate, who inspired me to undertake the dissertation.

## Epigraph

Ring the bells that still can ring.

Forget your perfect offering.

There is a crack in everything.

That's how the light gets in.

That's how the light gets in.

--Leonard Cohen, "Anthem," 1992

## CHAPTER 1: INTRODUCTION

In recent years, a substantial body of literature on the topic of assessment in higher education has emerged, indicating that this issue is central to contemporary pedagogy (Boud & Falchikov, 2007; Carless, 2015). In this study, *assessment* refers to the gathering information about students' knowledge and skills (Astin & Antonio, 2012; Gullickson, 2003; Mislevy, 2018) and the use of this information for the purpose of evaluation (Astin & Antonio, 2012). *Evaluation* refers to the determination of the quality or value of an object (Davidson, 2005; Gullickson, 2003), usually for the purpose of reporting or decision-making (Davidson, 2005). Assessment results are intended to inform students and teachers of students' progress, and also to help students develop the ability to identify gaps in their knowledge. Assessment, therefore, can and should be considered a reciprocal learning activity engaged in for teacher and students (Black & William, 1998). Research supports the importance of assessment in higher education for students learning and for program improvement (Bloxham & Boyd, 2007; Coutts, Gilleard, & Baglin, 2011; Higgins, Hartley, & Skelton, 2002). As a result, there is a substantial body of literature on the topic of assessment that describes various strategies of assessment and their practical development (Bloxham & Boyd, 2007), students' evaluation of courses and instructors (Spooren, Brockx, & Mortelmans, 2013), innovative practices, and recent discussion about appropriate assessment (Green & Emerson, 2007; Knight & Yorke, 2003). Much research focuses on assessment practices that help "monitor, control, and enhance quality in higher education" (Knight, 2002, p. 107). University teachers consider assessment to be a powerful motivational tool, and there is a

pervasive opinion that the majority of students seriously study assigned material when they are faced with assessment (Bloxham & Boyd, 2007).

Critical perspectives on assessment in higher education speak to the need to address issues concerning the questions of philosophy, social theory, politics, and complicated power relations that influence curriculum (Apple, 2012; Giroux, 1998). Both teachers and students should become aware of the ideologies that frame "taken for granted" knowledge and instructional practices (Giroux, 1998) and the pressure to judge through "performance assessment" (Apple, 2012 p. 16). *Ideology* refers to a set of beliefs which are taken for granted, such as the power relationship between teachers and students (Wallace, 2015).

Assessment as a judgement, a ranking of students through grading, that monitors and controls is incompatible with critical feminist pedagogy that is based on critical pedagogy and feminist theory, and referred to in literature as *feminist pedagogy*. Critical pedagogy "encourages the learner to reflect critically on issues of power and oppression" (Collins, & O'Brien, 2011, p. 112) while feminist theory is a system of theoretical discourse that addresses inequality with a central focus on gender (Buchanan, 2010, Collins, & O'Brien, 2011). *Inequality* refers to unequal access to various resources including economic, political and cultural (Griffin, 2017). Nowadays, radical feminists like bell hooks (2015) argue that feminist theory aims to radically transform society to end existing inequalities rooted in differences of class, race, gender, and sexuality and to eliminate power imbalance. Similarly, in education, feminist pedagogy, as noted by Woodlock (1995), recognizes power relations based on inequalities and inequities by

gender, class, race and sexual orientation is identified as unjust inequity. *Inequity* is defined as an unfair and unjust treatment (Griffin, 2017).

When assessment supports learning by engaging students in learning, it can be aligned with feminist pedagogy, whereas a patriarchal view of assessment used for monitoring and controlling students' knowledge runs counter to the idea of a feminist classroom. This study will examine assessment practices being used in the university today and analyze how they are compatible – or incompatible - with critical feminist pedagogy. In short, I would like to look at assessment in the university context through a feminist lens.

In this chapter, I situate myself politically, professionally, and culturally, because I strongly believe that no research is free of influence from one's personal set of values. In other words, my values, beliefs, assumptions, cultural background, gender, teaching philosophy, and political views have influenced not only my choice of the topic, but the direction and methodology of my research, as well as the presentation and utilization of its findings.

### **My Educational and Cultural Background**

I am a female educator with over 30 years of experience in both school and university settings. Originally from Russia, I have been living in Canada for over 20 years. Pursuing my dream of becoming a teacher, I graduated from the Moscow State Regional University for Teacher Training in 1987. Marxist theory, the official philosophy of the Soviet Union, and Vygotsky's (1978) social-cultural theory, which focuses on human development and learning through social interactions, constituted a significant part of the university's curriculum. Both theories have influenced my worldview greatly.

I believe in the political character of education, and I view learning as a co-constructed process that occurs through social interactions. My Russian collectivistic background and my fondness for Vygotsky's theory were the reasons for my adopting a social constructivist perspective toward learning theory, a perspective that rejects "the view that the locus of knowledge is in the individual" (Palincsar, 2005, p. 280) and considers learning and understanding to be "inherently social" (Palincsar, 2005, p. 281). Thus, I am still a strong proponent of Vygotsky's socio-cultural theory; however, the collapse of the Soviet Union and subsequent release of previously censored information about Russia's past seriously undermined my belief in the validity of Marxist ideology. As a result, after a critical review of all aspects of Marxist theory, particularly its privileged status as a national ideology and the crimes committed under its banner – I came to view Marxism as a social theory with both strong and weak points. I agree with some Marxist ideas, such as Marx's (1867/2011) notion of the exploitation of labour being the motivating force behind capitalism (Marx, 1867/2011) or Engels' (1884/2004) view of male domination of women as being not the cause, but rather "the effect of the economic oppression of women" (Engels, 1884/2004, p. 79).

Yet, I consider Marx's (1867/2011) ideas of revolution and the redistribution of property to be unjust coercion and a violation of human rights, and Engels's (1884/2004) proposed solution for ending the oppression of women by bringing them into to the labour force to be utopian. Indeed, women's involvement in production was merely an "effort to solve the short-fall of the labour force at women's expense," and it did not

eliminate men's domination over women in Russia (Watkins, Rueda, & Rodriguez, 1999, p. 93).

### **My Political Position**

I have come to consider myself to be a radical feminist. Since the understanding of the notion of *feminism* varies and the meaning of the term is not always clear (Aronson, 2003; Delmar, 1986; Thomson, 2001), I would like to elaborate on my current understanding of the term *feminism* and on *radical feminism* in particular. I view *feminism* as something more than a movement that focuses on political, economic, social, and cultural equality of the sexes. Gravett and Bernhagen (2018) claim that feminism addresses more than gender, it is “the interrogation of power, the honoring of perspective, the encouragement for reflection that makes us more aware of ourselves and our actions and more open and empathic to those around us” (p. 18).

I think that concentrating exclusively on women's oppression limits the feminist movement. Indeed, besides sexism, which refers to “prejudice or bias toward people based on their gender” (Glick, 2007, p. 859), there are other forms of oppression that negatively affect all members of society. Thus, “to be a *feminist* in any authentic sense of the term is to want for all people, female and male, the liberation from sexist role patterns, domination, and oppression” (hooks, 2015, p. 195). Boucouvalas and Lawrence (2010) argue that since “race, gender, and class are interlocking systems of oppression, particularly for people of color, they need to be considered as a collective” (p. 43). However, the position of women in society is my primary concern because of the global pervasiveness of gender inequity as well as the multiple dimensions of discrimination

women face if they are also unjustly treated due to other reasons, for example, their ethnicity, race, low socio-economic status, age, etc.

Two very inclusive definitions of radical feminism particularly resonate with me. The first, from Ware (1970), claims, “Radical feminism is working for the eradication of domination and elitism in all human relationships” (p. 3). I also agree with the following definition of Thomson (2001) who considers feminism “a social enterprise, a moral and political framework concerned with redressing social wrongs” (p. 7). I define myself as a *third-wave* feminist. The 19<sup>th</sup> century *first-wave* feminism focused on the full citizenship rights of women while *second-wave* feminism dated from the World War II until the 1990s expanded its focus to female solidarity, theorization of patriarchy, demand for equal pay and cultural representation (Mikell, Hawthorne, Moghadam, & Das Dasgupta, 2007). *Third-wave* feminism views gender activism within a broader range of problems to be addressed, and these concerns include social justice, and justice in such domains as the environment, the economy, race relations, sexuality, physical ability, and religion (Heywood, 2006).

As a radical feminist, besides being focused on gender justice, I am concerned with social justice and power distribution in all human relationships, respect for the common good, and community cohesion. Besides this broader agenda of concerns, the third-wave feminist movement attracts me with its two distinctive features: inclusiveness and a multifaceted perspective on individuals. Third-wave feminists include individuals belonging to diverse religions, races, ethnicities, economic status, and genders, and also comprises all humans as gendered beings (Heywood, 2006). In my everyday life, I strive to keep in mind the idea that no one can claim to be acting morally in every facet of their

life. From a third-wave perspective, in modern society one person can play different social roles and have different identities. The fact that the same person “can be exploitative in one context while being exploited in another, makes it impossible to ever see oneself in a position of an absolute moral high ground” (Heywood, 2006, p. xx).

### **My Pedagogical Position**

While working in the field of education as a Russian language teacher, both in school and in university settings, I have encouraged students to assume an active role in their learning, work in groups, participate in discussions, and share their ideas. The knowledge I gained through the Educational Psychology program at Simon Fraser University reinforced my commitment to constructivist teaching methods. In addition to the Vygotskian perspective on teaching and learning, I embraced Piaget’s cognitive constructivist theory, which also views learning as an active construction process. According to Piaget, knowledge acquisition involves self-generation built on prior learning experiences (Kaufman, 2004) , and students learn best when they are actively constructing their own knowledge, and when learning is individualized and student centered (Pardjono, 2016). Cognitive constructivist theory emphasized individualism and views teacher as a facilitator, while Vygotsky’s promotes social organization as an agent for change and views teacher/mentor as a proactive force that takes greater control in the process of learning (Glassman, 2001).

As a result, my teaching practices align with principles of constructivist foreign language teaching. These principles include learner-oriented instruction based on cooperative learning, learning by projects, intercultural awareness, and the use of authentic content (Aljohani, 2017). In my classes, I strive to create meaningful learning

environments for students with authentic contexts where students have diverse paths to discovery, collaboration, and personal control.

My teaching philosophy is also based on principles of critical feminist pedagogy, defined as classroom practices that “create classroom interactions that foster empowerment, community, and leadership” (Sandell, 1991, p. 181). Critical feminist and constructivist pedagogy are closely related. Both approaches value knowledge generated by students, and they both challenge the definition of knowledge as well as the process of acquiring it (Merriam & Brockett, 1997). Cronjé (2006) identifies the following features of constructivist pedagogy: experiential learning, cooperative learning, diversity, learner initiative and control over learning, the use of intrinsic motivation rather than external rewards, the teacher’s role as an egalitarian facilitator, and a balanced emphasis on process and product. I argue that these features are central to feminist pedagogy as well, and they will be discussed in detail in the next chapter.

I consider my classroom to be constructivist and feminist. I encourage students to construct their own knowledge by working in small groups and by engaging them in dialogue both with me and with one another, using their peers as resources. In my classes, I function as a facilitator who supports collaborative learning rather than a teacher who issues directives and dominates class time by lecturing. I believe that the primary goal of instructors is to provide students with opportunities to learn. Instructors should create a classroom environment that encourages students to explore and learn actively, while knowledge should come from the course content developed in cooperation with students.

One of the most prominent features that distinguishes critical feminist pedagogy from constructivism and other pedagogies is its strong focus on creating an egalitarian

classroom by addressing social justice issues, promoting inclusion, and power balances between teachers and students as well as between genders (Crabtree, Sapp, & Licona, 2009; Ellsworth, 1992; Manicom, 1992; Novek, 1999; Ropers-Huilman, 2009; Shackelford, 1992; Shrewsbury, 1993; Weiler, 1995). Critical feminist pedagogy emphasizes “the gendered nature of the traditional classroom” (Shrewsbury, 1993, p. 11), and feminist educators make sure that in their classrooms both women and men have equal opportunities and rights (Shrewsbury, 1993). Feminist educators attempt “to give students access to ‘better,’ more inclusive, socially just and nonexploitative knowledges” (Luke, 1996, p. 296). As a Russian language teacher, I provide this kind of knowledge by creating a positive classroom environment where all students are welcome and heard and by integrating material on social justice within my curriculum. As a radical feminist and a university language instructor who adopted a democratic teaching style, I strive to eliminate the hierarchy of authority in the classroom and to create an atmosphere of mutual respect in my classes. By *hierarchy* I refer to a system in which individuals are ranked according to their status, creating unequal relationships of subordination (Child, 2019)

I support the idea expressed by Renee Sandell (1991) that “each student has [a] legitimate entitlement and potential contribution to make to the subject matter” (p. 181). I encourage students to actively participate in the design of classes by eliciting their opinions on class activities and on the topics addressed in class. For example, for this purpose, I conduct an interest survey at the beginning of each course, and I find it particularly important because it creates opportunities for students to determine the course content and the topics for class discussions. By incorporating practices such as

this, I give students more choices and share the power in my classes. Giving students more choices and sharing the power ultimately reduce the power imbalance between students and me.

However, maintaining a balance of power has serious limitations in the current university setting where teachers may exercise their control over students by giving or withholding permission to submit assignments after deadlines, giving or withholding opportunities for make-up tests and exams, and by writing recommendation letters. Moreover, how can power be shared in university classes where the power so clearly lies with the instructor who evaluates students' learning?

Teaching has always been my passion, and I love every aspect of my work except grading which induces a lot of tension in me, and sometimes even anger. Indeed, evaluation is in conflict with my desire to create a classroom environment that relies on the principles of shared power. Maintaining a class average grade—that I should keep at a certain arbitrary number—is a considerable source of stress for me. I feel ashamed that part of me feels almost relieved when there are a few low-achieving unmotivated students in my class. For me, this means I can keep the class average grade around 75% and assign high grades to those who study diligently and perform well. As a result, I think that grading is not a meaningful activity from a pedagogical point of view. It meets the needs of the institution, not the needs of students -- who will ultimately lose because grading ranks and demotivates them rather than facilitates learning. Obviously, by evaluating and grading students, I judge students instead of supporting them, and I promote competition rather than encourage a sense of community. How can I be a feminist pedagogue if I am evaluating and ranking students by deciding who is the best

and who is the worst in class? How am I supposed to implement genuine peer-assisted learning in my classes if students know that by helping a schoolmate they may create a successful competitor? Thus, grading is harmful to the community because it generates competition between peers rather than cooperation and collaboration.

Furthermore, I strongly believe that grading ranks students, and this ranking “is inaccurate or unreliable, gives no substantive feedback, and is harmful to the atmosphere for teaching and learning” (Elbow, 1993, p. 188). *Feedback* is defined as “information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one’s performance or understanding” (Hattie & Timperley, 2007, p. 81). In contrast, feedback as ranking in the form of grading is harmful to students’ well-being because it creates a feeling of uncertainty and fear over a potential failure rather than interest in learning (Pulfrey, Buchs, & Butera, 2011). Thus, both my constructivist methods of teaching and my efforts to practice critical feminist pedagogy are hindered by the current evaluation system.

I believe that foreign language instruction is incomplete without teaching the culture of the language being taught. I therefore incorporate substantial cultural material into my lessons. My courses have a gender perspective on the Russian language and on Russian culture. The existence of the grammatical category of gender in the Russian language makes the gendered nature of the world and the presence of patriarchy more visible. In the notion of patriarchy, paternal power is the central meaning, i.e., “patriarchy refers to conjugal family relations or, more clearly, to generational gender relations (Therborn, 2004, p. 13). Students develop awareness of male dominance in Russian culture by reflecting on such common practices in Russia, as for example, the preference

of masculine forms of nouns denoting professions over their feminine peers when referring to women who succeed in these professions.

### **My Connection with the Topic**

My interest in assessment practices stems from my teaching experience. My perspectives on assessment underwent significant changes throughout my teaching career. I went through a phase during which I believed that by quantifying tests and assignments, I would develop perfect objective assessment tools with high validity and reliability. *Objectivity* or *objective judgment* is defined as one “that is free of judgments and bias” (Gaukroger, 2012, p. 4). *Validity* refers to the accuracy of assessment or whether it measures what it is supposed to measure (Bloxham & Boyd, 2007; Brown & Abeywickrama, 2010; Gronlund & Cameron, 2004; Yorke, 2008), whereas *reliability* refers to consistency of assessment results (Bloxham & Boyd, 2007; Brown & Abeywickrama, 2010; Gronlund & Cameron, 2004; Yorke, 2008).

However, at some point in my teaching practice, I realized that perfect assessment tools do not seem possible, at least when it comes to measuring students’ language abilities. All tests have limitations: “When we design a test, we cannot incorporate all the possible factors that affect performance” (Bachman, 1990, p. 31). Indeed, “in measuring language abilities, we are never able to observe or elicit an individual’s total performance in a given language” (Bachman, 1990, p. 34). Furthermore, objectivity also seems to be unattainable and in some instances not even desirable. As a constructivist, I believe that no judgment can be free of bias. Therefore, no matter how I quantify my assessment tools, my tests will be subjective because *I* create them and they will reflect *my* view regarding what should be measured and by what means. Hence, “a test score should be

interpreted as a reflection of a given test developer's section of the abilities included in a given syllabus or theory" (Bachman, 1990, p. 37).

I viewed assessment as a continuous process: I was assessing students' knowledge when I asked students questions or when I was listening to them working in groups. In class, I was assessing how well they understood new material. Outside of class, I made further assessment through home assignments and communicated it to the students by writing them feedback on what they had done well, but mostly on what they needed to improve. Years ago, I used to think that a good teacher was someone who "never ceas[ed] to assess students" (Brown, & Abeywickrama, 2010, p. 3). I tried to be always in control, while students did not participate in their own assessment of learning at all.

It was not until fairly recently, 2014, that my perspective on assessment underwent a dramatic change due to two important events: teaching upper-level Russian courses and my encounter with critical feminist pedagogy. The main challenge of teaching upper-level language courses was and still is establishing distinct and meaningful criteria for assessment and evaluation of students' knowledge. Unlike first-year Russian language classes, upper-level classes are designed to accommodate students with varying language levels including heritage speakers. In addition, even within these groups, there is considerable variability in students' pre-existing knowledge of Russian. Moreover, due to a various amount of exposure to the Russian language, every student in this class represents her or his own level of Russian. Therefore, the difference in prior experiences is the main factor that determines "success."

Thus, in addition to the lack of defined criteria, evaluation based on comparisons of students' language skills did not seem to be possible in this class. A heritage student

who had a rich regular exposure to the language early in her life would always perform better than a non-native speaker regardless of how hard the latter worked. My attempt to take student effort into consideration also failed because it was not clear how to determine effort. Moreover, for me the purpose of assessment is to encourage students to learn so that they become educated people. By *educated* people, I mean individuals who are empowered by knowledge and various skills, and who have developed abilities to think critically. The practice of assessing student's efforts creates a system of rewards and punishments that instructors assume students have invested.

Another important event that influenced my view on assessment was my encounter with critical feminist pedagogy. For various reasons, I always felt tension when I had to evaluate students' performance. At the beginning of my career, I did not question the benefits of grading for learning, but I doubted my capacity to provide "perfect" grading due to the lack of "clear" criteria. Knowledge gained from works of feminist pedagogues led to a revision of my perspectives on assessment and evaluation. I realized that my role as an assessor and evaluator contradicted my perspective on pedagogy. How can the balance of power in class be achieved if I am supposed to evaluate or in other words, judge each student's performance and assign grades? Moreover, how can I judge students' performance by using arbitrary criteria? Even grading scales developed collaboratively among all faculty members working in the program remain arbitrary numbers that do not represent all the skills and experience student gained and can use outside the class. This conflict has not been resolved and currently persists.

It is hard to deny: "Instructors are vested with power over a class from the moment class rosters are received until final grades are handed in. There is no escaping

the authority that comes with the roster” (Shackelford, 1992, p. 572). My role as an assessor and evaluator is challenging because it is in conflict with my desire to create a classroom environment that relies on the principles of shared power. Hence, it is crucial for feminist pedagogues to seek out effective democratic assessment practices that align with feminist principles. By *democratic* assessment, I mean an “assessment context in which normally marginalized voices are given a more active role and expression, i.e., a context in which traditional power relations are recognized and made more reversible and flexible” (Lynch, 2001, p. 368). I believe I am not the only educator who feels tension because of the conflict between the desire to achieve the balance of power in class and the role of assessor and evaluator. I also think that many educators would be willing to change their assessment methods to enact feminist assessment. However, to do so, it is crucial to find out whether feminist assessment is possible and what its features are.

Therefore, my study posed the following research questions:

1. What forms of assessment do instructors in the Department of Modern Languages and Literatures use that comply with principles of critical feminist pedagogy?
2. What tensions exist for instructors related to assessment in higher education?
3. How can these tensions be addressed using a feminist reconceptualization of assessment?

### **Positioning Myself as a Researcher**

I agree with Scotland’s (2012) claim that it is impossible to be involved in research without taking certain ontological and epistemological positions because they lead to the usage of different methodological approaches to the same phenomena.

Ontological assumptions are concerned with the nature of reality while epistemological

assumptions address the issue of the nature of knowledge (Wallace, 2015). In other words, ontology and epistemology are “what can be known and how it can be known” (Hamilton & Corbett-Whittier, 2012, p. 22). It is important for me to establish my ontological and epistemological positions and discuss how my view of reality and theory of knowledge has affected my choice of methodology and the topic I am choosing to research.

My constructivist approach to teaching and learning influenced my commitment to constructivism in educational research. I value constructivist perspectives because they “highlight subjective and intersubjective social knowledge and view this knowledge as core to understanding human phenomena” (Hershberg, 2014, p. 185). My personal experience of living in two different cultures, working in various settings and studying in two different graduate programs with different perspectives convinced me that there is no absolute truth. I started to appreciate the advantages of seeing the world as being comprised of multiple truths and different perspectives, which “need to be uncovered and unpacked in order to shed light on multiple and often competing realities” (Savin-Baden & Major, 2010, p. 1). As a constructivist, I believe that “truth is a consensus formed by co-constructors,” and that knowledge cannot be separated from its social context because it is “culturally derived and historically situated” (Scotland, 2012, p. 12).

This view of truth aligns with feminist research epistemology, which views knowledge as subject to the position of the researcher (Haraway, 1998). Thus, in research the following questions become important: “Whose truth is heard and validated? Whose perspectives are trusted and valued? Whose manner of communication is reinforced and whose is ignored?” (Ropers-Huilman & Winters, 2011, p. 671). I agree with Haraway

(1998), who proposes a new concept of “situated knowledge” as an alternative to positivistic *objectivity*. Haraway (1998) claims that positioning allows viewing from a perspective rather than from above and refuses “referring to science without differentiating its extraordinary range of contexts” (p. 591). Thus, Haraway (1998) emphasizes the importance of context and subjectivity for providing a better and richer account of reality.

Furthermore, feminist research emphasizes the *reflexivity* of the researcher, i.e., the capacity to analyze critically one’s own position (Griffin, 2019) and the importance of acknowledging the standpoints of researchers in the process of research and its outcome (Ropers-Huilman & Winters, 2011). Indeed, in order to produce a good qualitative study, researchers must be transparent not only about the design and execution of the study, but about their values, perspectives, and self-interests (Hiles & Čermák, 2007).

I recognize that individuals have a deep knowledge of their environment and their experiences are worthwhile for research. I therefore consider human interactions and interpretations to be a valuable source of research data. To explore the research questions indicated in the previous section with data gathered from multiple perspectives, I have chosen a qualitative study and a case study approach. I have used data obtained from six university instructors who volunteered to participate in this study and data from my teacher journal. I believe this study has the potential to produce social transformation by contributing to the development of better assessment practices and by changing the perspectives of all those who participated in this study. The open-ended interview questions aimed to explore the participants’ experiences in assessment allowed the participants to look more deeply at their assessment methods. Also, during the interviews,

a few participants shared some assessment and evaluation practices they had encountered in other educational settings, which I hope increased the participants' awareness and knowledge to what possible changes can be made in the current assessment practices.

### **Summary**

This study examined assessment practices being used in the university today and analyzed how they are compatible with critical feminist pedagogy. Assessment plays a vital role in learning and teaching, and the topic of assessment in higher education is at the center of attention in current pedagogy. I believe that my worldview impacted this study at every stage, from the choice of the topic to the utilization of the results. In this chapter, I have therefore provided information on my cultural, political and professional background.

As a radical feminist and a university language instructor with a constructivist perspective on learning, I see a disconnect between the notion of a feminist classroom that relies on principles of shared power and the current typical notion in university classes in my experience of assessment as a process that monitors and controls students. Furthermore, the practice of grading conflicts with my pedagogical perspective and calls into question the current evaluation system in higher education. I have conducted this study using a constructivist and feminist research epistemology, one that emphasizes context, subjectivity, reflexivity, and subjectivity. *Reflexivity* refers to the capacity to critically analyze one's own position (Griffin, 2019). I consider data collected from participants to be a valuable source of knowledge that may contribute to the development of better assessment practices.

## CHAPTER 2: LITERATURE REVIEW

### Critical Feminist Pedagogy

#### The Term

The term *critical feminist pedagogy* emphasizes the idea that this pedagogy has critical and emancipatory elements and emphasizes feminist theory, women's experiences and perspective. In education, the term is used as *critical feminist pedagogy* (English, & Irving, 2015), *feminist critical pedagogy* (Gur-Ze'ev, 2005), as well as its shortened version *feminist pedagogy*. English and Irving (2015) justify their use of the term *critical feminist pedagogy* "as it puts a deliberate stress on women and resistance to power in learning situations" (p. 105). However, the term *feminist pedagogy* appeared in the 1980s (Crabtree et al., 2009; Shackelford, 1992), and since then the majority of scholars prefer the original shortened version of the term. In this work, I will also use the term *feminist pedagogy* to align my work with the works of the majority of scholars I quote.

#### The Roots

A large body of literature supports the idea that feminist pedagogy has its roots in the critical pedagogy of the Brazilian educator Paulo Freire (1995) and in feminist theory (Allen, 2009; Crabtree et al., 2009; Luke, 1992; Shackelford, 1992). However, English and Irving (2015) consider Paulo Freire to be a bearer rather than the creator of the liberating ideas of critical pedagogy, and they argue that feminist pedagogy started much earlier with such women educators as Mary Arnold, Mabel Reed, and Jane Addams and with the social worker and philosopher Mary Parker Follett. The aforementioned three

educators worked in the field of adult education and were the original practitioners who inspired what later became critical pedagogy later (English & Irving, 2015), while Follett's theoretical contribution was identifying *power* as the key concept of critical feminist pedagogy –*power-with* as opposed it to *power-over* (Mendenhall & Marsh, 2010).

Yet, regardless of who should be considered the precursors of feminist pedagogy, feminist pedagogy and Freirean pedagogy have similar features, namely: affirming the existence of oppression, commitment to social transformation and the creation of a more just society, and a set of values based on “humanity and respect for students as knowers of the world” (Weiler, 2001, p. 74). Both pedagogies are learner-centred and produce knowledge essential for informed citizens who make responsible decisions in public and private life (Shackelford, 2003).

However, in contrast with feminist pedagogy, Freire dealt only with class-based oppression and failed to acknowledge the experiences of women (Weiler, 1991). Feminist scholars criticized Freire for not addressing such issues as the patriarchal assumptions of Western theorists and other forms of oppression based on gender and sexual orientation (hooks, 1994; Weiler, 1991). As an illustration of not addressing patriarchal issues, Weiler (1991) provides an example of the possibility of simultaneous positions of dominance and oppression that Freire did not look into—namely, when a worker is oppressed by his boss and at the same time is an oppressor of his wife at home. Hence, feminist pedagogy brings attention to the ways in which women are treated in various settings and how it affects women as a group (English & Irving, 2015).

Moreover, feminist pedagogy has expanded Freirean pedagogy by raising three areas of concern: the teacher's authority, personal experience as the source for knowledge, and the significance of recognizing difference and contradiction (Weiler, 1991). "What feminist pedagogy brings to the table is the power of questions, the use of inclusive teaching styles that challenge, and the stretch to have teaching reach to societal impact and change" (English & Irving, 2015, p. 106). The first feminist methods and techniques were used by instructors of women's and gender studies courses, but later they were "adopted by men and women teaching in various disciplines (Shackelford, 1992, p. 570).

### **Connection with Feminist Theory**

Feminist pedagogy refers to pedagogical approaches grounded in feminist theory and epistemology (Crabtree et al., 2009; Shackelford, 1992). Originating in the struggle of women for equal rights with men, feminist theory is considered to be a mode of critical theory with a focus on contemporary assumptions about gender, race, sex, sexuality, class, and nationality (Disch & Hawkesworth, 2016). Feminist pedagogy and feminist theory interact with each other because feminist theory is educational, and it has pedagogical and academic inquiry dimensions, such as "self-reflexivity, problematization of one's assumptions, the denaturalization of uninterrogated inherited knowledge, and so on" (McMahon, 2018, p. 14).

Feminist pedagogy and feminist theory share similar features. They both promote social change, avoid reductionist and essentialist conclusions, and revolve around the theme of power (Shackelford, 2003). *Reductionist theories* oversimplify complex phenomena while essentialist theories aim to distinguish which attributes are fundamental

to ontological status. The latter type of theorists, for example, identify sexual difference between men and women as an essential trait (Buchanan, 2010).

Feminist pedagogy and feminist theory are broad and multifaceted interdisciplinary fields that embrace many perspectives, which, however, share similar characteristics. Feminist theory and feminist pedagogy share three common features, namely: they make efforts to denaturalize socially constructed differences (in gender, race, classes etc.), they challenge the desire to produce universal impartial knowledge, and they consider the complexity of power relations (Disch & Hawkesworth, 2016). Similar to feminist epistemology that values community and claims that “epistemological community” generates knowledge (Nelson, 2015, p. 124), a feminist pedagogy promotes participatory classroom community (Hoffmann & Stake, 1998; Sandell, 1991). I will explore further these common features in a special section devoted to central themes of feminist pedagogy.

Both feminist theory and feminist pedagogy are continually in the process of development, and they share a basis in “ideologies, epistemologies, and methodologies that are negotiated and changing” (Shackelford, 1992). *Methodology* refers to the rationale of research, which is influenced by the epistemological and ontological views of the researcher (Wallace, 2015). Indeed, in the nineties, scholars rejected online courses as a potential feminist pedagogical environment because “the widespread electronic sharing of values has not resulted in a blossoming of relationships of care in the physical world” (Novek, 1999, p. 232). However, Hopkins (2009) challenges this assumption by applying feminist pedagogy in a live television women’s studies course and welcoming it “as an

opportunity to reach more people with the messages an analysis of society so important to women's studies" (p. 181).

### **Definitions**

There is no unified definition of feminist pedagogy. In the *Greenwood Dictionary of Education*, feminist pedagogy is defined as "a framework for teaching and learning that rejects the assumption of a universal, objective reality as perpetuated by socially privileged groups (e.g., straight white men). Feminist pedagogy emphasizes collaborative and self-critical analyses of the interaction of subjective experiences" (Collins, & O'Brien, 2011, p. 183).

Feminist scholars like Sandell (1991) tend to avoid definitions and focus on issues that feminist pedagogy deals with, such as "empowerment, community, and leadership" (p. 181). Yet there are a number of definitions of feminist pedagogy reflected in feminist literature. For example, Woodlock (1995) defines it as "teaching in a particular way: recognizing the relations of power—based on gender, class, race and sexual orientation—that permeate the classroom" (p. 4). Manicom (1992) argues that "feminist pedagogy is not a handy set of techniques," but a standpoint, and "whatever questions feminist teachers ask, they do so with a remarkable intensity, gazing inward, reflecting on their classroom practice, and outward, refining their critique of, and action in, the broader social world" (p. 365). A more recent definition by Crabtree et al. (2009) describes feminist pedagogy as "a set of assumptions about knowledge and knowing, approaches to content across the disciplines, teaching objectives and strategies, classroom practices and instructional relationships that are grounded in critical pedagogical and feminist theory"

(p. 2). Accardi (2010) defines feminist pedagogy as a form of critical pedagogy and “an educational approach informed by a feminist framework” (p. 27).

Thus, the existing definitions of feminist pedagogy are shaped by their authors’ personalities, values, interests and teaching styles, and there are two explanations for the lack of a unified definition. Accardi (2010) suggests that feminist educators deliberately resist limiting definitions, while Brown (199) thinks that “feminist pedagogy is still defining itself, largely through a process of questioning long-standing beliefs and practices in education” (p. 52). Nevertheless, despite the differences in definition, feminist educators agree upon several central themes or principles of feminist pedagogy.

### **Main Themes**

Shackelford (1992) points out that feminist pedagogy is not a monolithic movement. Consequently, instead of giving an exact definition of feminist pedagogy, she distinguishes its three main principles or themes. The first theme is interrupting the patriarchal order and giving power to all students (Shackelford, 1992). Instead of a reversal of power, feminist pedagogy “seeks to empower, to give voice and influence, to those who have been excluded from traditional power structures” (Shackelford, 1992, p. 571). The second theme is making the classroom more inclusive and providing those who are disempowered, e.g., women and minorities, with opportunities to express their ideas (Shackelford, 1992). The third theme addresses the process of teaching because “knowledge goes hand in hand with process,” which “suggests that how one teaches is as important as what one teaches” (Shackelford, 1992, p. 571). Thus, a fundamental principle of feminist pedagogy is the valuing of process over content by focusing on the “creation of a participatory classroom community that elicits full and open discussion

among students and faculty” (Hoffmann & Stake, 1998, p. 81). The process of teaching in the interactive feminist classroom includes attention to students’ personal experiences, encouragement of the questioning of widely accepted categories and authorities, and consciousness-raising activities (Hoffmann & Stake, 1998).

Manicom (1992) argues that the main goals of feminist pedagogy are social change and liberation. She puts forth three key themes of feminist pedagogy, namely, experience, collaboration, and authority. According to Manicom (1992), feminist pedagogical principles “include a critique of authority relations in the teacher-pupil relationship; a commitment to non-hierarchical, more communal co-operative classrooms; and the assumption that learning should begin in, and value students’ experiences” (p. 366).

Webb, Allen, and Walker (2002) synthesized previous works on feminist pedagogy, identifying six central themes or principles: “privileging voice, respecting the diversity of personal experience, reformation of the relationship between professor and student, empowerment, building community, and challenging traditional pedagogical notions” (p. 68).

The first two themes identified by Webb et al (2002), giving a voice to everyone, particularly to those who are oppressed, and respecting the diversity of personal experience are recurring topics in the feminist pedagogy literature (Crabtree et al., 2009; hooks, 1994; Noddings, 2013; Shackelford, 1992; Schniedewind, 1981; Webb, 2002; Weiler, 2001). According to hooks (1994),

The exciting aspect of creating a classroom community where there is respect for individual voices is that there is infinitely more feedback because students do feel free to talk—and talk back. And, yes, often this feedback is critical. (p. 42)

Feminist teachers put emphasis on collaborative learning experience, connectedness, and community (Novek, 1999; Schniedewind, 1981). By encouraging commitment to social responsibility and bonds of caring in students, feminist educators teach students how to view themselves as members of a community (Novek, 1999). Trust and student commitment can be achieved through various social gatherings outside of class too, for example, potlucks or field trips (Ellsworth, 1992). Finally, the sixth principle of feminist pedagogy is “challenging traditional views and practices” (Webb, 2002, p. 68). Feminist pedagogy considers education to be “inherently political and never neutral” (Monchinski, 2010, p. 16). Since “values are socially constructed,” they are “open to question and change” (Webb, 2002, p. 71). Knowledge in the feminist classroom “challenges the authority of traditional paradigms, showing them to be embedded in history rather than enshrined truths” (Maher & Tetreault, 1994, p. 5). Here a *paradigm* refers to a generalized model or to prevailing viewpoints in a discipline (Buchanan, 2010), where the discipline is understood as the practices of schooling students.

English and Irving (2015) identified four goals of feminist pedagogy. The first three goals are (1) building organizations in which students learn together, (2) creating social change, and (3) fostering social analysis by means of critically engaged pedagogy, and they overlap with the main themes described by Manicom (1992) and Webb et al. (2002). In addition to these three goals, English and Irving (2015) consider supporting

women's leadership as one of the main goals of feminist pedagogy "so that more women are positioned to make decisions and to lead change" (p. 105). Since feminist pedagogy is grounded in feminist theory, and "feminist theories are centrally concerned with the social and educational experiences of girls and women" (Luke & Gore, 1996, p. 193), supporting women's leadership should be included to the list of main themes.

Interestingly, neither Shackelford (1992) nor Webb et al. (2002) identify caring as one of important themes of feminist pedagogy. Yet Monchinski (2010) argues that "feminist pedagogy is firmly rooted in the ethics of care" (p. 10), and the traditional nurturing role is one of the central characteristics of feminist pedagogy (Accardi, 2013; Crabtree et al., 2009; Fisher, 2001; Lewis, 1992; Monchinski, 2010; Ropers-Huilman, 2010). Moreover, a feminist educator Nel Noddings (2013) considers caring to be a feminist approach to ethics and moral education. Noddings (2013) claims that "the primary aim of every educational institution and of every educational effort must be the maintenance enhancement of caring (p. 182). For Noddings (2013), in students' responses to teacher's questions, involvement and contribution are more important than whether the responses are right or wrong. "The student is infinitely more important than the subject matter" (Noddings, 2013, p. 186). One of the most important means of care is dialogue, through which a teacher learns more about a student, and having this knowledge is essential to being able to care effectively (Noddings, 2015).

### **Power as a Central Theme**

The theme of power is central in feminist pedagogy (Ellsworth, 1989; Manicom, 1992; Novek, 1999; Ropers-Huilman, 2009; Shackelford, 1992; Shrewsbury, 1993; Weiler, 1995). Ron (2013) defines *political or social power* "as generic terms to describe

a host of asymmetric social relations, such as force, coercion, domination, inducement, or persuasion” (p. 647). This definition of *power* aligns with Monchinski’s (2010) argument that traditionally *power* is perceived negatively as a form of repression, and it is associated with “violating the freedoms and rights of individuals” (p. 16). The definition of *power*, according to McCroskey and Richmond (1983), has a less negative connotation. They define *power* “as an individual's potential to have an effect on another person's or group of persons' behavior” or “as the capacity to influence another person to do something he/she would not have done had he/she not been influenced” (p. 175).

Tan (2004) identifies three notions of *power*: *sovereign*, *epistemological* and *disciplinary power*. Sovereign power cannot be shared, and it occurs when one individual has control over the actions of others while epistemological power can be shared or negotiated, and it occurs when teachers or institutions predetermine what students learn and what is valid knowledge (Tan, 2004). Finally, disciplinary power arises in discourses, and it is not limited to individuals (Tan, 2004).

In contrast to the notion of *power*, the notions of *powerlessness* or *disempowerment* received less attention in literature. TenHouten (2017) defines *powerlessness* as “means being subjected to domination by others and unable to live according to the dictates of one’s judgment and nature” (p. 139). Feminist pedagogues have always been concerned with power disparities existing within society and within educational institutions (Ellsworth, 1992). Feminist educators address the power imbalance between individuals on the basis of race and class, between privileged and disadvantaged individuals (Kenway & Modra, 1992; McClure; 2000), between men and women (Keyssar, 1985; McClure; 2000; Woodlock, 1995), and between teachers and

students (Ellsworth, 1992; hooks, 1994; Ropers-Huilman, 2009; Woodlock, 1995). For example, as far back as in the 1980s, Keyssar (1985) practiced various exercises to illustrate gender power imbalances in her university theatre classes. In one activity, students were assigned to imagine themselves being women or men with different amounts of power. Interestingly, a person assigned a card identifying “a woman with a ten power” had less perceived control of the situation than “a man with a five power” (Keyssar, 1985, p. 112). This exercise made visible that power depends on gender identity, and that even a very strong woman feels less powerful than a disempowered man.

A substantial body of literature on feminist pedagogy focuses on the analysis of power dynamics between teachers and students (Ellsworth, 1992; hooks, 1994; McClure; 2000; Ropers-Huilman, 2009; Woodlock, 1995). hooks (1994) claims that while she was a student, most of her teachers “were not the slightest bit interested in enlightenment. More than anything they seemed enthralled by the exercise of power and authority within their mini-kingdom, the classroom” (p. 17). In the same way, oppressive power relations occur not only in patriarchal classes but in classes considered by their instructors to be feminist. Manicom (1992) described several instances in which feminist educators admitted that oppressive practices occurred in their classrooms, such as privileging feminist voices over non-feminist ones, or lesbian-feminist voices over non-lesbian ones. Also, Keyssar (1985) shared her experience as an innovative theatre director at the University of California, and she mentioned that she selected students for a theatre course based on their understanding of feminism. Back in the eighties, she did not question her actions, whereas nowadays many readers would probably consider such selection

discriminatory because it violated students' freedom to hold their own political views as well as restricting their opportunities to learn.

In contrast to Ron's (2013) and Monchinski's (2010) negative view of power, hooks (1994) thinks that "power was not itself negative. It depends what one did with it" (p. 7). Many feminist educators create a positive image of power in feminist classes where power is shared, all voices are heard and validated, and students feel powerful because the professor creates an atmosphere of mutual respect (Ropers-Huilman, 2009; Webb, 2002; Weiler, 1995). Ropers-Huilman (2009) perceives power in a feminist class as fluid, therefore, always moving, and she shares her perception of teachers' and students' interactions as "a dance of power that was shaped by each person's talents, hesitations, limitations, and desires" (p. 45). In a feminist class, power and control shared between teachers and students change teacher-student relationships (Webb, 2002). Student or teacher status does not determine individuals' worthiness, and students perceive teachers' feedback as *a* perspective, not *the* perspective (Ropers-Huilman, 2009). On the whole, feminist pedagogy approaches power as potential, and feminist educators strive to increase the power of everybody—both teachers and students—in the classroom, whereas a patriarchally-constructed class treats power as domination, and the goal is to limit the power of subordinates (Shrewsbury, 1993).

However, McClure (2000) questions the possibility of creating a power balance in class, and she claims that feminist pedagogues "mask rather than expose power relation" (p. 54). To prove this point, McClure (2000) asks a rhetorical question: "What kind of message are we sending when we affirm and nurture students in class and yet send them home with a "C" or a "D"?" (p. 54). Feminist pedagogues do not provide a precise answer

to the question; nevertheless, feminist teachers have articulated their thoughts about power imbalances in the classroom (Allen, 2009; hooks, 1994; Shackelford, 2003).

hooks (1994) comments that at the beginning of her career as a university teacher she pretended that a power imbalance did not exist in her class, and she admits that this was a mistake. To decentre the power balance in class, feminist educators use various classroom practices described in the next section.

Another possible means of transforming the power dynamics of the academy from within is to encourage feminist leaders to move into administrative positions (Lloyd, Warner, & Baber, 2009). Lloyd et al. suggest that administrative positions provide feminist leaders with more opportunities to challenge hierarchies and put feminist ideas into practice. As administrators, feminists can create “democratic spaces so that the power to change is a ‘power-with’ others rather than ‘power-over’ others” (Lloyd et al., 2009, p. 295).

### **Feminist Classroom Practices**

There is no set of “rules” for feminist practices. There are, however, some common features of feminist classes that distinguish them from those of other pedagogical movements. To facilitate a non-hierarchical egalitarian classroom, feminist teachers employ various progressive and democratic practices: creating small groups, keeping journals, using role-playing and encouraging students to share their personal experiences (Freedman, 2009; Manicom, 1992; Weiler, 1995). Sánchez-Casal and Macdonald (2002) assert that an egalitarian non-hierarchical classroom can be achieved by pedagogical strategies that break teacher-student dyads. Feminist educators can create democratic classrooms by decentering the authority of the professor, creating discussion-

based classes, being able to manage conflicts, and questioning a pervasive assumption that the knowledge instructors produce “is automatically progressive” (Sánchez-Casal & Macdonald, 2002, p. 4). Progressive refers to favouring social reform or innovative (Progressive, 2019). Pedagogical strategies of feminist educators should create a collaborative classroom with democratic relationships between teachers and students, relationships of “equality, respect, and appreciation” (Sánchez-Casal, 2002, p. 68).

Allen (2009) discusses her experience of reflective teaching, in which she monitored her instructional practices and worked toward eliminating habits such as taking over class conversations. She also suggests the importance of “empower[ing] students by using strategies such as making assignments clear, giving later due dates, providing detailed instructions, offering extensive editorial and substantive feedback, and having writing workshops in and out of class” (p. 355).

In a feminist classroom, teaching methods using teacher-student dialogue prevail and small group discussions replace lectures, which encourage critical thinking and help recognize learners as knowers (Weiler, 1995). Feminist educators emphasize collaborative over competitive learning. They therefore use group projects and shared forms of leadership and decision-making (Shackelford, 1992; Schniedewind, 1981). For example, Shackelford (1992) established forums where students could express their thoughts as equals and could participate in the content of the course, while Schniedewind (1981) asked her students about their expectations from the course in a discussion at the beginning of the course.

Feminist classroom practices reflect feminist values (Manicom, 1992; Shackelford, 1992); therefore, feminist teachers prefer small classes, and connectedness,

sharing, and collaboration manifest in the spatial configuration of chairs in a circle (Manicom, 1992). Important distinguishing features of feminist classes are the inclusion of assignments and projects rather than lectures and the use of peer- and self-evaluation practices, which facilitate critical thinking skills and make feminist classes more egalitarian (Freedman; 2009; Manicom, 1992; Schniedewind, 1981).

## **Assessment**

### **The Concept of Assessment**

*Assessment* is a relatively new concept in education that came into use after World War II (Nelson & Dawson, 2014). The term originates from the Latin word *assidere* (to sit beside), and this was the way students used to be assessed, with a teacher sitting beside her students (Earl, 2003). At present, this term is difficult to define (Nelson & Dawson, 2014), and it has different meanings to different educators (Earl, 2003). According to Gullickson (2003), *assessment* is “the process of collecting information about a student to aid in decision making about the student's progress and development” (p. 5). Collins and O'Brien (2011) define *assessment* as “any method used to better understand the current knowledge that a student possesses” (p. 35), while Wallace (2015) describes *assessment* as the measurement of a learner's actual or potential level of knowledge (Wallace, 2015). Brown and Abeywickrama (2010) consider *assessment* as “an appraisal” of students' performance, which teachers may even make subconsciously, while Cox, Bradford and Miller (2016) define *assessment* as “means by which students' progress and achievement are measured, recorded and communicated to students and relevant university authorities” (p. 4).

Some scholars distinguish *assessment* and *evaluation*. For example, Brown and Abeywickrama (2010) view *evaluation* as the interpretation of assessment, and Gullickson (2003) defines *evaluation* as “the systematic investigation and determination of the worth or merit of [a student]” (p. 5). Bachman (1990) also distinguishes *assessment* and *evaluation*, arguing that the latter “is involved when the results of a test (or other assessment procedure) are used for decision making” (p. 22–23).

However, in some literature, the terms *assessment* and *evaluation* are used interchangeably (Collins & O'Brien, 2011), or some scholars consider evaluation to be part of assessment (Astin & Antonio, 2012; Hassan; 2011). For instance, Astin and Antonio (2012) claim that in higher education, the term *assessment* includes both activities: measurement, which refers to the gathering information, and *evaluation*, which refers to use of the information gathered for individual or institutional purposes.

In this work, *assessment* refers to collecting information about students' knowledge and skills (Astin &, 2012; Gullickson, 2003; Wallace; 2015) and the use of this information for various purposes (Astin & Antonio, 2012). I have chosen this definition for its breadth and inclusivity. I believe that any assessment includes evaluation because an assessor interprets, judges and determines the worth of each student's knowledge. Summative assessment involving *grading* that “takes place at the end of a course and provides the final judgement on, or “sums up,” the candidate's performance” (Wallace, 2015, p. 296), illustrates the idea that the term *assessment* includes both the process of measurement and evaluation of students' performance. This is the instructor's prerogative and imposed power by institutional hierarchy to determine the value/worth of students' work without the student deciding.

A significant number of authors in the literature on assessment distinguish three types of assessment: *summative*, *formative* or *diagnostic* (Bloxham & Boyd, 2007; Brown & Abeywickrama, 2010; Earl, 2003; Scaife & Wellington, 2010). However, Wallace (2015) classifies four types of assessment by including *ipsative* assessment in her classification. *Ipsative* assessment is designed to assess the learner's current level of knowledge and skills against their own previous level (Wallace, 2015).

### **Summative Assessment**

*Summative* assessment, as the term itself implies, summarizes what students have learned, and it has the following main functions: qualifications of proficiency and certification of students' knowledge in various fields, controlling of students' learning, and ranking students' performance (Brown & Abeywickrama, 2010; Gronlund & Cameron, 2004; Knight, 2007; Yorke, 2008). Brown and Abeywickrama (2010) define summative assessment as "a method of measuring a person's ability, knowledge, or performance in a given domain" (p. 3). A final examination with the use of numerical marks can be an example of typical summative assessment, which can be in the form of an essay or a test (Brown & Abeywickrama, 2010).

Summative assessment often involves evaluation and grading (Green & Emerson, 2007). Grading "is the process by which the work is assigned some code—usually A through F or a percentage—that represents the overall quality of the work" (Green & Emerson, 2007, p. 495). The use of grading involving either a letter grade or numerical marks in summative assessment is virtually universal, and "few (if any) students have escaped assessment that uses numerals" (Dalziel, 1998, p. 351). The first official record

of a grading system appeared in 1785 at Yale University, and since then grading systems remain highly debatable for their lack of coherence (Schinske, & Tanner, 2014).

Taking into consideration the pervasiveness of grading in higher education, the lack of theoretical frameworks for numeric scoring or letter grading in educational assessment is alarming, and it supports York's (2008) idea that "grading is not an issue that attracts a great amount of reflective thought in higher education" (p. 31). The lack of theoretical bases may be one of the causes that makes grading "one of the least liked, least understood and least considered aspects of teaching" (Green & Emerson, 2007, p. 495).

Available literature on the topic of summative assessment indicates that graded summative assessment does not convey reliable information about student achievement, and both letter and numeric grading systems are faulty (Dalziel, 1998; Green & Emerson, 2007; Knight, 2007; Sadler, 2014; Schinske, & Tanner, 2014; Yorke, 2008). Differential grading which refers to the phenomenon of students' receiving inconsistent grades for the same content and curriculum is rather a norm than exception (Beenstock & Feldman, 2016; Rauschenberg, 2014). Numerous studies have shown that grading criteria vary "between universities, between faculties within universities, between departments within faculties, and between courses within departments" (Beenstock & Feldman, 2016, p. 114). Yale University, where the history of the grading system started, has changed its grading system several times trying to achieve uniformity of grading system, but many inconsistencies persist (Schinske, & Tanner, 2014).

There is widespread evidence that written works given to various readers receive different grades, which proves the unreliability of grades and the lack of objectivity in

grading (Elbow, 1993). Bachman (1990) argues that even the so-called “objective” quantitative tests designed to assess student language ability are not perfect and completely objective. Therefore, Bachman (1990) recommends interpreting quantitative tests with caution because it is impossible to predict all factors that may influence student performance. Sadler (2014) claims that all attempts to standardize achievement when it can “be interpreted in a unique way by different people in different contexts at different times” failed because levels of student achievement in higher education have fuzzy boundaries (p. 275). Rauschenberg (2014) points out that differential grading occurs due to various factors, such as differences in instructor grading criteria, institutional policies, student behaviour, instructor bias, instructor quality, and curricula.

Besides the fussiness of grading criteria, Yorke’s (2008) analysis of the literature on summative assessment revealed the following problems with grading: a difficulty in attaining reliability, variations in grading outcomes between subject areas and institutions, bias not only in the humanities and social sciences where alternative views are to be expected, but also in science-based subjects. Thus, summative assessment is compromised by grading, which is not measurement, but rather judgment (Yorke, 2008); therefore, it can never be an exact science (Lewis, 2006). Yorke (2011) considers grading to be “quasi-measurement” because it is affected by a number of socially-driven factors whose influence is generally not well understood” (p. 251). Moreover, grading labels students as “winners” and “losers” in the grading hierarchy, and this labelling can demotivate students and compromise their self-esteem (Hughes, 2017).

However, despite abundant evidence for problems with the statistical manipulation of grades (Yorke, 2008), educational institutions consider this statistic

seriously when they take steps to prevent grade inflation while disregarding such factors as improvement in assessment (Yorke, 2008) or in teaching (Lewis, 2006). *Grade inflation* refers to “a trend toward students receiving higher grades, usually with the implication that higher grades are being given for the same quantity and quality of work that previously would have received lower grades” (Collins & O'Brien, 2011, p. 204). A review of research on grade inflation suggests a lack of support for grade inflation in higher education and a decline in students' proficiency (Kohn, 2002).

Yorke (2008) distinguishes two types of summative assessment according to the grading system used: *norm-referenced* assessment—when grades are normally distributed across a curve—and *criterion-referenced* assessment. Criterion-referenced assessment measures how well a student performs against an objective or standard for the course, while norm-referenced assessment is relativistic and “seeks decimation amongst students by placing their achievements in order of merit rather than by setting them against the kinds of norms that are developed for psychological and mass educational testing” (Yorke, 2008, p. 17).

Norm-referenced assessment was designed as a solution to the subjective nature of grading to minimize differences between grading criteria in different institutions (Schinske, 2014). Many scholars argue for criterion-referenced assessment as a more “objective and knowledge related grading system” (Hassan, 2011, p. 331). Moreover, some scholars question whether a form of assessment, in which students are “set against one another in a race for artificially scarce rewards” is ethical (Kohn, 2002, para. 24). However, norm-referenced assessment is pervasive in higher education, if not in policy but in practice with imposed curving of grades as a means of controlling grade

inflation (Hassan, 2011). This form of assessment is easy to conduct and it is used to sort out learners (Akyea & Sandoval, 2005).

In contrast to assessment that involves numeric or letter grading, assessment based on pass/fail or complete/incomplete grading focuses on the competency of an outcome rather than on earning a grade (Manning, Ference, Welch, & Holt-Macey, 2016). Pass/fail, or complete/incomplete grading intrinsically motivate students and significantly reduces stress (Manning et al., 2016).

### **Less Common Types of Assessment**

In contrast to summative assessment, diagnostic and formative assessment have received less attention in higher education (Scaife & Wellington, 2010; Yorke, 2003). Furthermore, both types of assessment are not as well understood in higher education as they are in the school context (Alderson, 2017; Yorke, 2003, 2005). Also, Scaife and Wellington (2010) report a lack of common understanding of the terms *formative* and *diagnostic* assessment in university policy documents.

Diagnostic assessment is usually carried out at the beginning of courses to evaluate a learner's strengths, weaknesses, and areas for improvement (Doe, 2014; Wallace, 2015). Diagnostic assessment is "designed to identify skills, knowledge and other attributes which can be used to decide on specific pathways of study, or difficulties in learning which require support" (Bloxham & Boyd, 2007, p. 233). Early diagnostic assessment is important for assessment of student progress, evaluation of program effectiveness, and for early identification of students with some learning difficulties to refer them to support services (Bloxham & Boyd, 2007). Diagnostic tests are widely used in foreign language classes in the form of placement procedures that include paper-and-

pencil and online tests, oral interviews, and combinations of these procedures (Brantmeier, 2006). There is evidence for the effectiveness of a post-enrolment diagnostic English language assessment program designed to identify students' strengths and weaknesses in the English language that provides recommendations for English language support such as English courses, tutorials, or visiting a Student Learning Centre (Doe, 2014).

In contrast to diagnostic assessment, formative assessment is performed during instruction, and it aims to inform students' competence by providing them with appropriate feedback on their academic progress (Brown & Abeywickrama, 2010; Gronlund & Cameron, 2004; Knight, 2007; Yorke, 2008).

Formative assessment is defined as the process of shaping students' competencies and skills with the goal for helping them to continue that growth process. The key to such formation is the delivery (by the teacher) and internalization (by the student) of appropriate feedback on performance, with an eye toward the future continuation (or formation) of learning. (Brown & Abeywickrama, 2010, p. 7)

Thus, in contrast to summative assessment that tends to be one-time, formative assessment is ongoing and blended with learning. The most important feature in the formative assessment is the quality of formative feedback (Black & William, 1998; Higgins, Hartley, & Skelton, 2002). The term *quality* is a value-laden term and means different things to different people (Harvey & Green, 1993). In education, the quality of assessment is connected with learning (Kvale, 2007).

There is well-documented research to suggest that formative assessment improves learning (Astin & Antonio, 2012; Black & William, 1998; Gibbs & Simpson; 2005; Wiggins, 1993; Yorke, 2003). Moreover, Black and William (1998) found that students read feedback without a mark more carefully, and they ignore comments accompanied by a grade (Butler, 1988). By comparing student performance after students were given numeric grades and no feedback, various types of detailed feedback and grades as a component of feedback, and feedback without grades, Lipnevich and Smith (2009) found that detailed, descriptive feedback in the absence of a grade is more effective than similar feedback accompanied by a grade. A growing number of scholars value descriptive feedback and formative assessment over grading (Fu, Hopper, & Sanford, 2018).

Nicol and Macfarlane-Dick (2006) identified seven principles of good feedback, according to which feedback provides high-quality information to students about what good performance is; facilitates students' progress in self-assessment; provides high-quality information to students about their learning; encourages discussions of learning with teacher and peers; inspires positive self-esteem and provides motivational opportunities for moving from existing to desired performance; and informs teachers on what may help improve their teaching practices.

Kvale (2007) and Yorke (2005) emphasize the importance of speediness and accuracy when providing formative feedback. There is well-documented evidence that "the more immediate and precise the feedback is, the stronger the effect of learning" (Kvale, 2007, p. 58). Yorke (2005) identifies sensitivity as another critical consideration in teacher feedback and perceives formative assessment as a form of conversation between

the teacher and student rather than an assessment form that fulfills technical requirements of reliability and validity. As for sensitivity, Yorke (2005) warns that some students may take a teacher's comments on their work personally. Furthermore, according to the qualitative study conducted by Ruper Taggart and Laughlin (2017), after receiving "taught" feedback from a teacher, some students felt disrespected, experienced confusion, the feeling of being low down in a hierarchy, and reported "that their agency or authority had been usurped rather than when productive collaboration, negotiation, and integration of new perspective occurred" (p. 1).

However, similar to summative assessment, formative assessment has some disadvantages. For example, it may discourage students from learning (Yorke, 2003). It is also time-consuming for teachers and becomes therefore problematic to use in large classes (Hassan, 2011). One of the possible solutions may be group learning assessments, like a group project subdivided into individual parts for each student (Hassan, 2011).

Recently, another type of summative assessment has re-appeared in North America, *ipsative* assessment, which in contrast to criterion-referenced and norm-referenced assessments determines a student's progress based on their previous work rather than on external criteria or on the work of other students (Hughes, 2017). Ipsative assessment can also be formative when ipsative feedback is provided (Hughes, 2017). When comparing ipsative feedback with conventional feedback, Hughes (2017) considers ipsative feedback more motivational and beneficial for learning because it gives student power in the assessment process rather than comparing to others addresses a student's progress rather than the gap between the student's knowledge and an external goal.

## Purposes of Assessment

*Summative, formative* and *diagnostic* assessment are not the clear-cut terms they might seem to be at first sight. Scaife and Wellington (2010) claim there is no clear distinction between formative and diagnostic assessment, while Yorke (2003) reports a lack of clarity in distinguishing between formative and summative assessment. For example, there are assessment tasks that are both formative and summative (Yorke, 2003). Graded assignments within modules that contribute to the overall grade “act formatively if the student learns from them” (Yorke, 2003, p. 479). For example, credit-carrying, low-stakes short reading quizzes promote student preparedness to work in class and inform faculty about student learning and possible challenges (Hodges, Anderson, Carpenter, Cui, Gierasch, Leupen, & Wagner, 2015). Implementing several summative assessments during the course period instead of a high-stakes final exam also refers as *continuous* assessment (Day, van Blankenstein, Westenberg, & Admiraal, 2018). Studies show that continuous assessment improves learning because it creates the conditions for distributed practice over time rather than cramming before the final exam, and it allows students to reflect on their learning (Day et al., 2018). Also, there are various forms of negotiated assessment when students enter the negotiation over the construction of their assessment is an approach that creates conditions in which student learn during assessment period (Brna, Self, Bull, & Pain, 1999). For example, one of the forms of negotiated assessment involves a contract between students and teacher, in which the teacher describes the work involved to get particular grades and students agree to do the work according to the grade they want to receive (Schniedewind, 1981). Therefore, Ussher and Earl (2010) argue that the purpose of assessment is more important than the

assessment task itself, and *summative* and *formative* assessment should be replaced by the terms *assessment for summative purposes* and *assessment for formative purposes*.

Although many scholars agree that enhancement of student learning is one of the main purposes of assessment, classifications of purposes of assessment in higher education vary (Astin & Antonio, 2012; Bloxham and Boyd, 2007; Ussher & Earl, 2010; Yorke, 2008). Astin & Antonio (2012) identify four purposes of assessment: admissions, placement, certification, and classroom learning. Summative assessment is used for certifying and screening, and in classroom learning, it operates as “a carrot and a stick, depending on whether the student is oriented toward attaining success or avoiding failure” (Astin & Antonio, 2012, p. 196). In contrast to summative assessment, Astin and Antonio (2012) consider formative assessment to be an effective tool for classroom learning assessment, one that is “intended to serve the goals of teaching and learning: to facilitate student learning and talent development and to make the learning process itself more rewarding” (p. 201).

Bloxham and Boyd (2007) selected four main purposes of assessment: certification, promoting learning, quality assurance, and encouraging students to undergo lifelong learning. In this classification, certification is the means to identify achievements, providing licensing and selection for further studying, whereas the quality assurance purpose serves as a gauge for program standards (Bloxham and Boyd, 2007).”

Yorke (2008) identifies only three significant purposes of student assessment: learning, certification, and quality assurance. However, each main purpose includes more detailed sub-purposes. Among those itemized by Yorke (2008), for example, are developing motivation, detecting strengths and weaknesses, providing feedback, offering

help in developing skills in self-assessment, and establishing a known level of achievement at the end of each unit into the learning purpose. Quality assurance also serves several more detailed purposes, from judging the effectiveness of the learning environment and the teacher's performance in it to protecting the public (Yorke, 2008). Interestingly, neither Yorke (2008) nor Bloxham and Boyd (2007) indicate what type of assessment should be used to serve these purposes; however, Black and Wiliam (2009) focus mostly on formative assessment, and they emphasize improving learning environment as well as student learning as an important purpose of assessment.

Assessment purposes sometimes support one another and sometimes compete (Bloxham & Boyd, 2007; Earl, 2003). For instance, Bloxham and Boyd (2007) argue that there is a significant difference between assessment for certification and assessment for learning. "An overriding concern for reliable marking may prevent the use of group assignments or may encourage use of assessments that usually foster low-level learning such as multiple choice question tests" (Bloxham & Boyd, 2007, p. 40). Fu, Hopper and Sanford (2018) suggest distinguishing assessment types by the audience they are intended for: formative assessment is more useful for students and teachers, while summative assessment can be helpful for policymakers and administrators.

### **Assessment *for*, *as*, and *of* Learning**

Earl (2003) proposes an alternative classification of assessment in relation to its impact on learning and distinguishes assessment *of* learning, assessment *for* learning and assessment *as* learning. Assessment *of* learning is summative and the results are recorded and become public, whereas assessment *for* and *as* learning are formative (Earl, 2003). Studies have shown that the value of formative feedback is more associated with learning

as opposed to summative feedback (Black & William, 1998; Fu et al., 2018; Wiggins, 1993).

Assessment *for* learning includes personalized descriptive feedback from the teacher, which provides the necessary conditions for teachers to adjust their instruction and for students to adjust their learning (Earl, 2003). Thus, in the assessment *for* learning approach, assessment information is used to improve instruction because teachers modify learning according to assessment results (Davies & LeMahieu, 2003). Furthermore, assessment *for* learning is concerned with creating transparent learning goals and encouraging intrinsic motivation. “Educational assessment may thus facilitate learning by providing extensive feedback, clarifying the goals for learning, motivating further learning, and encouraging active learning” (Kvale, 2007, p. 58).

In assessment *as* learning, Earl (2003) views the role of students as more important: “Students, as active, engaged, and critical assessors, can make sense of information, relate it to prior knowledge, and master the skills involved” (p. 47). This goal can be achieved by involving students in the assessment process through various means: self-assessment, peer assessment, authentic assessment, and portfolio assessment (Davies & LeMahieu, 2003; Klenowki, Askew, & Carnell, 2006; Kvale, 2007).

*Self-assessment*, or students’ participation in their own assessment, increases students’ involvement in learning, helps to develop critical attitudes towards their work, increases the capacity for lifelong learning after graduation (González-Betancor, Bolívar-Cruz, & Verano-Tacoronte, 2019), and has the potential to empower students (Tan, 2004). Formative self-assessment is the most useful and worthwhile form of assessment

for learning because it allows students to reflect on their efforts and progress (González-Betancor et al., 2019).

*Peer-assessment* refers to assessment occurring among individuals of equal status and at university these are usually students from the same class. The main purpose of peer-assessment is to enhance both the assessor and the assessee (Topping, 2012). Usually, peer-assessment is reciprocal, and the participants are each other's assessors and assessees (Topping, 2012). Panadero (2016) points out that peer-assessment is a collaborative learning exercise because assessees receive direct feedback that helps them to improve, while assessors, by evaluating the work of their peers, learn their own strengths and weaknesses.

Research reports inconsistent results regarding summative peer and self-assessment scores in comparison to the grades assigned by instructors. Most research indicates that students are able to make reasonable judgments of their performance, and their marks are consistent with faculty grading (Lew, Alwis, & Schmidt, 2010). Some studies report that student scores were higher than the scores assigned by faculty (De Grez, Valcke & Roozen; 2012) or lower than the faculty scores (Dochy, Segers, & Sluijsmans, 1999; Wagner, Suh, & Cruz; 2011) Students ability is an important factor that influences the scores, and student with higher abilities tend to assign lower scores to themselves and their peers (Dochy et al., 1999; Uttl, White, & Gonzalez; 2017). Also, students assign, for the most part, lower scores at the beginning of the term and higher scores as the term progresses because they gain more confidence in themselves (Dochy et al., 1999).

*Authentic* assessment involves methods that focus on real problems in real-world contexts (Gronlund & Cameron, 2004), while *portfolio* assessment is defined as “a systematic and selective collection of student work that has been assembled to demonstrate the students’ motivation, academic growth, and level of achievement” (Norton & Wiburg, 1998, p. 237). Portfolio assessment is a multidimensional process of gathering evidence to prove a student’s accomplishments and progress (Davies & LeMahieu, 2003). Portfolio assessment may include various types of work: assignments, presentations, group projects, research papers, and many other projects (Arnold, 1992). Recently, *electronic portfolios*, digitized collections of a student’s work, have risen in popularity in higher education as an assessment tool that promotes transparency and facilitates visibility in learning (Curtis & Wu, 2012). Jimoyiannis and Tsiotakis (2016) claim that electronic portfolios promote student engagement, self-regulation and communication. Van Aasist and Chan (2007) found that a student-directed electronic portfolio supports both individual and collective learning.

In assessment *as* learning, the role of students is more important: “Students, as active, engaged, and critical assessors, can make sense of information, relate it to prior knowledge, and master the skills involved” (Earl, 2003, p. 47). In assessment *as* learning, students must be able to set goals of learning, monitor their progress and be able to reflect on the results (Earl, 2003). Wiggins claims (1993), “Assessment is not something we do to students but something we do with them...” (p. 111). Both students and teachers are responsible for learning since students develop an ability to recognize and remediate these gaps (Black & William, 1998), while the teacher’s role is to guide students so that they become competent assessors (Earl, 2003).

Taking responsibility for their own assessment makes students independent lifelong learners (Davies & LeMahieu, 2003). While all three types of assessment, assessment *of*, *for*, and *as* learning have a place and purpose, Earl (2003) contends that if authentic learning for students is the goal, assessment *as* learning must receive the greater emphasis.

From a feminist perspective, assessment *for* learning and assessment *as* learning challenge traditional top-down assessment practices and facilitate the creation of a participatory classroom community. These learner-centered progressive assessment practices empower students by involving them in their own assessment and by making them responsible for their own learning.

### **Learning Theories and Assessment Practices**

Assessment practices are closely related to learning theories, which in turn vary depending on the school of thought adhered to (Behlol & Dad, 2010; Hassan, 2011; James, 2013). Since there are numerous philosophical and educational schools, each with its own perspective on learning, it is beyond the scope of this work to describe each of them. In order to illustrate how learning and assessment theories are aligned, I have chosen the following learning schools: behaviourism (Pavlov, 1928; Skinner, 1968), the cognitivist-constructivist school of learning (Piaget, 1973) and the cultural-historical theory (Vygotsky, 1998), which in the West referred to as socio-cultural theory (Collins, & O'Brien, 201; James, 2013). These three educational philosophies are considered to be the most notable and influential (James, 2013).

Behaviourism views learning as “the conditioned response to external stimuli” and asserts that “complex wholes are assembled out of parts” (James, 2013, p. 123). Learning can thus be achieved by the application of rewards and punishments and by deconstructing complex performances into elements and practicing them (James, 2013). Behaviorists redefined “schooling as specific objectives, individualized instruction, and the teacher as managers” (Calfee, 1994, p. 340). Behaviorist theory thus views learning as a passive process, which Watkins (2003) defines as “learning as being taught.” Proponents of behaviourist theories of learning value summative assessment by means of timed tests, which are interpreted as either right or wrong (James, 2013).

According to the cognitivist-constructivist school (Piaget, 1973), learning is a way of being and an ongoing pattern of actions and attitudes (Behlol & Dad, 2010). Since knowledge is not a copy of the world, it cannot be transferred, but it can be constructed by learners according to their understanding of the world (Kocadere & Ozgen, 2012). Modern theories of learning are closely related to the cognitivist-constructivist school of learning, and they have similar common characteristics: an emphasis on understanding rather than memorization, a deep knowledge of important theoretical frameworks, building learners’ knowledge on the basis of their previous knowledge, and learners’ active involvement and control of their own learning (Assessment and Reporting Unit, Office of Learning and Teaching, 2005). Since eliciting a student’s prior knowledge through dialogue and open-ended assignments is important, teaching and assessment are blended, and formative assessment becomes an important assessment component (James, 2013).

In contrast to the principles of the individual construction of knowledge proposed by the cognitivist-constructivist school, according to the socio-cultural theory (Vygotsky, 1998), learning occurs during interactions with other students and teachers, and during participation in activities (Hassan, 2011). Watkins (2003) defines learning in this school “as part of doing things with others” (p. 10). The concept of *the zone of proximal development*, which refers to the distance between learner’s actual and potential competence, plays an important role in the Vygotsky's socio-cultural theory (Collins & O’Brien, 2011; Wallace, 2015). The teacher monitors student’s knowledge, and provides tasks within the learner’s zone of proximate development, while the learner successfully accomplishes this task with the help of others (Wallace, 2015). To be congruent with the idea of creating knowledge through social interactions, the assessment tasks should be collaborative and solved by students and teachers together (James, 2013). For example, James (2013) describes assessment of presentations conducted by the teacher who is the main tutor, another supervisor, and a peer selected by the student. The criteria have been negotiated at the beginning of the preparation for the presentation, and fifty percent of the time is allocated to the presentations and the other fifty percent to discussions, which would include reflections on the work covered and on the learning process (James, 2013).

Currently, the behaviourist approach is used in science tests that involve the recall of facts and when the focus of learning is on the development of basic skills (James, 2013). When deep learning is essential, the majority of the educational researchers claim to take either the cognitivist-constructivist or the cultural-historical school’s perspective (James, 2013). In recent decades, learning is considered to be meaningful if learners construct their own knowledge, one that helps them make sense of the world, reflect on

their own and others' perspectives, and solve complex problems. Learners shape and are shaped by the community (James, 2013), and learners are considered to be effective if they are self-motivated, self-regulating, knowledgeable of various learning strategies, and are ready to continue their learning (Segers, Dochy, & Cascallar, 2003). Within this learning theory, learners take an active part in the assessment process by exchanging information, ideas and experiences (Fa et al., 2018). The information on learning and assessment theories is presented in Table 2.1.

**Table 2. 1**

**Assessment and Theories of Learning**

Learning Theory	Assessment Practices
behaviourism (Pavlov, 1928; Skinner, 1968)	summative assessment by teacher
socio-cultural theory (Vygotsky, 1998)	formative assessment by teacher/ group-assessment/peer- assessment
cognitivism-constructivism (Piaget, 1973)	formative assessment by teacher/ self-assessment

**What is Considered to be Effective Assessment?**

According to the Merriam-Webster dictionary, the word *effective* means producing desired results (Effective, 2003). Therefore, effective assessment varies depending on its type and purpose, and many of the characteristics of an effective assessment task conducted for formative purposes would differ from criteria for a

summative assessment task. For instance, Gronlund and Cameron (2004) primarily describe summative assessment when they claim that effective assessment has to satisfy eight conditions. These conditions include a clear understanding of all intended learning outcomes, usage of a variety of assessment procedures, the presence of instructional relevance, a representative sample of each student's performance, comprehensive grading, the specification of criteria for judging, feedback to students that emphasizes strengths and weaknesses, and fairness. By "fair assessment," Gronlund and Cameron mean an assessment in which irrelevant difficulties, ambiguous tasks and bias are eliminated.

In contrast to Gronlund and Cameron (2004), Boyle and Rothstein (2006) and Brown and Abeywickrama (2010) are more cautious in their description of fair summative assessment. Rather than giving characteristics of fair assessment they describe factors that influence student perception of fairness. For example, Boyle and Rothstein claim that students perceive a transparent grading system as fair when it gives them some control over their grades. Brown and Abeywickrama argue that faculty may increase student perception of their assessment as fair if their exam or test is well-constructed in an expected format, contains a reasonable challenge and clear items that have been rehearsed during the course, and can be completed within an assigned time limit.

In contrast to summative assessment, formative assessment has different criteria for its effectiveness. Earl (2003) considers formative assessment to be effective if it empowers students to ask reflective questions and consider a range of strategies for learning and acting . . . students move forward in their learning when they can use their personal knowledge to construct meaning, have skills of self-monitoring

to realize that they don't understand something, and have ways of deciding what to do next. (p. 25)

Similarly, according to Astin and Antonio (2004), effective formative assessment should

1. be multidimensional and involve observation of both the learning process and the learning outcomes;
2. be “intended to serve the goal of teaching and learning: to facilitate student learning and talent development and to make the learning process itself more rewarding” (p. 201);
3. be “an ongoing iterative” process rather than a one-time activity carried out at the end of the learning process” (p. 202);
4. constitute a significant part of teacher's activities;
5. be useful for teachers as feedback on their teaching practices;
6. be timely. Feedback has to be provided as soon as possible after the assessment task while it is still fresh in students' memory, whereas assessment conducted earlier than students are able to acquire particular skills may be confusing and discouraging.
7. Effective formative assessment has effective feedback (Black & William, 1998; Higgins, Hartley, & Skelton, 2002). Higgins et al. (2002) found that effective formative feedback has the following characteristics: it is timely, unambiguous, corrective in nature, and fosters high-order critical skills.

In summary, teacher assessment practices should not only inform students' about their learning it should also inform the teacher and the students about how to improve the learning conditions within the course of study (Black and Wilian, 2009).

## **Feminist Assessment**

Evidently, a large body of literature on feminist pedagogy focuses on its definition, curriculum content, and theoretical analyses of its goals and main principles. However, such an important topic as feminist assessment practices is overlooked and, in fact, under-theorized (Accardi, 2013; Akyea & Sandoval, 2005; Clifford, 2002). “Too few feminist scholars discuss how they assess students’ knowledge and products. Without a clear discussion of how feminist scholars assess students, their suggestion to share power with students leads to confusion” (Akyea & Sandoval, 2005, para. 12).

### **The Incongruity between Notions of Feminism and Assessment**

Clifford (2002) proposes that lack of attention to assessment theory and practices in feminist pedagogy “reflects the contradictory nature of the word *assessment* attached to the concept of *feminism*” (p. 109). Accardi (2013) attributes this gap in the literature to possible assumptions that “assessment is somehow anti-feminist because it is counter to the learner-centered, consciousness-raising focus of feminist pedagogy” (p. 75). More than any other aspect of education, assessment embodies a power relationship between the institution and its students, with tutors as custodians of the institution’s rules and practices.

Similarly, Campbell (2002) points out that there is a contradiction between feminist pedagogy and assessment. However, she emphasizes the controversial position of feminist teachers in the assessment process:

Power and authority are complicated issues for the feminist teacher because of contradictory locations: the institutionally sanctioned authority of a teacher’s

position, the disempowerment they can experience, their political struggles to survive in an often hostile domain, and the attempt to even out power hierarchies when they are expected to *grade* the “voice” of learners. (p. 32)

According to Flores (2015), feminist assessment practices—particularly in the universities of the United Kingdom—are limited by the institutional requirements and the standardization that disciplines both instructors and students who are “ranked and valued against pre-established criteria of assessment” (p. 47). Due to institutional constraints, assessment of feminist pedagogues is limited and consists of examinations and essays (Flores, 2015).

As such, some feminist educators merely “ignore a discussion of how they grade and provide learning feedback” (Akyea & Sandoval, 2005, para. 12). Other scholars attempt to overcome this incongruity by focusing on evaluation of programs rather than assessment of learners’ knowledge. They define feminist assessment as “a form of program evaluation in higher education built upon unique principles of assessment that exemplify the feminist ideals of student-centeredness, activism, and the impact of context” (Salkever, 2002, p. 88).

### **Feminist Assessment from the 1980s to the 2010s**

Feminist pedagogues of the eighties and nineties admit that assessment of students’ knowledge, particularly grading, is the main source of the power imbalance between students and teachers (hooks, 1994; Luke, 1996; Schniedewind, 1981; Shackelford, 1992); however, they do not specify how to reduce the power imbalance caused by traditional assessment and what exactly feminist assessment in a democratic classroom should look like. For instance, hooks (1994), while providing an honest analysis of her

teaching experience in academia and criticizing the existing grading system, does not give a detailed description of her own assessment practices in her women studies' classes. In fact, hooks offers a very general idea of her view of the process. According to hooks, assessment has to be more flexible, with high, but not absolute standards, and it has to provide students with a possibility to controlling their grades by their labour.

Similar to hooks (1994), Schniedewind (1981) shares her experience as a women's studies professor. In order to avoid grades, she establishes the complete/incomplete grading for one of her courses, and for those students who still need a grade, she provides a very detailed rubric with descriptions of what work it is necessary for them to complete in order to get a certain grade. At first glance, this approach to assessment and grading seems to be empowering for students because "each student decides the grade she wants to earn and does the appropriate work" (Schniedewind, 1981, p. 27). However, *appropriate work* is a questionable criterion, and it is not clear how Schniedewind determined the "appropriateness" of students' work.

Another proponent of feminist pedagogy, economist Shackelford (1992), proposes to employ feminist methods of teaching by establishing a forum where students can participate as equals and provide their "input into the process, goals, content, course design, and even evaluation" (p. 573). Shackelford considers assessment tools such as peer review of writing, projects, and assignments that "empower students to seek answers beyond the texts and lectures" (p. 573) to be feminist. In her economics classes, evaluation of assignments, group dialogues, and participation are "combined with constant and consistent feedback to establish a learning environment. Students should know where they stand at all times" (Shackelford, 1992, p. 574). Thus, evaluation uses

progressive bottom-up methods of assessment. However, by showing students where they stand, Shackelford ranks students, which makes her assessment problematic from a feminist perspective. Shackelford emphasizes discovery learning in her classes, and she states that in all assignments “students write to learn rather than demonstrate knowledge” (p. 573). However, Shackelford does not explain how these assignments are assessed. Thus, similar to hooks (1994) and other feminist educators, Shackelford does not describe her assessment practices in detail.

In the nineties, there was a dramatic shift in academic assessment from testing to a wide range of educationally meaningful approaches tied to learning approaches, such as portfolios, group interviews, student self-assessment, and the involvement of external examiners (Hutchings, 1992). This shift encouraged feminist pedagogues to overcome the negative associations with the term *assessment* and to start developing feminist assessment to facilitate women’s studies programs (Arnold, 1992; Hutchings, 1992; Musil, 1992).

Currently, feminist assessment emphasizes formative assessment of various learning activities and value practices in which assessment and learning are blended (Accardi, 2013). For example, Flores (2015) observed a feminist class where teaching was blended with formative assessment by teacher and peers. During that class, students shared their plans for their final essays, received feedback both from with the teacher and their peers, and the teacher commented on the work, providing students with some guidance.

## **Features and Principles of Feminist Assessment**

Hutchings (1992) argues that feminist assessment shares similar characteristics with innovative assessment practices, which are educationally meaningful for the students. The only difference between the two is that innovative practices are shaped by practical concerns whereas feminist practices are shaped by feminist theory and values. These shared characteristics make assessment an educational process rather than an administrative task, and they include the following:

1. focus on student experience;
2. focus on using gathered information for the improvement of programs rather than for external proof;
3. use of multiple methods when gathering information in order to obtain the fullest picture of what students have learned;
4. organizing discussions of the assessment results with faculty members, administrators, and students (Hutchings, 1992).

Shapiro (1992) describes nine principles of feminist assessment, which emerged as a result of a three-year research project on the subject. According to Shapiro, feminist assessment:

1. questions all evaluation and assessment paradigms;
2. is learner-centered, and it focuses on learning rather than on displaying adherence to someone else's values;
3. is participatory and interactive;
4. is deeply contextualized and values individualized assessment practices for students with various backgrounds;

5. is decentered by avoiding outside assessors who carry abstract understandings of excellence and by valuing knowledgeable inside assessors (faculty members, administration, and students);
6. complies with feminist activist beliefs, which results in an emphasis on discussions and journal-writing;
7. is shaped by feminist pedagogy, which in turn emphasizes the role of relationships in learning;
8. is a process designed for improvement rather than testing because it takes into consideration such feminist concepts as caring, relatedness, concern, and connectedness;
9. acknowledges values, and it does not presume to be value-free and objective.

The principles developed in the early nineties are still relevant today as they are concerned with power and social justice. Furthermore, they are timely as they are learner-centered. Yet, Akyea and Sandoval (2005) criticize these principles for providing little detail on assessment practices in the classroom and for not addressing student learning concerns. Moreover, these principles were developed exclusively for women's studies programs (Arnold, 1992; Hutchings, 1992; Musil, 1992; Shapiro, 1992). Notably, even the term *feminist assessment* was applied solely to women's studies or gender programs (Hutchings, 1992). However, both feminist theory and feminist pedagogy are continually in the process of development, and since the eighties, it "has moved from the margins, becoming a valued tradition in many academic settings" (Crawley, Lewis, Mayberry, & Editors, 2008, p. 1). Therefore, principles of feminist assessment should be understood in the light of these changes.

Along with the principles of feminist assessment described by Shapiro (1992), Accardi (2013) claims that feminist assessment should be reflexive, and it should use feminist strategies which acknowledge “the uniqueness of each learner and measures learning in a way that validates this uniqueness” (p. 76). These additions seem to be quite significant and should be considered part of feminist assessment.

Akyea and Sandoval (2005) highlight another important characteristic of feminist assessment: Feminist assessment should be effective in assuring student learning and improvement because lack of academic knowledge causes student disempowerment. This suggests that besides being compatible with feminist principles, feminist assessment should have effective assessment tools that facilitate learning.

### **Feminist Assessment and Grading**

Feminist scholars recognize that grading involves the exercise of considerable power and view it as problematic for feminist educators (Accardi, 2013; Kenway & Modra, 1992). However, Kenway and Modra (1992) claim that feminist pedagogy literature is “most unclear on the question of grades” (p. 154), and that grading is under-theorized. Indeed, assessment has an essentially hierarchical nature (Reynolds & Trehan, 2000), and the demands of the institution for standards and examinations conflict with the idea of liberation (Kenway & Modra, 1997).

Feminist scholars suggest various solutions for reducing the power imbalance inherent in the grading process, such as implementing graded portfolios (Arnold, 1992), including self-assessment elements in knowledge tests (Accardi, 2013), and various forms of negotiated assessment (Kenway & Modra, 1992; Schniedewind; 1981) described in the subsection Purposes of Assessment.

Arnold (1992) states that graded portfolios comply with feminist pedagogy. Since portfolios are multidimensional in comparison with single-score assessment, it is therefore less subject to misuse (Arnold, 1992). However, Arnold does not specify the criteria for grading, nor how and by whom grades are generated.

Accardi (2013) claims that she makes rigid and patriarchal methods “more feminist by adding an element of self-assessment and personal reflection to them” (p. 85). Accardi does not explicitly indicate the purposes of her knowledge tests developed for a library instruction course, but she provides a sample of her knowledge tests and includes knowledge tests in her list of feminist assessment methods that feature multiple choice questions tests, paraphrasing, and summaries. Accardi does not specify whether her knowledge tests are formative or summative, whether there is grading involved, nor how she grades students. She does admit, however, that formative assessment is a preferable assessment type in a feminist class.

Yet there are feminist educators who think that summative forms of assessment are needed and by avoiding grading or giving all students high grades, “feminist pedagogy can fall into permissive teaching that unintentionally produces a form of covert discrimination” (Akyea & Sandoval, 2005, para. 1). Akyea and Sandoval (2005) think that this approach to teaching “can be especially devastating for poor, minority, and first-generation college students” (para. 2), and they argue that standardized tests are worthwhile. Moreover, Akyea and Sandoval are concerned that by caring more for learners’ self-esteem than learning, feminist educators discriminate against students who come to lack necessary academic knowledge. Grading thus facilitates effective learning

that empowers students by giving them a chance for success in the world as it exists today (Akyea & Sandoval, 2005).

Some proponents of summative assessment in feminist classes share their ideas regarding how to reconcile feminist pedagogy and grading. For example, Coeyman (1996), a music professor and feminist pedagogue, argues that in a feminist class grading criteria should be mutually agreed with students and when grading, feminist educators should focus on the quality of work, avoid personal favouritism, and be ready to deal with personal emergencies that affect students' work habits.

Koh (2017) shares her experience in teaching and evaluating a graduate class where students signed a Grading Contract that included peer-evaluation. Every student was responsible for writing two anonymous evaluations of a classmate's work (a midterm and a final evaluation) by using up to 300 words in a satisfactory /unsatisfactory descriptive format. An unsatisfactory midterm evaluation was required to have recommendations for improvement that peers collaboratively determined, and Koh (2017) argues that midterm evaluations provide students with an opportunity to improve, and both evaluations allow students to receive the grade for which they contracted. Koh (2017) considers the contract as a very positive experience, more positive than any course she has ever taught, because

the students in this class used these key learning activities to take control of their own learning. The levels of engagement that resulted were very impressive, as well as the level of commitment most students showed both to the class itself as well as the learning and welfare of the other students. (p. 131)

Yet the grading ideas described in this section run counter to one of the principles of the feminist assessment identified by Shapiro (1992), which is taking into consideration such feminist concepts as caring, relatedness, concern, and connectedness, because grading creates a power imbalance between a grader and the one who is graded. Even if it is not a teacher, but rather a student who does the grading, there is the danger of imposing not the instructor's but another student's values and judgments in the act of measuring and ranking. Consequently, the forms of feminist assessment described in the sections below are not assessment tools directed at producing grades.

### **Forms of Feminist Assessment**

Developing assessment methods that comply with feminist pedagogical theory and practice is a challenging task (Arnold, 1992). Feminist scholars advocate journals, participatory assessment, group- and self-assessment, portfolio assessment, and participatory evaluation as feminist methods of assessment (Accardi, 2013; Arnold, 1992; Clifford, 2002; Hutchings, 1992; Soeiro, de Figueiredo, & Ferreira, 2011). In journals, students analyze their reading and material covered in class and then relate this material to their personal experience (Schniedewind, 1981). Clifford (2002) describes the following purposes of journals: the development of self-reflection and self-assessment skills; the reconstruction of new knowledge from existing knowledge; and the promotion of self-awareness, critical thinking, and self-reflections.

Clifford (2002) found that the journal as a learning tool demonstrated adherence to the principle of feminist pedagogy. Clifford, however, expresses some concern regarding the journal as an assessment tool. The position of the teacher who has access to journals is privileged because students are in a vulnerable position when they give their

personal information “to a person who has the power to fail them” (Clifford, 2002, p. 113). Since Clifford did not feel comfortable showing her own journal to her students, she wondered to what degree students’ journals were sincere. Clifford expresses her deep concern regarding the grading of journals as a form of feminist practice.

Along with the portfolio’s learning benefits, such as being more learner-centered by showing students not only how they achieved a particular outcome but the process that led them to it, Arnold (1992) argues that portfolio assessment is a feminist assessment tool. It is participative in nature because portfolios invite discussions between teacher and students and allow for the incorporation of other feminist methods, such as self-assessment.

Accardi (2013) thinks that the reflective research portfolio has great feminist potential because it values reflection and learners’ contributions to knowledge. In addition, it values both process and product, provides opportunities for students to express themselves, and “challenges traditional teacher/student power relations by empowering students to choose how to represent themselves for evaluation” (Accardi, 2013, p. 85).

Both Accardi (2013) and Arnold (1992) give a rather brief overview of portfolio assessment as a feminist method. Moreover, they do not take into consideration practical implementation problems uncovered by a number of studies (Hamp-Lyons & Condon, 1993). For example, Hamp-Lyons and Condon (1993) found that instructors tend to find shortcuts in time-consuming portfolio assessment by not equally attending to the entire portfolio. In addition, instructors overestimate their role in improving their students’ assignments, and they tend to assess the first draft rather the final work because they view

the revised version as their work rather than the learner's (Hamp-Lyons & Condon, 1993). This shortcoming is not compatible with the feminist belief in collaboration and collectiveness emphasized by Shapiro (1992). Therefore, having students produce a portfolio seems to be a promising feminist assessment tool, especially if the portfolio is an electronic one, because it provides students with an opportunity to share with peers easily and to receive feedback.

*Participatory* assessment, in which learners partake in their own assessment and in the assessment of the program in which they are enrolled, is another assessment tool that may be compatible with feminist pedagogy (Accardi, 2013; Soeiro et al., 2011). Student self-assessment may have potential for student empowerment; however, this assumption needs further research (Accardi, 2013).

Soeiro et al. (2011) conducted a study on an innovative strategy of assessment in which students evaluated the pedagogy of every class and participated in their own evaluation. Soeiro et al. (2011) claim that this strategy empowered students because instructors shared their decision-making power, which provided students with an opportunity to self-regulate their own learning. Soeiro et al. (2011) do not describe how this self-evaluation by students was conducted. However, Soeiro et al. (2011) share their concern that students might not be sincere when providing feedback because the teacher would assess them. This concern suggests that summative assessment played a significant role in the study.

Campbell (2002) examined assessment practices in computer-mediated conferences conducted in a variety of online courses, and she emphasizes conflicts

between the increased intimacy of online conversations and the need to assess them.

She found the following assessment practices compatible with feminist pedagogy:

1. Creating focus groups of five to seven students and evaluating their discussions by quality of points per thread rather than by number of postings.
2. Devolving of instructor's power by using students' own evaluations and reflection on their own postings as criteria.
3. Student involvement in designing questions for written assessment.
4. Use of self-determined formats such as presentations, research projects, multimedia group projects, or creation of an assessment rubric.

The assessment practices described by Campbell (2002) seem to be innovative and democratic since they encourage the construction of knowledge through collaboration. Not all of these methods, however, can be considered feminist. For example, evaluation of group discussion is done by the instructors, who determine the criteria for quality of participation. Yet even equipped with detailed criteria for evaluation, feminist instructors will still experience problems reconciling their values with university culture that quantifies engagement (Campbell, 2002). In addition, it is not clear how exactly student reflections are used and how they are judged or whether students have an opportunity to improve while being assessed. Thus, even those practices which seem to be promising for feminist assessment theory require a more detailed account in order to decide how compatible these practices are with the principles of feminist pedagogy.

## Conclusion

This chapter reviews literature on the origin and main ideas of feminist pedagogy, assessment in higher education, and how feminist pedagogy could inform assessment practices. Feminist pedagogy is grounded in critical and feminist theories. There is no one way of feminist thinking or one view on feminist pedagogy. However, a substantial body of literature on feminist pedagogy describes it as a pedagogical approach seeking to transform learners by fostering an egalitarian classroom environment. Despite the lack of common definition, there are common themes or principles that feminist educators agree upon, namely: creating non-hierarchical co-operative participatory classroom communities where individual voices and experiences are respected, empowering, reformation of the teacher-student relationship, challenging traditional pedagogical notions, supporting women leadership, and caring.

Generally speaking, there are three main kinds of assessment: summative, formative, and diagnostic assessment; however, the distinctions between them are not always clear. Some scholars claim that despite contributing to the final grade, summative assessment, for example, continuous assessment, can be used in a formative way. From a feminist perspective, summative assessment is flawed because it usually involves grading, which ranks students and creates competition. In contrast to summative assessment, formative assessment methods and instruction in which teaching and assessing are blended align with feminist principles. Though I disagree with the idea of summative assessment as commonly defined in the literature having a formative role, others have argued that summative assessment that is used in a formative way, that is to assist student and teacher decisions in regards to future learning activities could be seen as part of

feminist theory to assessment practices (Black & Wiliam, 2009, 2018; Fu, Hopper, & Sanford, 2018). Moreover, studies show that descriptive feedback without a grade is more effective than feedback accompanied by a grade, which may be interpreted as evidence for the usefulness of descriptive feedback and formative assessment as well as demonstrating a negative effect of grading on learning. However, some scholars believe that grading is necessary to prepare students for the existing patriarchal and competitive system, and that is not grading per se that contradicts the main principles of feminist pedagogy, but the way in which grading is used.

Analysis of the current literature indicates that there is a considerable gap in the theory of feminist assessment practices in higher education. Moreover, even this extremely sparse literature on feminist assessment is devoted to assessment practices in women's studies programs, whereas feminist assessment in other disciplines is neglected. Hence, it is critical for feminist pedagogy theory to develop feminist assessment strategies that will be compatible with the main principles of feminist pedagogy.

## CHAPTER: 3 METHODOLOGY

### The Rationale for the Methodology

The nature of my research questions, which required data collection and in-depth analyses of data through human interactions and interpretations, led me to choose a qualitative study. An interpretive study seemed to be a good fit for my research because it “acknowledges some degree of subjectivity and seeks to throw light on a particular situation” (Atkins & Wallace, 2012). Moreover, a qualitative study presents an opportunity to develop a *thick description* of a phenomenon I wish to study. In the literature on qualitative research methodology, the term *thick* describes the findings of research that features accurate description within a well-described context, captures the thoughts and motivations of the participants, and promotes *thick* findings that catch the reader's attention (Ponterotto, 2006).

I chose approaches widely used in qualitative research: a case study and since I consider my own pedagogical practice valid and worthwhile, I included myself as a participant in my research. This added an autoethnographic element to my approach (Ellis, Adams, & Bochner, 2011).

I conducted this study informed by underlying feminist assumptions defined in Chapter 1. These assumptions include the five principles of feminist research methodology identified by Cook and Fonow (1986): attention to gender issues, emphasis on personal experience, a challenge to the norms of *objectivity*, the importance of consciousness-raising, and the transformation of patriarchy.

One of the concerns of feminist research is to study “the ways in which gendered relations of power-knowledge in educational institutions affect women’s and men’s experiences” (Ropers-Huilman & Winters, 2011, p. 677). Studies have shown that subtle bias in higher education is pervasive. For example, female teachers must practise more time-consuming forms of pedagogy than men so that female teacher can get the same ratings as male teachers (MacNell, Driscoll, & Hunt, 2014; Sprague & Massoni, 2005). Similarly, science faculty tend to exhibit gender bias against female students by judging them “to be less competent and less worthy of being hired than an identical male student” (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012, p. 16477).

I addressed the way in which gender affects our perception of people and our behaviour by including a question on gender bias in assessment practices. As a feminist researcher, I treated the participants and myself as gendered beings, keeping in mind that the interview responses of female participants tend to be shaped by patriarchal power structures in our society (Devault, 1990). While working with the transcripts, I was reflexive and reflective about how participants’ gender influenced their responses. By *reflective* I mean that I reflected or thought over the interviews, trying to make sense of the information the participants shared. Being *reflexive*, in contrast, involves the self-awareness of researchers who “turn a critical gaze towards themselves” (Finlay & Gough, 2003, p. 3).

As a feminist researcher, I emphasized personal experience and reflexivity as a paramount feature of ethical practice. While I was writing the field notes, I explicitly recognized that my political position and pedagogical views influenced the research process and the findings. Therefore, I strived to monitor the impact of my biases and to

acknowledge the role of my personal experience and beliefs at the very beginning of my research. By choosing a qualitative study, I also challenged the norms of *objectivity*. I believe that the discussions of assessment practices have produced changes in the consciousness of the participants and myself by providing an opportunity to contemplate existing assessment practices. I therefore consider the study to be feminist for its potential to transform patriarchal forms of assessment into more progressive and liberating assessment practices.

### **The Rationale for a Case Study Approach**

The *case study* approach may be described as “an examination of a specific phenomenon such as a program, an event, a person, a process, an instruction or a social group” (Merriam, 1988, p. 35). A case study allows one to gain knowledge from single cases (Stake, 1994). Since my research questions involved people’s values and attitudes, I followed the recommendation of Atkins and Wales (2012), who claim that these topics are appropriate for a case study approach. Also, a case study approach is considered “particularly useful to employ when there is a need to obtain an in-depth appreciation of an issue, event or phenomenon of interest, in its natural real-life context” (Crowe, Cresswell, Robertson, Huby, Avery, & Sheikh, 2011, p. 1). Since I intended to obtain an in-depth examination of a particular phenomenon (assessment) in its real-life context (the university), a case study approach seemed to be the right choice.

Also, the case study approach is attractive because investigating cases “provides insight into an issue or refinement to a theory” (Stake, 1994, p. 237). Traditional forms of assessment focused on measuring students’ memorization abilities are grounded in the hegemony of patriarchal, hierarchical assumptions because traditional assessment does

not take into consideration and question power, social context, and politics. By providing some alternatives to prevailing assessment practices, I hope that my study will give insights into feminist assessment theory. When I selected the case study approach for this study, I took into consideration the argument of Merriam (1988) that the end product of a case study should be “a combination of description and interpretation, or description and evaluation” (Merriam, 1988, p. 35). The end product of this study, then, is an insight into feminist assessment theory, communicated through a description and interpretation of assessment strategies in one particular university department. This case study would offer a view of assessment strategies used in a higher education second language program that may resonate with instructors in other higher education and language learning programs.

Finally, as a flexible approach, case study offers a great variety of means to conduct small-scale research (Atkins & Wallace, 2012; Hamilton & Corbett-Whittier, 2012; Merriam, 1988). Interviews, focus group discussions, surveys, observations, and artifact and document analysis are among the most popular methods of data collection (Atkins & Wallace, 2012; Hamilton & Corbett-Whittier, 2012). Out of this wide range of data collection methods, I chose reflective journal, individual interviews, observations, and document analysis for my study. Keeping a reflective journal and carrying out individual interviews allowed me to use different perspectives in my study while the data collected from observations and document analysis added a further qualitative dimension of my research.

## **The Rationale for Autoethnography**

“*Autoethnography* is an approach to research and writing that seeks to describe and systematically analyze (graphy) personal experience (auto) in order to understand cultural experience (ethno)” (Ellis, Adams, & Bochner, 2011, p. 273). In autoethnography (“a narrative of self”), the researcher and the researched are one and the same, and this approach demands rigorous reflexivity from the researcher (Chan, 2008; Grant, Short, & Turner, 2013; Sparkes & Smith, 2016). Similar to qualitative case studies, “autoethnography is one of the approaches that acknowledges and accommodates subjectivity, emotionality, and the researcher's influence on research, rather than hiding from these matters or assuming they don't exist” (Ellis et al., 2011, p. 273).

I employed autoethnographic approaches in addition to the case study because I found the features of the autoethnographic approach mentioned above to be beneficial for my research. By utilizing autoethnographic approaches, I “chose to make my experiences, opinions, thoughts, and feelings visible and an acknowledged part of the research process” (Ortlipp, 2008, p. 695). Also, this approach agrees with my political views as a radical feminist as it treats “research as a political, socially-just and socially-conscious act” (Ellis et al., 2011, p. 273). Being a highly reflective method, autoethnography enabled me to address various ethical dilemmas (Ellis et al., 2011) and provided some valuable insights into the research process. Thus, I became both a participant and a researcher, which allowed me to use both the perspectives of my participants and my own in this research. Moreover, presenting the data collected in the autoethnographic tradition provided excellent opportunities to share my thoughts, to make my voice more distinct and to increase transparency in my study.

## Methods of the Research

I used four data collection tools in this study: a reflective journal of my teaching, individual interviews with my colleagues, observations made in classes taught by my participants, and document analysis of the course syllabi of my participants. The main methods of my research were the interviews and the reflective journal that enabled me to use my own experiences and those of my colleagues as resources for my research. Observations made while observing the classes of my participants and document analyses of their syllabi were useful in suggesting additional interview questions and provided supplementary research data. Utilizing more than one data collection method, known as *triangulation*, allowed me to collect richer and more comprehensive data (Atkins & Wallace, 2012; Hamilton & Corbett-Whittier, 2012). Also, employing various tools allowed me to view assessment from more than one standpoint (Cohen, Manion, & Morrison, 2002).

The reflective journal consists of two parts: a teacher journal in which I analyzed my pedagogical practices as a Russian language instructor, and a research journal in which I recorded my thoughts and events occurring while I was progressing with the interviews and data analyses. Writing my teacher journal enabled me to generate materials derived from my experience and my subjective emotions. I believe I benefited professionally from writing the journal as an educator and as a researcher. Keeping the research journal helped me understand the methods and content of my research, and it raised my awareness of my own thoughts, feelings, and biases (Etherington, 2004; Orange, 2016). As I was progressing through my study, my praxis shifted towards liberating assessment while my research became more flexible and open-minded. This

shift occurred due to constant reflections on my beliefs and pedagogy, and it illustrates the idea that “as important as methods may be, the most practical thing we can achieve in any kind of work is insight into what is happening inside us as we do it” (Palmer, 2007, p. 6).

While I was writing my journal, I was keeping the research questions in mind, trying to make my notes relevant to assessment practices and my attitude towards them. I was engaged in systematic sociological introspection, which involved a process of emotional recall (Ellis, 1991, 2008). Before making entries in the journal, I recalled my working day starting from the first morning class, and while visualizing it, I examined what I felt and why. Experiencing tension while recollecting class activities signified that there was a discontent between my teaching values and my pedagogy, one that I needed to address this problem in my analysis. I then made entries in my teacher journal identifying and summarizing the most important ideas related to my teaching and assessment practices. Thus, I was using introspection as an interpretive tool, and my thoughts and feelings as a source of knowledge.

By recording my conflicting and ambiguous thoughts and emotions, I achieved a deeper level of understanding of assessment practices. To make my study robust, however, I needed more perspectives than just my own. I therefore employed other methods of conducting research, such as interviews, observations and document analysis. I used classroom observation as a robust research method, which is valued for its authenticity and opportunities to discover issues that might be missed in interviews (Cohen et al., 2002). I was particularly interested in attending large classes since they seemed to be the most challenging for both feminist pedagogy and assessment.

Document analysis or systematic procedures for reviewing or evaluating printed or electronic documents are often used in combination with other research methods (Bowen, 2009). In this study, I analyzed course syllabi because they are a rich source of data. A course syllabus is worth studying because it is a powerful document, one that functions as a contract between the professor and students, a plan of action, a communication device in which an instructor answers anticipated questions, a plan of action, and a cognitive map of the class environment (Matejka & Kurke, 1994; Thompson, 2007). A syllabus represents the instructor's teaching philosophy and the culture of teaching and learning in the educational setting (Stanny, Gonzalez, & McGowan, 2015).

Four participants offered their course syllabi, which represent the corpus of documents that I analyzed. By analyzing the content of the syllabi, I gained insight into the structure of each course, the instructor's teaching philosophy, and the assessment and evaluation methods used in the class, or at least the instructor's intent to use specific assessment methods (Cullen & Harris, 2009).

The reflective journal, class observations, and reading the analysis of the course syllabi prepared me for the next step in my research: individual interviews. Indeed, the examination of my own assessment practices and observations of the assessment strategies of other teachers helped me to formulate the interview questions I would ask the participants. These questions centered on their perspectives on assessment and their experience and expertise in various assessment methods.

According to Atkins and Wallace (2012), the interview is a flexible research method designed to gather various types of information, including factual data, personal

beliefs, and possibly narrative anecdotes. The justification for the use of the interview as a research method came from researchers who considered an interview to be an effective way of gathering information on research questions and who claim that interviews offer the richest data for research (Hamilton & Corbett-Whittier, 2012; Magnusson & Marecek, 2015).

I chose an individual, semi-structured interview for my research, a method of inquiry that uses a prepared set of flexible questions (Hamilton & Corbett-Whittier, 2012). Using a semi-structured type of interview gave me some control over the discussions, while open questions allowed the participants a considerable amount of freedom in how they responded to the questions. In the informal semi-structured interviews, the participants could contribute their own agenda on top of the prepared list of questions, taking the interview “beyond the gathering of fact and allowing the participant an authentic voice” (Atkins & Wallace, 2012, p. 88). I expected that my research would benefit from the semi-structured interview method because of its active involvement of participants in the meaning-making process and organization of the research. I expected that interviews would be collaborative and informative because the participants and I were co-workers and we had often shared their perspectives on assessment before I started my research.

Interestingly, while I did not ask questions designed to elicit narratives from the participants, three participants nevertheless told stories during their interviews. Their narratives were short accounts of their university experiences either as a student or as an instructor. In these stories, the participants shared their insights about university policies

related to evaluation practices and their perception of students' attitudes towards the assessment and evaluation of learning.

## **Procedure**

### **Participants**

The participants were instructors from a department of modern languages and literatures at a large Canadian research university. In addition to myself (a researcher participant), there were six participants in the study, which I considered to be a sufficient number for the following reasons:

1. The number appeared manageable for conducting an in-depth interview with each participant.
2. Since a semi-structured interview is a rich source of data, a study that employs interviews needs fewer participants, and a sample of six participants was sufficient for qualitative analysis (Smith & Sparkes, 2016).
3. Since I participated in this study as the researcher and as a participant, the study collected data from seven participants, which appeared an appropriate number for identifying meaningful patterns while conducting analysis (Brown, Clarke, & Weate, 2016).
4. Seven is a sufficient number for providing some diversity of opinion. Indeed, there were five female and two male participants, aged 40 to 65, who belonged to various cultures: North American, Eastern- and Western-European, and Middle-Eastern. Four participants were full time contract faculty from the teaching stream, and three participants were tenure faculty from the research stream. I

consider myself to be a feminist researcher; therefore, gender and culture diversity of the participants is particularly important because it challenges androcentric research based on the experiences of white, middle-class, predominantly Western men.

The choice of participants for the interviews was justified by my familiarity with them, because in a case study, the researcher should be familiar with the research setting (Hamilton & Corbett-Whittier, 2012). By the time I started collecting data for the research, I had been working in close collaboration with my colleagues for over eight years, and I considered myself to be familiar with the department.

Since I had non-hierarchical collaborative relationships with the participants, and they volunteered to participate in the study, I believe they had an authentic interest in my research. To protect the confidentiality of the data and the participants, I assigned them pseudonyms that I used throughout the research process, and I do not use names in this dissertation.

## **Gathering Data**

### **The reflective journal.**

On July 13, 2016, I obtained ethics approval from the University Research Ethics Board, and I followed its guidelines at each stage of the study. The data collection took place from the end of August 2016 to the beginning of May 2018. On August 19, 2016, I started to write a reflective journal, which I kept for over a year and a half. The journal consists of two parts: the teacher journal and the research journal. The teacher journal is

62 double-spaced pages long, and it covers the period from August 19, 2016, to December 18, 2016.

The teacher journal includes autoethnographic data: my field notes on my pedagogical practices, my thoughts on past events, occasional emails from students and colleagues, and still images. I created entries approximately three to four times per week, and the focus of the notes was the exploration of my pedagogical practices from a feminist perspective. At the beginning of January 2017, I started to write a research journal, in which I reflected on my observations and on the interviews with the participants. The focus of the entries switched from reflections on my teaching to the pedagogical practices of others and to researching ethical issues. Also, in contrast to the teacher journal, the entries in the research journal became shorter, less frequent, and sporadic.

### **Recruitment.**

Examination of assessment practices of other pedagogues and their attitudes towards various types of assessment was the next stage of my research. While choosing between several ways of recruiting participants, I dismissed recruitment through face-to-face contact to avoid pressuring faculty members into participation. Therefore, I chose email as a non-intrusive and low-pressure way of recruitment in which potential participants had time to consider whether they had sufficient interest and/or spare time for an interview.

On January 3, 2017, I sent an email to all faculty members. In the email, I told them about the topic of my research, described the methods I was planning to use, and invited colleagues to participate in my study (See “Letter of Initial Contact” in Appendix

A). Six faculty members agreed to participate in the study. To minimize potential risks to the confidentiality of participants, I used an informed consent procedure (See “Informed Consent Form” in Appendix B and “Obtaining Consent Email” in Appendix C).

Two participants expressed a particularly keen interest in my study and invited me to attend their classes. I was interested in observing their classes for a few reasons:

1. They taught courses with over 80 students, and I was particularly interested in the challenges of implementing assessment that was timely, unambiguous, and corrective in nature.
2. I was familiar with the pedagogical views they had shared at departmental meetings and colloquia, where their teaching philosophies seemed to be aligned with feminist pedagogy.
3. The effectiveness of their teaching was recognized both by the administration and by students in their evaluations.

Thus, in addition to the interviews, I had an opportunity to use classroom observation as a research method. Since four participants provided me with the syllabi of their courses, I used the document analysis as a complementary research method to achieve triangulation. I examined the core elements of the syllabi required by the university: course learning outcome, course overview, evaluation and grade distribution, and course policy. In addition to the information on how the participants evaluated students' work and the required assignments and examinations, I examined attendance policy (whether it is part of the grade), the description of specific assignments and class activities in relation to feedback (multiple drafts of papers, peer review as a class

activity), and course policies (acceptance of late works). For some pedagogues, attendance is an indicator of students' effort; for others it is a possibility to evaluate students' thinking in class discussions or their foreign language abilities in group work. Multiple drafts indicate the willingness of teachers to provide formative feedback while acceptance of late works is an important indicator of a teacher's philosophical stance on pedagogy or their capacity to care about students.

### **Classroom Observations.**

I conducted observations in two third-year literature classes with the focus on assessment practices. According to previous research, three observations is a sufficient number to record accurately the classroom environment (Shih, 2013). I conducted 15 classroom observations in the span of one term. I expected that observations at different times of the semester would allow collecting richer data by capturing the assessment practices of participants when they taught a variety of topics. Also, I predicted that at different periods of the teaching semester (the beginning, the middle, and the end of the term) the participants would use different types of assessment.

I participated as a student, although students were informed that I was a faculty member. I became immersed in the course, and I did the assigned reading for each class and participated in the classroom activities as if I were a student. This strategy permitted an in-depth collection of information on assessment practices of the participants and how these practices affected students. I took notes on my observations either after classes or while the participants lectured. This strategy enabled me in some way to put myself in the students' shoes because I had to prepare for classes so that I could participate in group

discussions. However, the immersion method had one disadvantage: it was time-consuming because of the extensive reading assignments for these classes.

### **Interviews.**

Autoethnography, analysis of the syllabi, and classroom observations allowed me to get a better understanding of the topic of assessment and to develop meaningful open-ended questions (See “Interview Questions” in Appendix D). Since I wanted to avoid imposing my view of “feminist assessment” on the participants and could not be sure that my participants and I understood “feminist assessment” in the same way, my questions were aimed at investigating the assessment practices of my colleagues in general, without explicit reference to feminist pedagogy. For example, I used such questions as, “Is there an assessment method in your teaching practices that you find particularly effective and interesting?” or “Do you use diagnostic assessment in your practice? If yes, please tell me how you use it.”

The interviews were scheduled at the participants’ convenience and held from February 28th, 2017 until May 1st, 2017 (See “Interview’s Schedule” in Appendix E). Lasting from 15 to 56 minutes (201 minutes total), I audio recorded the interviews with the permission of the participants. I conducted interviews face to face with feminist interview methods in mind. Interview methods can be considered feminist if “the relationship of interviewer and interviewee is nonhierarchical and when the interviewer is prepared to invest his or her own personal identity in the relationship” (Oakley, 1981, p. 41). I emailed the same set of interview questions to all participants a few days prior to each interview. I added some questions related to the challenges involved in assessing large classes and online classes for the participants who taught these courses (Rideout,

2018; Soffer, Kahan, & Livne, 2017). For instance, in addition to the 13 questions, I sent the participants a question on assessments of large cohorts: “According to research, providing detailed and timely feedback in large classes can be challenging. How do you assess learning in your class?”

All interviewees allowed the use of a recording device. I therefore used an audio recorder and later transcribed all the interviews in full. Within a week of each interview, I sent the transcript to the participant via email to ensure its accuracy or, in other words, to ensure that the interviews validly represented what the participants said or meant to say. In my email, I asked each participant to provide corrections and clarifications, if necessary. In their responses, the participants indicated that the transcripts were accurate, and they had no revisions to make.

### **Summary**

In this chapter, I describe the approaches and methods of my research, the rationale for using them, and a detailed procedure of how I used these methods. I employ two approaches in this study: case study and autoethnography. My choice is based on the fact that these are flexible approaches that allow an in-depth, qualitative exploration of assessment in higher education. Also, to gather data that yield various perspectives on assessment, I use four data-collection tools with writing a journal and conducting interviews as primary tools, and observation and document analysis as complementary tools that are particularly helpful in the preparatory phase for the interview.

The chapter contains a detailed description of all elements of the research development process, as follows: receiving ethics approval from the university where the study was conducted, writing the teacher journal, recruiting the participants, collecting

observations, exploring the participants' course syllabus, preparing the interview questions, interviewing the participants, transcribing the interviews, and finally seeking the participants' approval of the transcripts. All these stages of the study are grounded in feminist values and in the following research principles: attention to gender issues, emphasis on personal experience, challenging the norms of "objectivity" by seeing the world as being comprised of multiple truths and different perspectives, the importance of consciousness-raising, and of acting so as to transform the patriarchy.

## CHAPTER 4: THE CATEGORIES

### Developing a Structured Category Process

In this chapter, I present the categories developed from the four data sources I designed to evoke responses to my research questions:

1. What forms of assessment do instructors in the Department of Modern Languages and Literatures use that comply with principles of critical feminist pedagogy?
2. What tensions exist for instructors related to assessment in higher education?
3. How can these tensions be addressed using a feminist reconceptualization of assessment?

While reading the findings of this work, it is important to consider the context of the study—a context that influenced the data and the findings. The topic of assessment has always been a subject of intense discussion in our department. When I was conducting my research, however, the administration started a campaign against “grade inflation,” and the topic of assessment and grading became a major concern for our faculty.

In qualitative research, recording study results and analyzing data are interconnected (Chang, 2008; Hennink, 2014). Describing and organizing the data and developing categories is the first step in the process of examining the data to answer the study’s research questions. Therefore, organizing and labeling data was the first step in the analysis. I also included some preliminary analysis in my findings chapters.

I chose a cluster analytic approach, namely, thematic analysis, because it is a robust and flexible method that “offers the potential for nuanced, complex, interpretive

analysis” (Brown et al., 2016, p. 191). According to Brown et al. (2016), this approach can “provide analyses of people experience in relation to an issue” (p. 193). Since the nature of my research questions related to the experience of the participants (people experience) in assessment (issue), thematic analysis suited the research questions of the study. Finally, Brown and al. (2016) recommend thematic analysis for analyzing semi-structured interviews and diaries or journals, which are the data types I used in my study.

I conducted the data analysis in multiple phases. First, I looked for patterns and relationships within each data collection in the order the data was obtained, namely, my teacher journal, the syllabi, field notes from the observations, and the interviews. Then I searched for relationships across each data collection that allowed me to interpret the data.

I analyzed the primary data sources, the interview scripts, the teacher journal, and the observation notes using the qualitative data software program NVivo 11. I approached the data coding by using inductive or “data-driven” analysis. According to Smith and Sparkes (2016), this analysis is driven by the data content rather than by theory. Coding refers to assigning labels to the fragments of data that describe what each segment is about (Brown et al., 2016). I performed a data reduction process, during which I coded the collected data by labeling frequently patterns of data that established categories and later implied themes (Hamilton & Corbett-Whittier, 2012; Hennik, 2014). Categories refer to “analytic units that conceptually match the phenomenon portrayed in the data set” (Given, 2008, p. 71) Themes refer to “any principle recurrent in a number of domains, tacit or explicit, and serving as a relationship among subsystems of cultural

meaning” and considered as “cognitive principles that people believe and accept as true and valid” (Spradley, 1980, p. 141).

While I proceeded to the explorative phase of the data sources, I used Brown et al.’s (2016) guidance for analysis, which included reading through the entire documents for an overall sense of the data, identifying “meaning units,” and organizing these units by determining whether they fit together. I thus read the data multiple times and coded the content into main categories and subcategories. At this stage of the data analysis, I looked for the patterns in the data. I explored the data, identified clusters that emerged, and created nodes.

At a later stage in the analysis, I decided to eliminate the data that did not address the research questions. The rationale for this decision was my expectation that since the data collection methods were intended to answer the research questions, the data should be relevant to the topic of the research, even though this relevance may not be necessarily obvious. For example, assessment is the focus of my journal but the category *Pressures, Tension and Discontent* includes data on pedagogical challenges I face in my teaching in general. At first sight, these challenges did not seem to be directly related to assessment. Yet after a closer examination, I realized that all aspects of pedagogy (content selection, delivery, and assessment) were interconnected, and any change in one of the components affected the other elements. I found a similar idea in my teacher journal entry dated November 11<sup>th</sup>, 2017: “I am amazed by the importance that assessment plays in the Russian language curriculum, and how every change in assessment changes my teaching style, my relationship with students, and the students' learning strategies.”

### **The Categories in relation to the Research Questions: Findings**

While I was exploring the relationships between the categories, I used Mind Map, a software brainstorming tool offered by NVivo. I visually organized all the categories as unconnected floating ideas, and then examined how the categories might relate and overlap. I then reread the nodes and created new refined categories, which were connected on the same level or in a hierarchy. At the next stage of the research, I refined my analysis by commenting on the importance and implications of the categories I had established in the previous stage and by developing finalized themes. Smith and Sparkes (2016) suggest that at this stage researchers move from descriptive to interpretive commentary and answer their research questions.

While creating categories that describe and summarize the information obtained from the participants in the interviews, the main challenges were: 1) detaching mentally from the interview questions I asked the participants, and 2) creating names that captured the essence of each category. My first attempt ended with categories generated according to the interview questions. This was one of the most frustrating stages of the data analysis because the nodes had no structure, and all categories seemed unrelated. I overcame this problem by ignoring the interview questions. This approach helped me mentally to detach from the questions and to focus on what the participants actually said rather than what I expected them to say in response to my questions. It also allowed me to include information that might not obviously relate to the research questions. During this stage of the analysis, I tried to avoid formulating an interpretation of the information that I was organizing.

My analysis of the course syllabi of my participants involved multiple readings of the documents and coding deductively in relation to the research questions and the clusters that had already emerged. To do this, I wrote notes on the sections related to assessment, evaluation, grade allocation, and feedback from the instructors. The syllabi contained information about the requirements of each course and evaluation criteria to be applied, which I found to be useful additional sources of data and which I included in suitable categories.

In the interviews and the journal, the data clustered around two distinct overarching categories: *Participants' Reported Assessment Practices* and *Pressures, Tension and Discontent*. These overarching categories capture the main idea of the next layer of categories, mapped in Figure 1 (See also Appendix F)



**Figure 1** Assessment through a Feminist lens

The category *Participants' Reported Assessment Practices* provides information on forms and features of assessment that may comply with principles of feminist pedagogy referenced earlier in this dissertation. The category *Pressures, Tension and*

*Discontent* responds to the second research question related to tensions for instructors associated with assessment. The participants reported that assigning students grades was both a source of frustration for them and part of their practices that motivated students to learn and improve. I therefore included participants' comments on grading in both overarching categories. Initially, Grading was a subcategory that belonged to the category *Pressure from Administration*, and this category still overlaps with the sub-category *Grading Policy*. However, unlike pressure from the administration to grade on a curve, educational grading is almost universal in higher education. It is therefore not the administration that puts pressure on the faculty, but society in large. Moreover, during the analysis it became evident for me that the rich data collected on the participants' attitudes to grading was too diverse and needed to be broken down into a separate category.

Some categories more informed by data than others. I assume that a high number of references related to particular assessment strategies in the category *Participants' Reported Assessment Practices* indicates the importance of these assessment approaches for the participants. In the chapter below devoted to the *Participants' Reported Assessment Practices*, I therefore described the categories in the order from the highest to the lowest number of references. However, I disregarded the number of references on grading and participants' feelings of pressure and discontent because there the scarcity of references in a category may suggest the unwillingness of participants to talk about this topic. These chapters, therefore, describe the categories according to their similarities and differences.

In my descriptions of the clusters, I presented abundant excerpts from the interview transcripts along with my commentary. I believed that showing the

participants' exact words to be important not only because quotes provide evidence for my interpretation, but also because such illustrations deepen the reader's understanding of the topic by giving additional nuance to the emotions the participants express. I believe that taking in the participants' own words will give readers an opportunity to draw their own conclusions about the trustworthiness of my interpretations. In addition, for me as a feminist researcher, it is important to create spaces in the study to share insights from my participants using their words, albeit selected by me, because I view them as co-authors and the study to some degree as our mutual production. Thus, by citing my participants' exact words, I give the reader an opportunity to understand how they expressed their opinion and feelings about assessment, which is a topic of great concern for them.

Since writing qualitative research is an ongoing process that influences a researcher's beliefs, in the chapter entitled "My journey to become a feminist educator," I often use citations from my teacher journal rather than summarizing the content of the entries. By quoting myself, I emphasize the idea that in this research, I am *also* a participant who wrote that information in 2016 rather than *only* a researcher who makes her statement on assessment two years later.

### **Summary**

In qualitative research, the analysis of data and description of findings are interconnected. I have therefore included analysis and findings in the same chapters. I have chosen thematic analysis as being a robust and flexible method that researchers often use to analyze semi-structured interviews and journals. One of the main challenges during my analysis was detaching from the interview questions to capture the essences of participants' utterances. This allowed me to create the categories according to what the

participants said rather than to what I expected to hear. As a result of the analysis, I identified two distinct overarching categories: *Participants Assessment Practices and Pressures, Tension and Discontent*.

These two overarching categories create layers or clusters that elaborate on the meaning of their corresponding overarching categories. I use different approaches to the sequence I present the data. For example, in the next chapter, the clusters with the highest number of references are considered more important, and they are described before the categories with fewer references in the data. However, in the other chapters, the frequency with which a coded category is found in the data is not necessarily proportional to its significance, and in those chapters I present the categories based on their relevance to each other.

In this chapter, I also explain the abundance of quotations from the interviews and my teacher journal. I use the participants' exact words as evidence for my interpretations and as a means of deepening the readers' understanding of the topic so that they could make their own conclusions. I also prefer direct citations from my teacher journal to descriptions of the journal entries to emphasize the idea that in this research I am also a participant sharing my personal experience rather than merely a researcher making her statement about assessment.

## CHAPTER 5: PARTICIPANTS' REPORTED ASSESSMENT PRACTICES

### The Overview

The cluster *Participants' Reported Assessment Practices* has 111 coding references, which is the highest number in the study. The participants indicated the following forms of assessment as being among the most effective for their students' learning assessment practices: summative, formative, pass/fail, peer- and self-assessment, and diagnostic assessment. The participants also talked about essential qualities that assessment practices should have that apply to all assessment forms. I coded these features to the category *Essential Aspects of Assessment*.

### Summative Assessment

#### Assessment Modifications

All participants considered summative assessment to be effective for learning. Five out of seven participants, however, used modified forms of summative assessment. For example, two participants allowed students to rewrite their work. One of them said: "Even after giving students their grades, I allow them to upgrade. To me, it is important that they learn, and not just to grade them. This is a somehow diagnostic assessment."

The other instructor also said that he consistently encouraged students to revise their work "on the assumption that they get better." All participants said that they used summative assessment in combination with extensive written feedback, and two participants indicated that they allowed students to send the drafts of their work for their feedback before submitting the final work for a grade.

Three participants shared another distinctive feature of summative assessment they used in their classes: 4. One instructor gave students six questions ten days before the final examination. She let students know that they would be given either two topics and be asked to write on one of them, or three topics and be asked to write on two of them. At first blush, it does not seem helpful to students who would be preparing for double the number of topics they will need for the test. This instructor's intention was to use this practice to serve as a motivator for students to prepare for the exam while making their study manageable, especially in comparison to instructors requiring knowledge of all material.

Another instructor gave students several essay topics in advance, and students had an opportunity to choose the ones that they would like to write on at their in-class examinations. Finally, the third instructor let students know beforehand one out of two questions:

It is a legitimate thing. Once they know the question in advance, of course, you correct it with the knowledge. For example, when I sometimes teach a generic literature course, I say, "Look, you've got two essay questions. One I will certainly tell you, but the second question is a fairly free essay question. And it might be a very simple question, like, "What was the best text in class?"

### **Participation**

One participant excluded the evaluation of participation entirely when assigning final grades, considering the assessment of students' observable involvement in the subject to be problematic and subjective. Five participants did include the class participation mark as part of their overall assessment criteria but assessed and evaluated

student participation differently. One participant said that since she did not teach large classes, she had an opportunity to take notes on how students participated after every literature and language class. She discussed the participation grade at the very beginning of the course and emphasized that their mark depended on how actively students were involved rather than on the correctness of their answers. If she suspected that there were shy students in group discussions, she assigned those students to different groups.

One professor who taught large classes assigned from 10% to 15% of the final grade to participation. For example, in the syllabus of his course, participation was worth 10%, and it was stated that participation would be marked “according to your attendance and contributions to class. This includes visits to office hours with the instructor and online feedback.” He said that students often did not get high participation marks, estimating that on average students received six out of ten marks. In his evaluation of participation, he took into consideration what students said, how often they spoke, “the level of what they said,” how often they came to his office hours, how often they wrote emails, and whether they raised any additional questions. The professor liked to be involved in conversations with students in class and outside of class: “I like what they say in class. I like to read their writing. Anything they can send me outside of class like jokes, or references... I encourage that.”

Interestingly, many faculty members in the department used the term “soft mark” when they referred to participation grades, but at first I did not know its exact meaning of this term. My initial search on Google brought back the expression *soft grading*, which means “easy grading.” One of our faculty members told me that in her mind “soft grading” meant allowing “some flexibility for teachers,” and she mentioned a

participation mark as an example. From another faculty member's point of view the term "soft mark" could refer to both "the easy grades some teachers give and the grades they give for participation." He considered the difference between "hard" and "soft" tricky, because for some shy students, participation is very difficult and therefore cannot be considered "soft."

In the language classes of one of the instructors, class participation was worth 5 % of the final grade, whereas in her literature classes she did not include participation in the final grade. She did, however, reserve the right to add up to 3% for students who were with her "throughout the entire semester," mostly by answering and asking questions and "thinking" with her: "There is no assigned grade for participation but students who are very active in my class and whom I personally enjoy having in my class ... I give them extra few marks at the end." In her literature classes syllabi, she explicitly states her right to increase the mark: "There is no participation mark, but the instructor reserves the right to honour above-average participation with an overall mark increase." Thus, the instructor does not include "the soft grade" in her evaluation, but she can raise the final grade in cases of superior in-class participation.

Another educator used technology for evaluating students' participation in her large online literature course, which had over one hundred students. She mentioned that she had some activities that encouraged participation. For example, in her class all student essays are available online for all students to see, and students can comment on each other's work.

### **Important Aspects of Summative Assessment**

Two participants emphasized the importance of assessing a variety of skills in their summative assessment. Also, both of these participants thought that using various forms of assessment to measure different kinds of performance is important, not only because it allows a teacher to get a sense of what students know, but also because it is a fairer approach to summative assessment:

If we assess only speaking, it's not fair. If we assess only writing, it's not fair. If we assess someone's ability to write good essays, it's not fair. I mean all are fair but only as a portion, not as an entire class. Among other things, I'm trying to make sure that they read what I require. There are students who can write brilliant essays without reading. If they know a little bit about the work, they are so advantaged in comparison with others. There are good writers, and there are bad writers. But students who diligently read everything that is required and who are not writers also have the right to exist.

Two instructors said that their practice of grading anonymously was a positive feature of their summative assessment practices. Anonymous grading is an option available on learning management systems such as Canvas and Blackboard. Instructors grade students' assignments in this way without knowing their identities because in software such as SpeedGrader students' names are hidden. For the instructor who teaches a large online literature course, the opportunity for anonymous grading was one of the benefits of an online course which she liked, implying that it reduces the likelihood of bias.

### **Criteria for Summative Assessment**

Interestingly, the issue of developing clear criteria for summative assessment practices was not raised in the interviews. One participant included a few sentences in the syllabus, stating that she had certain expectations: “The summative assessment is designed to measure progress and judge whether the submitted work meets the expectations outlined in the learning objectives.” The learning objectives included only three points, stating that upon completion of the course, “students will be able to discuss historical, social and cultural aspects” of the course content, “recognize recurrent themes and motives in literary and cinematic works that treat subjects relevant to the course,” and “examine conflicting discourses on” the course content. However, the participant communicated her expectations to the students by posting a few graded works online as samples to inform students about her requirements and expectations.

In fact, none of the participants’ syllabi included scoring rubrics or other forms of determining how well the learning objectives had been met. As I mentioned earlier in this dissertation, rubrics do not solve the problem of the ambiguity of criteria. However, rubrics provide some idea of instructors’ expectations, reduce student anxiety, and help students reflect on their progress. Lack of transparency about teachers’ expectations disempowers students by creating feelings of uncertainty and a lack of control over their efforts. Therefore, students should know the instructors’ expectations from the very beginning, rather than after the first feedback on their work.

### **Summative Assessment as the Best Motivator for Learning**

One participant considered summative assessment to be the best motivator for learning: “I believe students need to be pushed somehow, and this is the way we push them. Otherwise, I don't think they will be studying.” She said that other motivators are not sufficient enough in comparison to graded summative assessment with strict rules. The instructor shares her own experience as a student: “I know how I am. If I don't absolutely have to submit something by a certain date, then I just wait.” During the interview, the participant went back to the topic of graded summative assessment a few times when we discussed various forms of assessment. For example, she considered continuous grading scale more motivational than the complete/incomplete grading. Also, in her comparison of formative feedback with grades, she stated:

[Students] are disappointed if they don't receive enough feedback. But at the end of the day, it is still the mark that they care about. Without this motivator, I'm saying honestly, I do not see how we could motivate them enough to study under such difficult conditions they are. Again, I'm thinking about myself. In the end, a good mark will console me with bad feedback in comparison with a bad mark and good feedback.

### **Summative Assessment through a Feminist Lens**

When describing the conceptual clusters, I inevitably looked at the findings through a feminist lens by applying the nine principles or themes of feminist pedagogy identified in Chapter 2: 1) privileging voice, 2) respecting the diversity of personal experience, 3) reformation of the relationship between professor and student, 4)

empowerment, 5) building community, 6) valuing the process of teaching over its content, 7) challenging traditional pedagogical notions, 8) supporting women's leadership and 9) caring. By *conceptual clusters*, I mean the major organizing concepts and the themes that aggregate around these concepts when they are discussed with participants.

As I mentioned earlier, it should be recognized that there is a power imbalance between a teacher and a student in a university class, and summative assessment—in which the teacher exercises complete control over a grade as well as the means by which students acquire their knowledge—is among the practices most contributing to power-imbalanced assessment approaches. Moreover, the more ambiguous criteria are, the more power the assessor has, because lack of criteria gives the teacher more opportunities to judge subjectively.

For example, the term *soft mark* that one of the faculty members defined as *flexible* in reality suggests its entire subjectivity, and lack of complaints from students regarding participation grades does not mean that they agree with the instructor's opinions. Also, depriving students of transparency in grading criteria is analogous to a lack of transparency in an employment contract, with basic wage terms and work hours unclear to the employee. In such economic cases, the value of work is hidden, capital is empowered, and labour suffers. It seems to me that none of the participants provided students with the criteria for their grading. Allowing students to revise work and encouraging them to “get better,” as one of the participants put it, seems to me less disempowering for students than not allowing students to revise their work at all. However, encouraging students to send in their drafts for feedback before submitting

their final work to be graded can be potentially more empowering than giving them an opportunity and time to rewrite their work to improve their grades.

Summative assessment of participation clearly violates principles of feminist pedagogy. As Jacobs and Chase (1992) point out, there are issues with including a participation score as a component of a student's final grade raises important issues such as a lack of instructions about strategies for improving participation, the disadvantaging of shy students, the subjectivity of the grade given, and problems with record-keeping. The participation grade is susceptible to bias because in assessing someone's behavior, people are prone to notice and remember things that confirm their preconceptions (Sullivan, 2012). Furthermore, such contextual factors as the gender of students and of the instructor, class size, the gender ratio in class, and the teaching style of the instructor affect student participation (Crombie, Pyke, Silverthorn, Jones, & Piccinin, 2003). For example, according to the study conducted by Tatum, Schwartz, Schimmoeller and Perry (2013), in classes where males outnumber females, female student participation decreases, whereas male participation is not affected by the number of female students in class. The right exercised by one of the participants to increase marks "to honour above-average participation" also seems to be questionable because it rewards outgoing students while punishing shy ones. Considering all these findings, allocating up to 15% of the final grade to participation significantly compromises the fairness of summative assessment, while assigning shy students to different discussion groups would probably put these students on the spot rather than increase their participation.

Finally, the participant's statement that "a good mark will console [her] with bad feedback in comparison with a bad mark and good feedback" speaks to a very

externalized and behaviourist notion of assessment with the narrow and limited type of learning it generates. Also, her claim that graded summative assessment is the best motivator for students to learn is neither in line with a feminist approach nor is supported by evidence. Indeed, how we assess depends on our perspective on the process of learning. According to Behling, Gentile, and Lopez (2017) graded summative assessment motivates students to outcompete their peers and avoid embarrassment rather than to learn. Studies show that such factors as the perceived enjoyability of the subject being taught, the importance of the subject to the student's sense of identity, self-worth and accomplishment, as well as its relevance to a student's personal life positively influence motivation and student academic engagement (Harackiewicz, Canning, & Tibbetts, 2015). Moreover, Darnon, Jury, and Aelenei (2018) found that students' motivation to learn (*mastery-approach goal*), perform and outcompete peers (*performance-approach goal*) do not even predict final grades, because the social class of students plays a more important role as a predictor of students' performance. Performance-approach goals and higher grades are correlated only for upper-class students, while mastery-approach goals positively predicted final grades for lower-class students (Darnon et al., 2016).

Student motivation to learn, perform and outcompete peers is shaped by the environment (O'Keefe, Ben-Eliyahu, & Linnenbrink-Garcia, 2013). O'Keefe and colleagues found that students become mastery-goal-oriented and interested in learning when teaching is engaging, and the focus is not on evaluation. Furthermore, students' motivation to learn is positively influenced by a mastery-oriented educational context that employs lectures and small-group work discussions, with an emphasis on creativity and

collaboration rather than high grades (O’Keefe et al., 2013). Thus, an instructor should therefore use various strategies to inspire student interest in the subject rather than disempower students with the anxiety of being punished by a low grade. Even a progressive evaluation practice as in contracting for a certain grade and the work it entails still involves grades that can be seen as associated with achievement orientation goals rather than mastery orientation goals.

### **Formative Assessment**

All the participants were asked the following interview question on formative assessment: “Formative assessment is assessment performed during instruction, and it aims to form students’ competence by providing them with appropriate feedback.

Usually, formative assessment does not involve grading. Do you use any formative assessment tools? What formative assessment tools do you find the most effective?” All the participants considered formative assessment to be important for improving student learning.

### **Feedback on Written Assignments**

As was mentioned earlier, two participants encouraged students to email drafts of their final work for instructor feedback before grading them. Another instructor provided formative feedback in her first-year language classes as preparation for essay-writing.

Before writing an essay for a grade, students write a few assignments in the form of an essay, and she corrects all their mistakes and writes her feedback without grading them.

The instructor provided the following example of her comments: "Try to create more complex sentences. Your sentences are too easy." She found these assignments useful for

two reasons: students built self-confidence as they wrote 100 words in a new language, and they learned what the instructor's requirements were. In response to my question: "Can you share with me any thoughts of how a teacher can provide more feedback and do less grading?" her answer was "Comments and ungraded feedback." She also said that she considered "personal" feedback to be the most useful in her practice. I interpreted "personal" as "individual," because the former means belonging or relating to a particular person (Personal, n.d.)

Both professors who taught large classes confirmed the importance of specialized feedback. One of them claimed that in large classes, instructors should make absolutely certain that they are available for feedback. Furthermore, the participant said that in certain cases, teachers have to approach students and proactively give feedback, and he used extensive feedback on students' essays as a formative assessment tool. The other professor (who also taught classes of over 80 students) provided students with formative feedback during his office hours: "You want to meet with [students] as often you can. I don't know if you've seen it, but I have a long line during my office hours."

### **Ongoing Feedback during Class Activities**

Two instructors reported that they used ongoing formative assessment during their teaching. One of these educators used formative assessment in her language classes by involving students in class activities, constantly asking them questions, trying to engage every student. She also tried to keep students actively involved, asked students questions, and conducted discussions in her literature classes:

And when we're doing a discussion, of course, they say such beautiful things that they take me by surprise half of the time. If it is something not very clever, I still

try to make a point out of it. But if it is something that gives me and everyone else food for thought, I always praise them.

The other proponent of giving ongoing feedback prefers collective feedback during her lectures rather than individualized personal feedback. In her language classes, she provided feedback to the whole class by identifying common mistakes and explaining them in class. She never singled out students who made those mistakes. If she used a sentence from students' essays, she changed them to such an extent that students did not recognize the sentences they had written. In her literature classes, she did a workshop using research proposals from the previous academic year (with the permission of students). She asked for students' feedback on those proposals by ranking them as "strong, stronger, weak, weaker and the weakest." She said:

There are no grades. Do you know what students like about this? They say, "We see others' mistakes, but we do not see our own. Even [those] who did not write good research proposals could make out the faults of a bad research proposal.

### **Formative Assessment through a Feminist Lens**

From a feminist perspective, practicing less grading and more ungraded feedback not only improves teaching and learning but also plays a crucial role in reducing the power imbalance in class. Although all the participants expressed a positive attitude toward formative assessment, they did not elaborate on exactly how they used it. From the examples provided by the proponents of ongoing feedback during class activities, they both appear to focus on students' errors – in other words on their weaknesses rather than their strengths. This practice is referred to as a deficit-model approach, which

assesses knowledge deficits of students after assuming a norm for them (Collins & O'Brien, 2011).

Also, the shortage of personalized feedback in the classes of the proponent of collective feedback may hinder student learning because common mistakes do not apply to all students. Furthermore, ongoing assessment of students' foreign language skills is a form of formative assessment. As one of the educators describes it, it informs her as the instructor about students' knowledge and skills but does not include feedback to every student on their performance. Indeed, in high enrolment courses, it does not seem possible to provide every student with caring and detailed feedback during class instruction. Yet written formative feedback outside a classroom seems to be problematic as well, because in high enrolment classes instructors have no time to focus on deficiencies or the development of individual students.

### **Complete/Incomplete Grading**

Five participants used complete/incomplete or pass/fail grading along with attendance and home assignments, and one participant used the complete/incomplete grading to evaluate group discussions, participation, and presentations. One participant who taught large classes did not use a complete/incomplete grading. He said: "I would use complete/ incomplete more on a preparatory level but not for a finished exam." He, therefore, thought that in a class of over 80 students, additional assessment seemed to him not possible because it would make assessment and evaluation more time-consuming.

### **Home Assignment**

One professor found the complete/incomplete grading effective for establishing whether students had read their assigned text. In one of her courses, 10% of the final grade was assigned to the completion of reading reflections on the texts, and students had an opportunity to receive full marks even if they missed one reflection. The instructor thought the complete/incomplete grading to be a fair method of assessment for reflection assignments because “it is either done on time or is done on the next day or it can be zero if it is not done on the next day.” She chose the complete/incomplete grading for this activity to motivate students to read the assigned works. Since I did not ask her why she preferred this option to continuous scale of grading, I can only speculate from her comments and my experience about the reasons for her choice. According to the instructor, complete/incomplete grading motivated students to read. Taking into consideration her previous statement that she does not like grading (discussed in more detail in Chapter 7), it seems reasonable to find alternative ways to give feedback.

### **Presentations and Group Work**

One participant reported that in her literature courses, 15% of the total grade was assigned on the basis of participation, presentation, and group discussions, and they were pass/fail—or as she called them “soft marks.” (As I mentioned earlier, in our department a “soft mark” refers to either flexibility for the evaluator or a so called “easy grade,” which is an easy assignment for which students can get a high score.) From the instructor’s comment, it is obvious that she used the term to mean “easy mark” because the mark increased the average grade in this class. For example, to earn full marks for

participation, students were required to make three presentations. Students recorded their participation in groups, and if students presented three times during discussion, they received full marks. She claimed that her approach “helps to get students involved,” and she considered this form of evaluation “quantitative” in the sense that “students have to do a certain number of presentations per term to get points. So, I will not grade the presentation or the quality of the group work because I see it as an ongoing process.” However, the participant did not specify how she provides feedback to help structure student learning. She continued with the following statement which switched the focus of the interview to grading: “The good thing is it really helps to get students involved, but on the other hand, it pushes your averages up.”

### **Complete/Incomplete Grading through a Feminist Lens**

From a feminist perspective, the complete/incomplete grading for home assignments and reading reflections seems to be a valuable assessment and evaluation option that does not create competition in class, elevate stress from students, or increase the power imbalance by assigning subjective and arbitrary grades that rank students. Continuous assessment based on complete/incomplete grading and combined with instructor feedback complies with feminist pedagogy and can be considered a feminist assessment method because it has a formative competence that guides the students to their next learning experience.

### **Peer and Self-Assessment**

Three participants reported that they had used peer assessment in their classes, and one participant commented that he would use it under certain conditions. All three

participants used peer assessment as summative rather than formative. Self-assessment has seven references, and three participants stated that they practiced or used to practice self-assessment in their classes.

### **Peer Assessment as Peer Evaluation**

Three participants had ambiguous attitudes towards peer assessment. On the one hand, they saw an important benefit in using peer assessments because it engaged students. On the other hand, they doubted the ability of students to judge each other appropriately. I believe this problem arises from terminology confusion. From their comments, it seemed that they understood the term *peer assessment* as *peer grading*. For example, one participant was very cautious in using peer assessment. He said that he would use it “only in small classes under the condition that everybody has a fair share, and under the condition that it is mutual. So, students assess each other, and then we assess the assessments of the students... You gave them marks because of that... You have a two-way street.” For this instructor, *peer assessment* seemed to mean *peer grading*, and considering incidents described by other participants of students assigning inflated grades to each other, his hesitation and caution regarding peer assessment is understandable. Inflated grades will inevitably increase the class average grade. According to the participant’s statement, to give validity to grades, instructors need to “assess the assessments of the students.” Instead of the reduction of the instructors’ assessment load reported by Hughes (2001), peer grading actually adds more work for the teacher. Therefore, the instructor sees the possibility of using peer assessment only in small classes, only if the criteria is very clear (“You give them marks because of that...”),

and only if both sides (teacher and learners) put equal amount of effort in providing valid grades (“You have a two-way street”).

Two participants who used peer assessment as summative reported that students assigned very high grades to each other. One participant used formative peer assessment along with students’ reflections, which had complete/incomplete grades. She provided students with a detailed rubric with questions such as, “Have you considered who the author is?” Besides the formative peer assessment of reflection papers, students graded each other’s presentations. The instructor “gave half of the mark while the other half came from students.” In addition, all students had to write an assessment of the presentation they evaluated. In this way, the professor encouraged students to listen to the presentations of their fellow students, and according to her, “it worked,” i.e., the audience listened.

One professor claimed that in his class peer evaluation in group discussions negatively affected students’ relationships by creating “resentment” and “competition:”

[Students] all evaluated each other’s work, and they graded each other's work... Fred should get a 60, and Josephina gets a 90. And what I found is that this type of work created a very antagonistic type of competition. And I did not want it to be the result of the discussion. And eventually, I dissolved all this group work business.

### **Students Directed Seminars**

In addition to mentioning negative experiences with peer assessment, as discussed at the end of the previous section, the professor also talked about successfully using peer

summative assessment in Students Directed Seminars (SDS), a program that allows students to organize and run their own courses on the topics of their interests:

That actually works very well. If it is the case from the start that students know that they are the ones who evaluate each other. That's it. Right? So, I have done that actually for ten years in a row now. And they are much better with that because that's the original deal going in.

The professor mentioned only one shortcoming of summative assessment in SDS, namely, the problem of having grades assigned both by students and by the instructor:

For me the problem is this mixture. A faculty gives the mark, and the students give the mark, and the confusion arises from the split there. In the classes where I am the faculty, I just give them a mark in the classes where students have the topic, and they actually enjoy it.

By the "split," the participant means the gap between the grades assigned by the students and by the faculty, and this gap seems to be the issue that creates some tension:

By "split" I meant the gap between the students' opinion of their work and my opinion. Their opinion was based on their sense of how hard they tried, and their sense of their own level of achievement. My opinion was based on 38 years of reading students' papers. There was, inevitably perhaps, a major difference: the average grade in my classes is 76, whereas the average grade in the SDS classes is 90 or even higher. The SDS classes are very specific examples, however. They are the only ones in which students largely have the final say on their grades.

Ultimately what tends to happen in SDS classes is that the students' grades are the ones that are entered. As a result, departments are increasingly less willing to

support SDS classes. Departments often feel that students' self-grading is extreme, and so they don't want to support that option.

### **Self-Assessment**

One participant reported that she had implemented online self-tests for students after every module. The self-tests were optional, and no marks were assigned for this activity. However, the instructor noted that in contrast to reflections that had complete/incomplete value, students were reluctant to complete self-assessment tests because their focus was on high grades rather than on learning:

Do they do it? No. They don't. Because they know if they don't do the reflections, they will lose a point and a half. But if they don't do the self-check test, they don't know how much they are losing.

Another professor shared his experience with students' self-assessment of their participation in classes in which he had used a participation mark. He used to invite students, after grading them, to tell him what they thought of their grades. The professor thought that by reflecting on their participation and their participation grades, students were doing self-assessment. He considered it to be "a good way to get [students] to think about what they were doing."

### **Peer/Self-Assessment through a Feminist Lens**

From my standpoint of a feminist pedagogue, peer and self-assessment are valuable forms of assessment that encourage student involvement and responsibility. In contrast, peer grading practices are even less reliable than grades assigned by instructors because students may unintentionally assign unreliable grades due to lack of experience

in grading. Also, students may give inflated grades due to the peer pressure of grading fellow students that cannot be alleviated by rubrics and detailed instructions. Using peer assessment as summative rather than formative creates an unhealthy learning environment. Lack of nurturing and collaborative environment not only contradicts feminist pedagogy but prevents learning.

In contrast, formative online self-tests support students' learning by enabling students to take more control of their learning. However, the use of self-assessment of participation, as described at the end of the previous subsection, does not seem to comply with feminist pedagogy. To me, this practice looks like teacher-assigned participation grades rather than self-assessment. I can see a power imbalance in this practice because it does not only involve grading that can be avoided, but also includes a discussion between a teacher and students who are not equal in their power to assign this grade. It was also not clear to me whether the instructor changed the grade after the discussion but even if he did, students might feel embarrassed to self-assess their participation.

### **Diagnostic Assessment**

Six participants expressed a positive attitude towards diagnostic assessment, and they either used or would use it in their classes.

### **Graded Diagnostic Assessment**

One participant shared her experience of using diagnostic assessment in her language classes. When assessing the written work of students, she combined diagnostic assessment with formative assessment by providing detailed feedback and grades without the assignment counting toward students' final grades.

Two other educators said that the option that students had of rewriting their work to get a higher grade turned her summative assessments into diagnostic assessments. One of these instructors expressed a similar idea by calling his first examinations “a combination of diagnostic plus formative assessment.” For him, it was important to provide feedback to students on the first examination to give them a sense of what needed to be improved and what students could do to perform better.

Only after I transcribed the interviews did I realize that I should have included a definition of the main features of diagnostic assessment in the question. When the participants claimed that by giving students an option to get a higher grade they turned their summative assessment into diagnostic, they did not take into consideration that diagnostic assessment does not grade students’ performance. Usually it is carried out at the beginning of courses or used throughout courses to help students and teachers see where to go next.

### **Ungraded Diagnostic Assessment**

Another educator conducted diagnostic assessment in both literature and language classes, and there were no grades involved. In her language classes, she asked students to write a short essay on any topic to identify the level of their German language proficiency: “There are no marks for it. But then I immediately know ‘OK, this person has definitely lived for some time in Germany’ or ‘This person only has studied through books.’” In her literature classes, she used diagnostic assessment “through brainstorming” on the subject taught. She also asked students to write an essay at the beginning of the semester on the topic “Why did they take this subject?” This essay

helped the instructor learn more about students' backgrounds and the level of their knowledge about the subject.

One instructor said that he would have used diagnostic assessment in his literature class if it had been smaller:

Just as if I were to go into an advanced language class, I need to see how good the language is. If I go to a small class, what is their background? What courses are they taking? That what will determine to a certain degree the level of their interaction. And then I would say, "That is the diagnosis of where you are. I am going to systematically week by week tighten the screw."

I interpret this comment regarding diagnostic assessment in upper-level language classes as participant's frustration about advanced students with a high level of language proficiency who register for courses to boost their GPA without putting efforts into their learning. Instructors from all language programs, including myself, encounter these students. They demoralize those who study diligently but still cannot achieve advanced language proficiency. The participant's last comment indicates his hope to use diagnostic assessment as a tool for spotting advanced students and making their learning more challenging.

### **Diagnostic Assessment through a Feminist Lens**

I find the practice of diagnostic assessment to be aligned with feminism because it creates opportunities for instructors to adjust classes according to each student's level of knowledge and interests. By assessing students' prior knowledge, this type of assessment focuses on student experience, and it is deeply contextualized because it considers students' backgrounds. Diagnostic assessment that includes feedback to learners is

particularly valuable because it creates an opportunity for students to interpret the instructor's feedback and use it in their learning. Graded assignments with an option to “upgrade” the mark received is not a diagnostic assessment because diagnostic assessment should not be scored. Graded diagnostic assessment may be considered a pre-test that prepares students for an upcoming test and builds esteem in students. It should not hold weight for calculating a student’s final grade.

### **Essential Aspects of Assessment**

Essential Aspects of Assessment is a small category that contains only four references, and initially it was a response to the question “Is there an assessment method in your teaching practices that you find particularly effective and interesting?” Most of the time the answer to this question was related to a particular form of assessment. However, there were a few participant statements that are not related to specific assessment practices, but that refer to essential or desirable features of assessment and evaluation practices.

### **Focusing on Learning**

Three participants emphasized that assessment should facilitate student learning. One participant stated that he viewed assessment as “the learning experience” for students, and suggested that assessment should create opportunities for students so that they might “learn and benefit from it.” Another professor also highlighted that the purpose of assessment was not just assigning grades, but “it is important that [students] learn.”

### **Small Classes**

For one participant, an important quality of assessment meant “understand[ing] of the student and student’s progress.” To achieve this, sufficient contact with students is needed, which was only possible in smaller classes. Another educator also conveyed the idea that the quality of assessment increases in smaller classes. He expressed appreciation for the experience he had at a small, inexpensive community college where he took a German course. The limit for the language class size in his college was ten students, and he considered this class the best he ever had. In the late seventies, the teacher practised self-assessment when students videotaped and analyzed their performance in German and saw their progress.

### **Contextualized and Individualized Assessment**

Besides providing students with an opportunity to assess their progress, another professor thinks that instructors should assess knowledge of the kind of material that is interesting to students. In his classes, final exams included personal questions that allowed students to bring in their own background to make these questions interesting for them. These questions engage students more and help them overcome the problem of overestimating the difficulty of the questions asked.

The link between students’ background and assessment is obvious in one participant's course syllabus, according to which there are two six- to eight-page essays assigned that each comprises 25% of the final grade. The first essay is an “Essay on the Self,” in which students are supposed to describe their “understanding of the self, or yourself, in terms of a personal experience” in relation to a text from the course. In this

essay, the professor asked students to answer the following questions: “How does your experience relate to specific texts in particular and philosophy in general, in our view? What is the personal or the private, in the first place? Who are you, and how do you know who you are?” In the second essay, an “Essay on the Other,” students were expected to describe their understanding of the other or of others. Similar to the first essay, the questions of the second essay were connected to students’ background: “How does our experience of others or the public relate to texts? What or who is the Other, what or who is the public, and what is the difference between the public and the private?”

This professor suggested that various assessment forms needed to provide the teacher with sufficient information on students’ level of knowledge:

The way I know that they know the material is mostly from their writing and to some extent from their speaking. Gradually what you start noticing in the course of the term is that students start to speak differently, and they start to use different cons [arguments in opposition (Con, 203)] and use different connections even as they are speaking. This happens even more in their writing. They start to create arguments that are different, far more complex, much more informed...

In my experience using individualized assessment, I found that students learn better when assessment tasks such as essay, presentations and group discussions topics are relevant to students' background and interests. The challenge of individualized assessment is in the diversity of students. Every class is different, and assessment tasks and topics that were relevant to students in the previous academic year often do not satisfy students’ needs in the next year.

## **Caring**

One participant identified himself as a proponent of the pedagogy of care, and he emphasized the idea that assessment needs to be caring. In the form of a story, he described how he applied a pedagogy of care to assessment by extending the deadlines for students. Usually, he had five or six students in his classes who needed extensions. After the term was over, an Arts Advising officer usually contacted him via email regarding these students because they did not have grades. Interestingly, the professor described their correspondence in the form of a dialogue, which made his story more dynamic and in a very condensed and powerful way demonstrated his attitude towards the Advising officer:

And so very regularly by mid-May C, an Arts Adviser, writes to me, “OK, you still have five or six students. Why don’t you give them their mark?” I say, “Well, they have not finished their work. Let me check with them.” And every summer I have students who just hang on until June, and I finally get their work from them. What I try to do is to hold off the pressure from Arts Advising to give the students time to finish. And that’s part of this pedagogy of care.

The participant thought that students submitted their work that late because “taking five or six courses is punishing” for them. “The more we can do to help them out, to help them remember things, and learn things, is better. I really hold on to their mark as long as I can,” he concluded.

### **Assessing Teaching**

One participant was not sure whether I meant assessment of students' knowledge or assessment of his own teaching when I asked him about effective methods of assessment. Later, he emphasized the importance of assessment of teaching: "I'd like to add another dimension to this, and this is an assessment of the assessor, assessment of the professor." Then he suggested that assessment in the university should include assessment of teaching in addition to student evaluation conducted by the administration. Since the participant was a member of our Peer Review of Teaching Committee, I assume he meant assessment of teaching as done by colleagues.

### **Essential Aspects of Assessment through a Feminist Lens**

All essential aspects of assessment reported by the participants comply with principles of feminist pedagogy because they have the potential to empower students. Assessment that facilitates knowledge empowers students because knowledge is power while by making assessment questions interesting and relevant to students' background, assessment empowers them by prioritizing students' personal experiences. Similarly, assessment in small classes would be more feminist because as Manicom (1992) asserts, small classes create more opportunities for discussion, for connectedness, and for sharing (Manicom, 1992). Tatum et al. (2013) observed that in smaller classes, instructors tend to call students by name and repeat students' answers, while student participation in smaller classes tends to increase particularly. Caring is also an essential aspect of assessment because it reflects the feminist value of caring. As mentioned in Chapter 2, some feminist

educators consider caring for students to be more important than correcting them (Noddings, 2013).

Assessing teaching is important for creating a positive learning environment and for maintaining a power balance in class. Early evaluation of teaching organized by the instructor has the advantages of having such evaluation go directly to the instructor without involving the administration, and of being empowering for students because they are involved in designing their learning and providing timely feedback when changes can be made during the current course (Keutzer, 1993).

Instructors' self-reflection on their teaching is another assessment tool that allows instructors to interrogate assumptions about teaching, to analyze their teaching in a social context, and to contribute to social change (Smyth, 1989). Formative assessment of teaching is a collaborative process that supports instructors in teaching by providing them with formative feedback and by forming a community around teaching. If peer reviews are organized by the administration, the preferences of reviewees in choosing reviewers should be considered. However, summative assessment of teaching conducted by the administration in order to maintain institutional standards or to promote faculty members creates anxiety, fear and resistance among faculty (Taylor, Atwood, and Hutchings, 2000).

I found formative assessment of my teaching by my colleagues to be useful for my professional development. In contrast, summative assessment had a negative impact on my well-being without even being conducted in our department. During the transition period of transferring our twelve-month lecturers to three-year contracts, there were discussions on organizing massive summative assessment which would involve observers

from other departments. These discussions created both anxiety and resistance among lecturers, and the whole idea was finally abandoned by the administration.

### **Summary**

In this chapter, I have described some positive experiences of participants with regard to assessment. The participants shared their perspectives on all forms of assessment in the university, including the assessment of teachers. According to the participants' reports, both summative and formative forms of assessment were the most often used forms of assessment (even in large classes), while a complete/incomplete grading, peer- and self-assessment, and diagnostic assessment were used less frequently. However, the participants did not elaborate much on their formative assessment practices. This is an area that needs further examination.

Most of the participants adjusted their assessment practices so as to make them more effective for learning. Also, a few participants combined several assessment approaches into one assessment task, such as implementing both summative and formative assessment. They did this by allowing students to rewrite their work or conducting formative assessment on students' essays and using these essays to diagnose student difficulties.

According to reports of those participants who taught large classes, all forms of assessment are possible in larger classes, including online classes. However, two participants said that they would use more diverse assessment methods and more individualized approaches to assessment if they had smaller classes. As for identifying important features of assessment practices used by the instructors, the participants

indicated that assessment practices should assess various skills, and assessment tasks should be perceived as interesting in order to promote students' interest.

Through a feminist lens, ungraded assessment practices, such as formative teacher's feedback, peer- and self-assessment, complete/incomplete grading, and diagnostic assessment reduce competition between students and mitigate the power imbalance between teachers and students. However, avoiding grading is not a universal solution to all problems in a graded classroom. For example, a deficit-model approach to formative assessment that lacks positive feedback about students' progress suppresses learning and disempowers students who learn not for personal needs but for the mark, for what they think the instructor wants from them in order to give them marks—the currency of higher education.

## CHAPTER 6: PRESSURES, TENSION AND DISCONTENT

### Overview

While reading the findings presented in this chapter, it is important to consider the background of the study. Shortly before the interviews there was a meeting devoted to “grade inflation” in our department where instructors with high class averages expressed frustration about being suspected as “GPA boosters.” The topic of grading became a subject of intense discussion in our department, and faculty expressed feelings of discontent with measures that might be taken by the administration against “grading leniency.”

The overarching category *Pressures, Tension and Discontent* has 99 references in the data coded into five categories: *Students’ Values and Attitudes*; *Pressures from Students*; *Pressures from the Administration*; *Discontent with Grading*; and *Self-Perceived Limitations*. I arranged my findings so as to link the categories in meaningful ways and to present the material in a narrative format.

Similar to the categories belonging to Participants’ Assessment Practices, these categories emerged from participants’ responses to thirteen open-ended questions related to their experience in various assessment practices (See Appendix D). Except for Question # 2, no question contains judgmental or evocative wording intended to elicit responses about a feeling of tension or frustration. I did, however, expect Question # 2 to lead the participants to express some tension. I worded the question as follows: “Administration discourages instructors from high class averages grade. This means you are supposed to keep in mind the class average grade when you are grading. Do you

agree with this university approach?” As I anticipated, the participants expressed their tension about grading and the policy of requiring the predetermined class-average grade

### **Students’ Values, Attitudes, and Capabilities**

Two participants expressed various levels of discontentedness with students’ values, attitudes and capabilities listed below under four sub-themes.

#### **Focusing on Grades over Learning**

Two participants articulated an intense frustration with student attitudes toward assessment and learning. One instructor’s primary concern was that “students are focused way too much on their grades than on the learning process.” Instead of concentrating on learning the subject, they “develop an attitude that is expected from them.” Moreover, she thinks that students want good grades, but they do not want to work for them:

They need to work. I allow them to repeat the same work. To improve it. They can work on their problems. This is something they cannot get at all. “How many points will I get for this? How many points will I get for that?” I teach a class that requires reading. I am fed up with talking to the class, talking to the people who don’t have a clue what I am talking about because they did not read the text.

The professor used a few instances from her practice as illustrations of students’ attitude towards learning. For example, despite her continuous encouragement, only 20% of students rewrite their final papers to get higher grades or send their first drafts for her feedback.

The educator also reported that students are reluctant to complete online self-assessment tests. She attributed their unwillingness to the fact that the online tests do not

count towards students' final grades. The professor was concerned as well that students used various excuses for late submission of their reflection assignments: "Students email me [saying] 'I am sorry, my grandmother is ill. How much will I lose if I do it later?'" The instructor expressed her discontent with students because she thought that they did not read the syllabus: "Students tell me in their evaluations: 'But you did not tell [me] this, or you didn't tell [me] that.' They complain that they didn't know. If I write too much, they complain it is too much to read."

From the participants' point of view, students' knowledge is superficial because they want a "shortcut." As I mentioned in the previous chapter, in the educator's classes, students know their exam topics ten days in advance. She highlighted that even though the instructions are very clear and students have ten days to think about how to answer the questions, students do not demonstrate critical thinking and tend to "offer what they know instead of answering the questions."

According to the participants, students' fixation on grades and grade rubrics prevents them from thinking and deep learning. The reason for the lack of deep learning is that, as she puts it, "thinking is not a requirement." She continues:

We create rubrics. In technical terms, rubrics are for grading. Can we say that a requirement for this class is thinking? Can we grade it? In a university level class, we cannot require thinking because everything we require from students' needs to be in the rubric.

Likewise, the professor seemed frustrated with peer assessment in her class, which probably was not as efficient as she hoped. On the one hand, she considered peer assessment of presentations useful, because students listened more carefully to each other

when work was presented in class. On the other hand, she was not satisfied with the quality of the assessment, because students assigned inflated marks to their friends. Moreover, there was limited use of peer assessment in her class, because students did not want to make an extra effort for this kind of activity: “[Students] are assessing each other only twice. And we learned it the hard way because students wrote complaints that they did all the work.” According to the participant, all the students wanted were grades not learning, and she could “make happy only those who are above the average.”

The other professor who encourages students to send their drafts in for his feedback and allows revisions of their essays with a chance of upgrading their marks, also found that students focus on grades over learning, and they desire to work less for their grades: “Because [students] are not really that concerned with the whole notion of improving or learning. They are more concerned with getting the number quickly and finishing leaving it like that.”

In his class of 85 students, only three students revised the examination essay by the time of the interview, and the professor expected that there would be a maximum of ten students by the end of the term. An opportunity to upgrade one’s mark therefore was not motivation enough:

They don’t see it generally as encouragement. And that’s a shame. That’s something I am always working against. I am always trying to encourage them to revise their work on the assumption they get better. This is really the darkest aspect of our approach to grading.

The conclusion of his statement implies a relationship between students’ attitudes toward learning and grading. The participant attributed the lack of students’ interest in learning

and unwillingness to revise their works to the university's flawed university evaluation and grading system policy:

Through this combination of administration-faculty-student involvement, students themselves impose their own numerical value on themselves and accept that instead of seeing it as a platform for going forward and for making improvements. An overwhelming number of cases. An overwhelming. I have been doing it for twenty-five years. Most students do not do the revisions...

Thus, this professor went further than an account of students' values and attitudes when he identified the origins of these attitudes.

### **Individualistic Values**

Besides dissatisfaction with student attitudes toward learning and grades, there are a few references to the research participants' concerns over other student attitudes, values, and shortcomings related to assessment. The participant who identified himself as a proponent of a pedagogy of care earlier expressed his discontent over individualistic values that contributed to an atmosphere of competition, resentment, and distance in his classes:

It is a very interesting psychological tight rope that you get on when you make your pedagogy of care. And what is underneath it is the feeling of encouragement... And promotion, and support... And students are not used to that. They are more used to being treated at a distance. Right? In a sort of cynical way, or skeptical way. Very often students just seize this attention and try to get as much attention as they can.

Students' individualism regularly causes detachment between students and teachers.

Students' unwillingness to connect is particularly frustrating for the professor because it is a barrier to applying a pedagogy of care.

### **Lack of Self-criticism, Gender-sensitivity, and English Language Fluency**

Besides discontent with students' fixation on grades and individualism, there were comments on students' other shortcomings, such as a lack of self-criticism, gender-sensitivity, and English language fluency. One participant commented that students have limited capacity to judge and evaluate their own or their peers' work: "Students are very lenient when it comes to them, and very critical when it comes to us... They see excuses for themselves." She indicated that students prefer to justify their errors in subject areas important for understanding the material being taught. The professor claimed that students argued "that it was a matter of opinion" when they, for example, confused the names of German special forces and squads of prisoners who worked in German concentration camps. In addition to a lack of self-criticism, students speaking English as a second language demonstrated limited gender-sensitivity and English language fluency in their work. The participant deducted marks for mistakes in English because she thought their limited language skills influenced their thinking abilities:

Instead of an SS person, [students] write an SS man. I write a comment: "Do you think only men do it?" I think I get more Asian students with lower language fluency or who were not taught to be gender sensitive. Not native speakers. And I think: "Do I have the right to include this into my marking?" I think yes because these things have an impact on the way they think. I always tell them, "I do not mark your thinking because I do not have access to your thinking." That means I

cannot evaluate thinking. When students say, “But I thought...” I say, “I am sorry, but I am not evaluating what is in your head, but I evaluate what you wrote.”

In the excerpt above, it is not clear whether students lack English language fluency or they make this error due to some cultural differences. It is evident, however, that the participant was questioning her right to deduct marks for this kind of mistake.

### **Discontent with Students’ Values, Attitudes and Capabilities through a Feminist Lens**

I attribute the reported lack of self-criticism in students to their fixation on grades because it is likely that their desire to bargain for a higher grade motivates them to justify their errors. As a feminist educator, I share the frustration of my colleagues with student’s tendencies to focus more on grades than on learning. However, I am against laying the blame for fixation on grades and reluctance to revise their work entirely on students. There is a need to consider the context in which students have developed these attitudes and how they have been trained to focus on grades and not learning. There is widespread evidence for a high level of stress and enormous pressure experienced by Canadian students in dealing with the challenges they face in higher education (Durand-Bush, McNeill, Harding, & Dobransky, 2015). By focusing on grades, students may merely try to get through the stresses of a heavy course load and the pressure of being graded on a curve. Moreover, in addition to such factors as distress and an intense workload, there is evidence that students' academic activities are influenced by critical factors such as teacher-student and peer relationships, as well as a student's sense of purpose in studying are critical factors that influence (Xerri, Radford & Shacklock, 2018). Thus, the lack of

motivation in learning and obsessive focus on grades perceived by participants needs more in-depth exploration so that a reconceptualization of assessment practices can properly address the various factors that contribute to this problem.

Students' fixation on grades fits the Western higher education culture, in which grades have become a currency rather than an indicator of progress and guide to decision making in the learning process. The discontent with excessive individualism among students should be addressed by society at large since it was the influence of societal forces that brought individualism and competitiveness into the classroom. Indeed, a person cannot be free from society while living and being a part of it (Lenin, 1905/1965). Yet, I believe that feminist pedagogues should facilitate students' awareness of how individualistic values have been cultivated over time by the western economic system and of the negative impact of individualism on interpersonal relationships and on the learning community. Similarly, the discontent with student gender-sensitivity should be addressed by conscious-raising rather than by deducting marks for lack of skill in using gender-neutral language.

### **Pressures from Students**

This category includes eight references from five participants, and initially, these references belonged to the category *Student's Values and Attitudes*. Indeed, when pressure from students is related to grading, these categories are very closely connected, and some references are coded in both categories. While I was reporting the findings, however, I concluded that not all kinds of pressure from students would be appropriate to include in the category *Student's Values and Attitudes*.

### **Pressure on Female Instructors**

A discussion of gender-bias in the assessment of students reminded one male professor about gender inequality in faculty promotions that the administration justified by citing the extra pressure that students put on female instructors. Students may seek academic and non-academic help from instructors, so this pressure is not limited to grading:

Because students talk more to women. That is a maternal thing. They are going to like you and trust you more than they are going to trust me. And that takes a certain [amount of] time. So, for me, it is almost a question of the other way around. Do the students take [into account] the gender of an instructor? Is that important for them? I just think that (probably overall this is an awful cliché) women still care more towards listening and coping with the problem by empathy than by rule (?). That simply means that students are going to take more time of a female faculty member than a male faculty member.

### **Pressure for Higher Grades**

Three participants stated that they experienced considerable pressure from students for higher grades. Two professors withstood the pressure and stayed firm with their decisions regarding students' grades, while one non-tenured instructor confessed that sometimes she gave in to students' demands. Interestingly, her statement regarding pressure from students was a response to my question about whether she had any gender bias in her assessment practices:

And also it is not about prejudice, but pressure. It is very hard to resist these overachieving students, mostly females, who want an A+. They put a lot of pressure on you. "I need 95%, and I want to go to med school. And after a while you think, "It is not 95%, it is 90%, but whatever."

### **Pressure from Students through a Feminist Lens**

Research evidence does not support the claim that extra pressure from students causes gender inequality in faculty promotions. The participant's suggestion that students would like and trust me more than him due to our gender difference appears biased. Studies show that such factors as instructor expertness and personality motivate students to communicate with the instructor (Myers, 2017), while underrepresentation of female faculty in senior academic positions is due to biased decision-making in appointments (Brink & Benschop, 2012). David (2015) presents strong evidence for patriarchy in higher education and that "the white male remains legitimately in power in HE and this remains the case because of the limited shifts in the overall political culture despite educational and socio-economic change" (p. 15).

The roots of the pressure from students for higher grades originate in the pressure students experience from the administration and from our competitive society, which was discussed in the previous section. The female student who is planning to apply for medical school asks for a higher grade because the medical school's decision is based on a high GPA, therefore "most successful medical school applicants come from wealthy, urban families and have high grades in addition to a resumé full of other achievements" (Vogel, 2018, p. E809). It is not clear whether the participant consented to assign a higher

grade because she felt compassion for the student and wanted to reward her efforts or because she was not sure that there was a distinct difference between 90% and 95%.

### **Pressure from the Administration**

In addition to pressures from students for higher grades, participants reported pressure from the administration and its assessment and evaluation policy. The category Pressure from the Administration has the highest number of references in the overarching category *Pressures, Tension and Discontent*. Six participants stated that they experienced various kinds of pressure from the administration, but mostly it was pressure related to students' grades. Out of 25 references coded as *Pressure from the Administration*, the overwhelming majority relate to university grading policies, particularly the "recommended" class average grades that attempt to curtail grade inflation.

### **University Grading Policies**

All three tenured participants who taught large classes reported that they had no problems with high class average grades in their classes. However, all three professors expressed discontent with university grading policies, or at least a disapproval of interfering with the instructor's domain of assigning grades.

For example, one participant articulated her discontent as a feeling that grading was out of her control, stating: "...grading is something that is taken from us." One professor, who was both a teacher and administrative faculty, claimed: "I disagree with the idea of discouraging instructors from high class averages quite simply because I disagree with the idea of interfering with instructors." Moreover, the professor shared his technique of grading students without keeping in mind the class grade average:

Sometimes I go so far as to not calculating the average. I do. Because then I would not put myself in the position when I treat students unfairly because I feel I should put the overall grade down. I will then say, “No, maybe the next test will be more difficult.” But I do not put myself into that position. It will make me a bad teacher.

The professor did not acknowledge that he felt pressure from the administration. However, the vocabulary he used, such as “go so far” and “I should,” indicate that he felt he was under the same pressure. Moreover, the character of these statements is worth commenting on it. “Go so far” implies a kind of rule breaking or stepping over the line. When the participant shared his thoughts that he felt like he “should put the overall grade down,” the “I should” might also come from some injunction imposed from above.

Another participant expressed the idea of pressure from the administration about grades directly. At the beginning of the interview, she stated that the combination of complete/incomplete grading and detailed feedback on students’ essays created more work for her, and at the same time it increased the class average grade: “That’s the frustrating part. I create more work for myself to help my students improve. Then I get higher averages, and then I’m getting punished again.” Thus, administrative pressure places limitations on the choice of assessment and evaluation practices used by instructors. In response to my question about whether she considered complete/incomplete grading without grading a fair method of evaluation, the participant stated she was “struggling with that all the time,” and would like to give more for group discussion than five percent of the final grade, but it “pushed [her] average up,” and for

“political reasons” she kept it “low.” She stated that she considered keeping the class average low to be discouraging.

Some references related to grading are coded in two categories, *Pressure from Students* and *Pressure from Administration*, which suggests that the participants felt pressure from both sources. One professor disagreed with the process of using grading as an instrument for comparing students’ academic performance. She also explicitly felt overwhelmed by the contradictory demands made on her: “University pushes us to keep grades somewhere, as they say, ‘on the curve.’ And on the other hand, students push us to ‘Give me more, more, more, more...’”

Similarly, another professor expressed the idea that faculty were caught between two fires:

You see, we are in a very interesting situation here... We have what the administration wants and have what we feel we’re forced to do. And we also have students’ expectations. And students’ expectations contribute to the grading system, to the grading pressure.

It is worth noting here the participants’ choice of words: “forced” when he describes the faculty feeling “pressure” in combination with “grading” here. I interpret this as indicating that the institution uses grades as a source of power and as an instrument for manipulation.

### **Conforming to Students’ Expectations**

The collected data also indicates that pressure from the administration can be applied in favour of conforming to students’ expectations. One educator recalled two conflicts with students that she considered unfairly resolved by the administration in

favour of the students. Once the administration allowed her former student to drop a class almost two years after the end of the course because the student had “resisted” accepting a lower grade. Another time the administration changed a failing grade assigned by the participant by giving the exam to another instructor to grade. She states:

I failed a student once. Once in my entire career because the students wrote a paper not on the topic. My colleague ...gave her a higher grade because they didn't want her to complain to the department.

Indeed, any grade modification after two years is unfair to the teacher who assigned the grade, and the whole incident looks like an example of treating a student as a customer. This places the instructor at the intersection of competing and antagonistic demands, one of which is that the institution and student are in a business-client relationship. As for the second expert from the interview with the participant, I believe it was a legitimate request from the student. All educational institutions have grade complaint policies that allow students to be re-evaluated if they consider the grade unfair.

The participant apparently resented her colleague for assigning a higher grade. This excerpt is an example of the instructor's frustration about pressure from the administration and the negative impact the administration policy can have on faculty relations.

### **Creating Competition between Instructors**

The participants shared their negative feelings arising from market-driven competition between instructors for enrolment and high evaluations of their teaching. The administration puts pressure on faculty by setting enrolment targets. The more lenient an instructor's evaluation practices, the more likely students are to register for this

instructor's courses. Thus, teachers who assign higher grades and who are more flexible with deadlines and late submission are suspected of "buying" higher evaluations and higher levels of enrolment by lowering their assessment and evaluation standards.

For instance, one instructor expressed resentment towards faculty members who assigned higher grades to students, considering this marking policy unfair to colleagues: "It is not fine if some people believe that they could give students a higher average and some people still have to do it on the lower scale because it sets an unfair situation between the instructors." Thus, she implies that by assigning higher grades, some faculty members get higher student evaluation scores and the higher enrolments that are necessary for being promoted. At the same time, the participant stated that she had no problems with the class average and grading "if everyone follows the same rules more or less."

### **Large Classes**

Initially, there was a separate category, *Large Classes*, that belonged to the overarching category *Pressures, Tension and Discontent*. However, the categories *Large Classes* and *Pressure from the Administration* were so closely related that the issue of large classes became a sub-category of *Pressure from the Administration*. Indeed, faculty tend to increase the number of seats in their popular courses due to pressure from the administration to teach more students at lower cost. One statement from an interview captured the resentment and frustration of being under administrative pressure to teach large classes and nevertheless maintain the high standards of teaching: "I always feel slightly resentful that on the one hand the university expects us to up enrolment, and on the other hand they expect us to teach as if classes are 25 people."

The participant describes this situation as pervasive because other universities—including Harvard, which had “all the money”—experience the same problem. Moreover, according to this professor, in comparison with other departments in our university, our classes of 80 students were not “so large,” and our enrolment numbers relative to the number of full-time faculty members, “did not look good.” This statement indicates the pressure the participant experienced from the administration that demands high enrolment numbers. At the same time, the educator expressed frustration about the limitations in assessment due to his class’s large size. He stated that he would like to use formative and diagnostic assessment, and a complete/incomplete grading without assigning a grade value. He believed it was unrealistic to use all these assessment strategies in large literature classes. On how to follow the basics of best assessment practices and provide feedback as often as possible in large classes, he stated that “mathematically” it was not possible, and there is “nothing you can do on a satisfactory basis in a class of 80.” According to the participant, only in small classes is proper assessment possible:

Once you teach smaller classes I think that you can come to a fairly objective assessment based on your understanding of the student and the student’s progress. But that objectively presupposes that you’ve had enough contact with the student. And that would be possible in smaller classes. And then I would certainly provide much more feedback than an actual grading.

He also expressed the idea that he longed for the times when he taught smaller classes and knew every student by name.

Another participant implied that high enrolment negatively affects the quality of assessment. She said that in her class of 110 students, she did not have enough time to write feedback that was “done well,” and she focused on their deficiencies or as she called it “where they lose.” Thus, two participants who taught classes of over 80 students expressed dissatisfaction with the quality of their assessment in these classes. In contrast to them, another professor who taught a large class did not talk about lower standards of assessment in his classes, but he still thought that there should be a limit to the number of students in a class. He said that in that term, the total number of students he taught was 230, and he implied that he felt overworked: “You have to take care of yourself. But if you don’t, you burn out. There is a tremendous danger of burn-out in this job.”

All the participants expressed their concerns about the limitations of assessment in large classes. For some participants, a “large class” is a course with more than 80 students enrolled; for other educators, a “class” becomes large when there are over 30 students in a language class or 40 students in a literature course. For example, one participant considered a literature class of 40 too large for individualized meaningful assessment, while another instructor admitted that in a literature class of 50, she could not conduct assessment comparable in quality of that possible with a small class. In a class of over 40 students, she was able to give formative feedback only to those students who made an effort to write her emails and attend her office hours. From her point of view, a teacher can provide more feedback and less grading only with “commented and ungraded feedback,” which would be overwhelming in a large class. For this reason, she does not teach large online literature classes. Finally, all the participants considered a class of 30 students to be a large language class.

### **Pressure from the Administration through a Feminist Lens**

Even given the amount of literature about the inappropriateness of grading on a curve, the policies based on "recommended" average grades exist. Designed to minimize subjectivity inherited in grading systems, these policies distort students' grades, making graded summative assessment even more arbitrary and less reliable in conveying information about students' academic preparedness. Besides erroneous messages about students' knowledge, the policy prevents cooperation and sharing knowledge between students and creates stress and unhealthy competition in class. Recommended average grades disempower students because students' grades depend on the performance of other students in class rather than on the quality of their work. This policy also disempowers instructors by ignoring good teaching and creating the conditions under which instructors face a moral dilemma: having to decide between their goal of helping students master their subject and the need for some students to underperform so that the assigned grades will comply with the "recommended grade averages."

The reported pressure from the administration about conforming to students' expectations probably implies the metaphor "student-as-customer" (Xu, Lo, & Wu, 2018). I believe that the student's role as *customer* or *client* is the reality in our market-driven society with high tuition fees in higher education. Xu et al. (2018) point to other metaphors used to define students, metaphors that co-exist and are not mutually exclusive: *student-as-partner*, *student-as-client*, *student-as-employee*, *student-as-manager*, *student-as-co-producer*, and *student-as-citizen*. Xu et al. (2018) argue that there is no evidence for the existence of these roles that reduce student and faculty experience to pure business. I believe that students' needs and expectations should be

considered while it is essential to make sure that the faculty does not label students as *customers* when students seek fair treatment.

The sources of tension between instructors due to competition for high enrolment and their discontent with large classes are symptoms of the marketization of the university, which refer to as the *neoliberal university* (Darder, 2019). The term *neoliberalism* “symbolizes a general trend toward greater marketization and the upward transfer of wealth as well as power to the financial elite” (Bloom, 2017, p.3). The culture of neoliberalism is characterized by competition, preoccupation with ranking criteria, managerial transparency, accountability measures, and diminishment of job security (Darder, 2019). The process of marketization and corporatizing higher education began in the 1990s with the university’s increasing focus on cost efficiency, and this process raises concerns as to whether the university is able to serve the public interest (Newson, 1998). The data in the sub-category *Large Classes* is consistent with the claim of feminist educational researchers deNoyelles, Milanés, & Dunlap (2016) that in large classes pedagogues make sacrifices at the expense of their educational values. The rise of student enrolment numbers without an increase in university budgets forces the administration to put extra demands on instructors (deNoeyelles et al., 2016) This solution, however, negatively affects the quality of assessment in both literature and language classes in the department.

### **Discontent with Grading**

The category Discontent with Grading contains responses to the interview question “What are your perspectives on grading?” Five participants expressed tension and discontent with grading, while three participants faced a dilemma considering

grading a necessary but flawed method of evaluation that does not match the pedagogical goals and values of the university. Since the responses of many participants indicate that instructors are struggling with dilemmas related to grading, and since this form of evaluation is prevalent in higher education, focusing on these dilemmas is important for my research questions.

### **Grading is Disliked but Necessary**

This subcategory turned out to be the category with the highest number of references. Three participants considered the practice of grading a necessary component of teaching and learning. One participant emphasized the importance of grading as part of the university system and as both feedback and mark:

Definitely, grading is necessary. Students want to see results. They want to see improvement. It is absolutely necessary to give them feedback or some numbers. I think it's important also because it's part of the whole university system.

The participant even justified the university's control over the class average grade as a measure to prevent "grade inflation." However, she disagreed with the university grading policy from a pedagogical point of view:

There're three aspects of it: political, business, and the pedagogical. Political and business aspects - the university has to have certain standards in order to have a good reputation and not to be a lightweight university to get a lot of good students. But pedagogically I don't like it at all because it discourages things that I like. For example, participation marks or ungraded assignments where students experiment expressing their thoughts.

Another instructor also expressed the idea of the importance of grading practice, but she clearly emphasized her dislike of it: “Grading is something that I don’t really like to do. As much as I love my job, I hate this part. But I recognize that it is important.” When I made a comment that grading is a reality, and grades will influence students’ lives and future, she disagreed:

Yes and No. Nobody hires on grades. Who wants somebody who has straight A’s? There are students who know how to get A’s. Everybody tries to get somehow in the middle. The evaluations, the recommendation letters. I think at this point, grading for us is something that is part of the job, and for students it is harmful.

The instructor viewed grading as a practice of ranking students and she contrasted the current transcripts that reflect students’ ranks with the academic record practices used in universities when she was a student. She explicitly favoured the previous record policies:

When I started learning... My transcript showed the course I took, my grade, and the teacher’s name. Those names meant something. Now what students get are the name of the course, the grade, the class average, and the class size. So, you have to be above average. It means if the class average is 97%, students who got 96% are below the average... I can make happy only those who are above average.

Thus, both participants seemed to be not entirely opposed to grading as an assessment and evaluation method, but they expressed discontent with the current university assessment and grading practices.

Another participant's perspective on grading as a disliked but necessary practice resonates with the attitudes of the first two participants described in this subsection:

Grading is like paying taxes or going to a dentist. It is something that I don't like doing, but it is something which given the fact that my teeth decay and that the state needs money to build streets, it is necessary. So, it is necessary. I would rather not be doing it, but I do not see a way not doing it.

### **Grading is Harmful and should be Avoided**

Two instructors expressed an unambiguously negative attitude towards grading and did not consider grading to be necessary. In response to my interview question regarding perception on grading, one participant provided a very brief answer: "Grading is stressful," and then she recalled teaching a non-credit continuing studies foreign language course. She contrasted university courses grading with this course's assessment and evaluation practices that did not involve grades. The instructor argued that students in this class were highly motivated, and she was satisfied with the assessment practices she conducted there. Among the factors that motivated students to learn a foreign language were students' desire to learn the native language of their partner or their plans to go abroad. "[Students] just come to learn the language..." She concluded her story with the following suggestions: "I think we should go in this direction when we do not have the language requirement."

One professor positioned himself in terms of his view of assessment and evaluation before I even asked him any questions. He shared his positive experience as a student, and he told about evaluation and transcript policies of the university he graduated:

There were no grades, but what you got was the narrative at the end of the course about what you did in the course. And so, my transcript on graduating from the university was thirty pages of narrative. And really that is the form of grading I prefer. I do not prefer using numbers and letters. I prefer the narrative style of grading, but this is very rare...

The professor criticized the depersonalized alphanumeric depersonalized grading system. He considered it to be harmful for students' learning, and he regretted that instructors did not have any other alternatives to grading:

In our schools today, especially in big universities, we are driven to very numerically based kind of grading. Simply by technological rationality which looks for speed, efficiency and numerical equivalence, and really in my experience, [grading] flattens out the differences between people and ruins in my view the learning experience for students. On principle, I am against numerical grading. In practice, however, at [our university], that's the form we have to use.

### **Grading through a Feminist Lens**

Grading continues to be a serious challenge to feminist pedagogy, which makes me question the possibility of creating a feminist classroom in a contemporary educational institution. Summative assessment that involves grading as it is conducted in universities is flawed because it ranks students and contradicts the equitable principles of feminist pedagogy. The subjective nature of grading contributes to the unfair treatment of students due to bias. The power imbalance between students and teachers and grading as a manifestation of this imbalance is one of the most serious discrepancies between feminist pedagogy and teaching praxis. Also, as was mentioned in the description of the

subcategory “Pressure from Administration”, bell-curve grading creates an atmosphere of competition in the classroom and becomes a serious obstacle for creating community.

### **Self-Perceived Limitations in Assessment Practices**

Four participants reported their limitations and shortcomings related to assessment.

### **Gender Bias**

Assessment bias without conscious intention has the highest frequency as responses to the following question: “Does students’ gender impact the ways in which you provide feedback or grades?” (See Appendix D Interview Questions). *Bias* here refers to targeting “social groups, encoding stereotypes or reflecting prejudicial evaluative hierarchies” (Holroyd, Scaife, & Stafford, 2017, p. 12410).

The following extract from an interview indicates that a participant acknowledged gender bias in her assessment practices also it contradicted her explicit values. She admitted: “I think I somehow evaluate girls harsher than boys. What is really good about Connect [is] that I can grade students anonymously. So this [is] something I really like.”

Another instructor also acknowledged that she had some gender bias. For example, she wrote a different kind of feedback to female and male students, but she could not explain the difference. She said:

I can't really explain how different. And I don't think it translates into higher or lower marks. It's just a different approach. I'm not ready to answer this question.

You know, I noticed at some point, I noticed that I expect that girls are more

diligent than boys. I don't do it consciously. I assume somehow that girls are more diligent.

The educator also stated she thought that there was “no absolutely 100 percent fair way of assessing people.” She was open about her experience of conducting biased assessment:

We are just not equal by nature. I mean, we try... we strive towards fairness as much as possible. But I reserve the right to like a student not on a personal level and giving them a little bit... Maybe I'll do that one way or another. I mean, it could be hidden under something else. You may not notice the mistake when it is there and not with someone else.

Other participants claimed that they did not notice any gender bias in their assessment practices.

### **Bias on Prior Achievements**

This question on gender bias elicited responses on other biases as well. One instructor said that her language classes were “fairly mathematical,” once in which the teacher “calculated mistakes.” But in literature classes, she might be prejudiced towards quiet students and those who had not done well in the past, or so-called “C” students, by not expecting them to perform well academically: “If a student did not do well on the first work, you have a prejudice. You think, “Oh, this's just a “C” student again.” The educator defeated her bias based on students’ prior achievements by marking students anonymously. Interestingly, along with the reported lower quality of assessment in large classes, the instructor mentioned one benefit of a large- class assessment, namely, its

anonymity: “But with many students now I do not know who they are. When you see a student's name, you do not see this student's face.”

### **Intolerance of Uncertainty**

Finally, one participant confessed that such personal limitation as having difficulty tolerating uncertainty could also negatively affect her assessment strategies and cause a discrepancy between her pedagogical beliefs and pedagogical practices. Creating innovative assessment strategies required changes, which were challenging for her because they entailed risk-taking. For example, after she developed a syllabus for a new language course where she replaced grading of a considerable number of assignments with complete/incomplete grading, she felt uneasy because she was not sure whether the course would be successful after she implemented them. She admitted that she had problems tolerating uncertainty, and taking the risk of changing the method of assessment in her classes was stressful for her because she did not know what kind of results she would get by using the new method.

### **Self-Perceived Limitations in Assessment Practices through a Feminist Lens**

The cluster of “Self-Perceived Limitations in Assessment Practices” turned out to be the most difficult to describe and analyze due to the complexity of the issues involved. According to the meta-analysis of research findings from 20 studies conducted by Malouff and Thorsteinsson (2016), conscious or unconscious bias in grading is pervasive. This bias can be related to sex, race, physical attractiveness or prior experience with a student (Malouff & Thorsteinsson, 2016). In this study, only three out of six participants admitted that they had a bias in assessment. Those participants who did not report their

bias might not feel comfortable confessing inappropriate assessment practices or they might not be aware of their biases. Also, it is not clear whether all the participants were aware that they were working in a patriarchal institution that has an immense influence on us. Discussions aimed at raising faculty awareness of the pervasiveness of stereotypes in our patriarchal society will create the conditions necessary for combating these biases.

Many scholars suggest anonymous grading as a method of combatting bias in assessment. Indeed, this practice is widely used in UK higher education (Malouff & Thorsteinsson, 2016). Two participants expressed a positive attitude towards anonymous grading because they say that it prevents bias in their assessment practices. Yet anonymity in grading weakens the teacher-student relationship so necessary for feedback to be effective, and as a result, students who need help may stay unnoticed (Protivínský & Münich, 2018).

So how should a feminist pedagogue practice relational pedagogy and address potential biases? And if we are aware of our biases, how can we confront them? Is anonymity in assessment a good strategy for confronting biases in assessment? Does it actually reduce gender bias? First of all, it is important to recognize our own biases so that we can address them effectively. As a feminist educator, I think that anonymity in assessment undermines the type of personalized feedback discussed in Chapter 5 devoted to the assessment practices the participants shared with me. Furthermore, assessment carried out without knowing who our students are when giving feedback does not align with the feminist principle of care.

Finally, the last limitation described in this subsection, namely, the limited ability of instructors to tolerate uncertainty is universal to varying degrees. I believe that the

support of colleagues is a vital resource for addressing this limitation. What is important for faculty is creating a supportive community at work where instructors can share their concerns and limitations and get help from their colleagues.

### **Summary**

This chapter is devoted to reviewing the various kinds of pressure and assessment-related discontent that the participants shared during the interviews I recorded. Faculty are under pressure from a student body that desires higher grades and an administration body that requires keeping class average low. Apparently, obsession with high evaluation scores is pervasive, and it affects not only students but also instructors, if not our entire society.

A feminist view of human relationships is in opposition to the individualistic market-driven relationship model. Feminist pedagogues promote positive learning communities where everybody-- students and teachers alike--is interconnected, and relationships ensure all members' well-being within communities (Llewellyn & Llewellyn, 2015). However, pressure from the administration in the form of contradictory demands such as low class average grade combined with high enrolment and high student evaluation scores create tension between colleagues and obstacles for achieving interconnected feminist communities.

There are tensions and contradictions in what the participants say about their pedagogical beliefs and the university policy. On the one hand, they are very concerned with supporting their students' learning, but on the other hand, they still must provide arbitrary and coerced summative grades. The expectation from the administration to have high student evaluation scores and high enrolment numbers result in competition among

faculty members. In addition to these pressures, participants express their concern about the quality of assessment and the overwhelming workload laid upon them when teaching large classes. Competition among faculty harms collaboration in the academic community while large classes negatively affect the quality of assessment and the relationship between students.

The participants reported the following self-perceived limitations and shortcomings related to assessment: implicit gender and expectation biases and a dislike of uncertainty. Gender bias is the most reported limitation probably due to the question on gender bias having been included in the interviews. The participants' concerns can be summarized as centring on the mismatch between pedagogical values and pedagogical practices caused by the way in which university policies conceive of the student as a client and rank students hierarchically. These limitations can be combatted by raising faculty awareness of patriarchy in our society and by providing mutual support.

When discussing the negative aspects of assessment practices that cause frustration and tension, the comments of participants indicate that instructors are aware of the conditions necessary for meaningful assessment, but they are not able to create these conditions because of external factors such as large class sizes or because of unreasonable demands from the administration. Participants' personal limitations also likely contribute to this gap between what is needed and what is possible under present conditions, but insufficient data were collected on the faculty's personal limitations related to assessment to be able to draw strong conclusions on this issue.

Grading is the largest cluster, and it indicates that this view likely represents the perspective of most teachers on campus. Ambiguous perspectives have the following

common features: grading as it is presently practised in accordance with university policy is unavoidable and is generally disliked. Participants also share some moral dilemmas about fairness and the pedagogical value of grades. Interestingly, those participants who have some direct experience with ungraded classes explicitly express a negative attitude towards grading and voice their preference for ungraded classes. From a feminist perspective, summative assessment that involves grading as it is conducted in universities is flawed because it contravenes most of the equity principles of feminism and feminist pedagogy.

## **CHAPTER 7: MY JOURNEY TO BECOME A FEMINIST EDUCATOR**

### **Introduction**

This chapter describes the findings from my teacher journal that I started to write in the hope that recording my pedagogical practices and my reflections on the issues of assessment might shed some light on my research questions. I found very few empirical studies on assessment methods currently implemented by proponents of feminist pedagogy that would inform the theory of feminist assessment. Thus, I chose to keep the journal not only as an autoethnographic tool to answer my research questions but also as a means of improving my teaching practices so that they would align with feminist pedagogy.

I kept my journal from the end of August until the end of December 2016, which corresponds to an academic term at the university. I intended to reflect on my teaching starting from the syllabi I posted on the course websites at the beginning of the term to the final grade submission and my farewell letters to students. During that term, I was teaching four courses, extending over all four levels of the Russian language program: first-, second-, third-, and fourth-year courses. Two of them were large classes. I knew that the term would be challenging for me. Teaching four courses is always overwhelming, and I do not enjoy my work as much as I do when I have a regular three-course teaching load. Extreme fatigue negatively affects my pedagogy, particularly in high enrolment classes where I have problems remembering students' names, let alone giving enough attention to every student. Knowing this, however, I volunteered to

increase the number of seats in the classes and decided to teach an additional course in term 1 instead of term 2 so that I could get more data -- and more importantly, so that I might practice feminist assessment in the less favourable conditions that are so prevalent in higher education settings. I decided that if in that term I were able to align assessment in my classes with feminist pedagogy then other instructors could implement a feminist approach to assessment despite teaching large classes and having heavy workloads. Thus, on August 19, 2016, I made the first entry in my journal, and my journey to become a feminist educator started:

My literature research and long hours of reflections on the compatibility of assessment and feminist pedagogy did not bring the desired result, and I am still struggling with the question of whether feminist assessment is possible at all. In this academic year, I would like to make considerable changes in my assessment strategies, and I will try to align them with feminist pedagogy.

### **At the Beginning of my Journey**

Inspired by literature on feminist pedagogy that I was reading during the whole summer, I redesigned assessment and evaluations in the syllabi of all my courses, trying to reconcile assessment with feminist pedagogy. The primary concern I was hoping to address was a power imbalance between myself and my students, which I hoped to reduce. I intentionally use the verb “reduce” and not “eliminate” because a complete power balance does not seem to be achievable in higher education. At university, besides assigning grades, instructors have a lot of power over students by making important decisions on the content, communication flow, and pace of the course.

Literature on feminist assessment is extremely scarce, so I postulated that feminist assessment should be based on the characteristics or themes attributed to feminist pedagogy, namely “privileging voice, respecting the diversity of personal experience, reformation of the relationship between professor and student, empowerment, building community, and challenging traditional pedagogical notions” (Webb et al., 2002, p, 68), the value of process of teaching over content (Shackelford, 1992), and being caring and nurturing (Accardi, 2013; Crabtree et al., 2009; Fisher, 2001; Lewis, 1992; Monchinski, 2010; Ropers-Huilman, 2009).

While preparing my courses, I was keeping these themes in mind, and I was focusing on strategies that could help me teach according to these principles. For example, to facilitate privileging voice and respecting the diversity of personal experience, I was planning to develop assessment strategies that would take into consideration students’ individual differences, and that would allow students to demonstrate their progress and achieve their learning goals. I was planning to involve students in the process of creating learner-centered feminist assessment. It seemed to me that I had a clear vision of feminist assessment strategies, and as I noted in my journal:

These assessment strategies will match students’ level of knowledge, learning styles, and they will consider students’ interests. In order to achieve it, I will share my views on assessment strategies with my students and I will initiate discussions with them. I hope to get students’ insight on assessment practices and to get information on their learning goals so that I could adjust my practices accordingly  
09.01.2016.

## **Challenges and Pressures**

### **Can Summative Assessment be Feminist?**

While working on the curricula of my courses, I paid particular attention to graded summative assessment. I considered it the most challenging type for a feminist educator because of the power imbalance it created between instructor and student. Even before my encounter with feminist pedagogy, I felt tension when I was grading, and as a result, I was unhappy with myself. I attributed the problems I was experiencing with graded assessments to personality flaws, particularly to my desire to be liked -- even by those students to whom I assigned low grades. While working on this dissertation, I gradually became convinced of the incongruity between the principles of feminist pedagogy and the existing summative assessment practices. However, before I started to write my teacher journal, I thought that it was not graded summative assessment that violated principles of feminist pedagogy, but rather the way in which grading was conducted at the university. I therefore decided to develop and implement feminist summative assessment in my classes that would be objective, fair, and that would establish a power balance between students and myself.

Before even making any changes in the curricula, I started my teacher journal with reflections on summative assessment in all my courses. I expressed satisfaction with assessment in my first-year classes where I thought that I had a better power balance between students and myself than in my other courses. I faithfully believed that our tests and examinations were reliable and valid (See the definitions of these terms on p. 11). I valued these tests for being designed and developed collaboratively among faculty members working in the program. Students in all sections wrote the same final

examinations, which from my point of view validly demonstrated that students who passed the examination could handle communicative tasks on the topics we covered in class. I valued our tests for being based on recent Russian-language assessment and evaluation research, and for capturing various aspects of the Russian language. I considered our tests to be objective because the scoring criteria were specific. For example, answers for fill in the blanks questions with missing endings and prepositions could be unambiguously scored.

Those were my thoughts at the beginning of the academic year. However, during the time I was writing my journal, my attitude toward assessment practices underwent dramatic changes. By the end of that term, I started to question summative assessment as a feminist tool more and more often due to the inconsistency between the practice of grading and the feminist objective of eradicating domination in human relationships.

I became aware that we were using tests merely because we were required to grade and not because they were objective and useful in indicating the level of student language skills or improving learning. For example, students' final scores depend on the marks assigned to various tasks, and assigning a high weighting to a more challenging task would negatively affect students' grades. In short, the fact that a test is written by a human makes it subjective because it is the author of the test who decides how long and how difficult the questions should be, and whether the task containing a certain number of questions is long enough to assess and evaluate a student's knowledge of the material.

Also, after I scrutinized the relationships between students' grades and performance in class, I realized that our tests privilege those students who excelled at taking tests, regardless of their depth of knowledge and understanding of the topics. So

third- and fourth-year students tended to outperform first- and second-year students just because they were more experienced in taking tests. Considering all these factors, the fact that in our first- and second-year classes the tests and the final examinations are worth 75% of the final grade raised my concern.

At the beginning of my journal, I proudly described our program's practice of exchanging students' final exams between colleagues so that we did not grade our own students' works as helpful for decentering my power:

During the first weeks of classes, I let my students know that I will not be grading their final exams. I think it has a positive impact on our relationship. Thus, in my first-year courses, I view myself as a facilitator who helps students do well rather than an instructor who controls them.

After reflecting on this practice, I concluded that the idea of reducing the power imbalance by exchanging students' exams is flawed since the anonymously graded final examinations were worth only 35% of the final grade. The final grade also included in-class quizzes and tests that could not be evaluated anonymously because they were designed to provide students with their instructor's feedback rather than a score of an instructor who does not know them. Moreover, anonymous grading did not relieve pressure from students while they were writing the exam, nor did it decenter the power of the teacher who determined the difficulties of the exam. Also, giving my students' exams to another instructor was not a better approach but merely my escape from the responsibility of grading the tests I had written for my students. Moreover, this action removes a very important aspect of learning, which is the relationship between teacher

and student. This relationship is meaningful for both teacher and student because it allows catering learning to the needs of the student.

Furthermore, when I realized that our tests were subjective anyway, the practice of anonymous grading as a fair assessment tool and a solution for re-establishing the power balance in class ceased making sense to me. This was because of the instructor's role in determining the difficulty of the summative assessment tools used. The instructor's right to choose examination questions shifts the power balance towards the decision-maker.

Tests and examinations are not the only tools that are prone to bias. Indeed, no matter how elaborate my rubrics for written assignments or oral presentations are, the criteria remain ambiguous. When I tried to use a sample of a "holistic rubric" suggested by a well-known university to evaluate presentations, I found such criteria as "a clear and well-structured delivery" or "well prepared" and "obviously rehearsed" problematic. How should I know whether students have rehearsed their presentations or not? Thus, any assessment is subjective, but when it is summative, it becomes harmful because it is used to make important decisions about students.

In my lower-level Russian language courses, computer-based testing was the most promising type of summative assessment from a feminist point of view because it had features that made the online tests qualify as assessment for learning. Students could use these open-book tests at their own pace, and they had an unlimited number of attempts that allowed them to improve their results with practice and receive a full mark if they practiced enough. Unfortunately, however, while this seemed to be very promising as a feminist assessment tool, it has serious flaws. Initially designed as an assessment tool that

allowed students to try out and practise various kinds of tests, it had to be reduced to multiple-choice questions due to technical difficulties with the Cyrillic alphabet keyboard.

I believe this technical issue could have been resolved if we had not found out that multiple attempts made the test vulnerable for cheating. Cheating, or academic dishonesty, is any dishonest act intended to produce a more favourable outcome (Murdock, Stephens, & Grotewiel, 2016). The problem was that students did not get a new version of the quiz for each attempt. So some students completed one attempt with all answers wrong, printed out the correct answers provided after the attempt, used these answers as an answer-key for their second attempt, and received 100% for these tests. So the online tests that we considered to be summative assessment for learning failed. However, it can be used as a self-assessment tool for mastery learning if its score is excluded from the final grade.

I believe some students cheated because it was a graded assessment. Numerous empirical investigations of cheating indicate that performance-goals are positively associated with cheating (McCabe, Butterfield, & Trevino, 2012). Findings from one of the largest two-decade-long studies indicate that 65% of North American university students admitted to at least one incident of cheating (McCabe et al., 2012). The prevalence of graded classes may explain these findings.

In my upper-level classes, I was even more concerned with graded summative assessment than in the beginner's courses because in my third- and fourth-year courses I assessed and evaluated learning achievements of students with different Russian language backgrounds. As a possible solution, I made some attempts to evaluate students'

progress, effort and sometimes even attitude towards my subject. However, after some time I realized that for some students, it might take very little time to write an assignment while for other students the same assignment could have taken hours to write. Only those who wrote the assignments knew how much time and effort it took them to complete their assignments. Thus, determining the criteria for students' progress, effort and attitude proved to be an impossible task. Thus, ipsative assessment that involves grading is questionable not only from a feminist point of view but also as a promising alternative kind of assessment for learning described in Chapter 2 of this dissertation. The key is identifying progress in own learning that is based on tools that help you understand your own learning, like competency grading in digital portfolios.

At some point, I thought I implemented summative feminist assessment in my fourth-year Russian language class in which students were involved in developing evaluation criteria for the final project. Students actively participated in creating assessment criteria, and I found students' suggestions very useful. I was not surprised by that because years of learning have made students experts in their own education. Students suggested quantifying my grammar expectations. Together with students, we developed some distinct grammar criteria, such as “verbs of motions are used correctly nine out of ten times – 90%.” Also, we attempted to create clarity and transparency of the instructor's expectations based on the grammar material we learned. We tried to be specific and estimated an approximate number of sentences that should include complex grammatical constructions.

While I was applying these mutually developed criteria instead of my rubric, my assessment practices in this class seemed to be aligned with feminist pedagogy. However,

my encounter with the Common European Framework of Reference for Languages (CEFR) (Cephe & Toprak, 2014), an assessment test called the Test of Russian as a Foreign Language (TORFL) made me question the rubric that took so much time and discussion in class. I consider TORFL to be a holistic test that assesses four components of linguistic skill: listening comprehension, writing, speaking and reading. In contrast to our rubric, the test focuses on communication and positive features of learners' performance rather than on errors, and students can get bonus marks for using complex constructions and rich vocabulary that outweigh the marks deducted for errors. In addition, grammatical errors for which my students suggested deducting 10% do not count much in TORFL because they do not hinder communication. Thus, I realized that no matter who created the criteria, they are still arbitrary. However, process and ownership are essential even if categories and scoring process could be improved.

Furthermore, if I chose to use the TORFL criteria for grading, all my students would get over 90% for their graded assignments and tests, and their final marks would be only "A's" and "A's +." The university would consider such distribution as a sign of "grade inflation," and I would be asked to "raise my standards of assessment." In short, no matter how progressive and democratic this method of getting students involved in the development looked at the beginning, assessment remains subjective, and it stays summative, which means it ranks students. Moreover, if a summative assessment tool allows students to get higher grades, it becomes questionable for the administration.

One entry describes my contemplations on combining summative peer, instructor, and self-evaluation of group presentations in my beginner's Russian language that I was preparing for the next term:

Unfortunately, in my Russ 102, the oral group presentation has to be graded.

I have to follow the evaluation breakdown all the instructors agreed upon.

However, there are the following problems with oral group presentations that make them particularly challenging to grade:

- a. Some students cannot perform well not because they did not practice, but because they feel on the spot.
- b. There are no precise criteria for their performance.
- c. Some students do not contribute to group work, but they get the same mark for the draft of the presentation as other group members.
- d. We do not practice enough this type of activity; therefore, I do not feel comfortable in including this score to count toward students' final grade in the course.

I should probably persuade my colleagues at our next Slavic meeting that the complete/incomplete grading with an opportunity to repeat using the feedback on how to reach the required level is a better choice for presentations.

In the journal, I wrote about my decision to combat the lack of participation from some students with a participation mark. In the next term, I implemented a new evaluation breakdown. The presentations were worth 10% of the final grade: students received four marks for the content and four marks for performance while two marks they assigned to themselves and to each other for participation. I gave students evaluation sheets after their presentations, and they anonymously gave a participation grade to every participant, including themselves.

I used this practice as a measure for preventing what students call “free riding,” i.e., the situation in which those who did not contribute to a group presentation receive the same grade as those who did. At first glance, it may seem inconsistent to avoid participation marks in all my courses but to implement them in group presentations. There were a few reasons behind this choice. I thought that it was a graded assignment, so adding a participation grade to it would not make it more damaging for the power balance in class as it already was.

I also thought that this mark did not require any special training or detailed criteria from students. Students’ grading seemed to be reliable because in the grading forms students submitted to me, the grades that students assigned to themselves and received from their peers were similar or very close. Furthermore, a peer-grading of participation acknowledged the contribution of those who invested more time and effort into the project. Finally, I justified this practice by the idea that no matter how serious 20% of the presentation mark sounded, it was worth only 2% of the final grade.

Despite some seeming advantages of using the peer grading of participation in group presentations, this practice is questionable from my point of view as a feminist instructor, and that was the reason why I hesitated before starting to use this practice, and finally refused to use it in 2018. Looking back at this approach to peer-grading I used in 2016, I now realize that it is not a feminist practice. Indeed, it is not right to include more grading when it can be avoided, particularly peer grading, which creates a type of tension that is damaging for relationships between students. Moreover, grading as a participation motivator seems to be questionable no matter how small the weight of the participation score as a percentage is.

One of the possible solutions for this dilemma could be to discuss this matter with students, to tell them about the practice and ask their opinion about what would motivate them to participate in group projects. The only measure I know of at this point that would prevent free riding is offering students the grading of making individual presentations for which a participation grade would be irrelevant because there is only one participant. Usually there are three to four students in class who prefer to present “solo.”

In my journal, I analyzed my assessment practices in every class, and I gradually arrived at the attitude towards grading I described in Chapter 1 of this dissertation as being an act of ranking and labeling students according to their performance and as a source of power over them. Usually, my discontent with graded summative assessment is the most overwhelming and intense at the end of the term. The final grade submission can ruin gratifying feelings at the end of any successful course I have taught. Instead of satisfaction and joy, I have a bitter aftertaste and a feeling that I did something wrong. I feel especially uneasy when I assign low grades because I am sure that they have disempowered students rather than motivated them to learn more in subsequent courses.

One of my journal entries about an incident with a fourth-year male science student with a severe disability is particularly illustrative of grading’s disempowering effect on students. That student took my first-year Russian language class to satisfy his program requirement, but he failed a few quizzes in my class, and he was afraid that he would not be able to complete his degree. So that he could pass the course, he asked me to allow him to rewrite the quizzes, which I did. This incident made me realize how much power I had over my students. Indeed, I could have answered: “If I make an exception, it

will be unfair towards other students who had no opportunity to rewrite their quizzes.” But life is not fair. It was not fair towards the student who, due to his disability, had to use special devices to be able to write. This journal entry ends with a promise I made to myself to pay special attention to the issue of power imbalance and grading, how they are interconnected, and what can be done to give students more power and more control over their learning.

I must admit that despite all the measures designed to align summative assessment to feminist pedagogy, the mismatch between feminist pedagogy and graded summative assessment remained. After looking at graded summative assessment through a feminist lens during my attempt to become a feminist pedagogue, I came to a clear realization and became firmly convinced that summative assessment cannot be feminist because grading is harmful to learning, to the learning community, and to teacher-student relationships. The fact that that the number of universities in which grading is the core of their assessment policies, and there is even a “recommended” average grade, raises my concerns as a language instructor and a feminist educator.

### **A “Recommended” Class Average Grade**

Besides unavoidable graded summative assessment, the university policy that establishes a “recommended” class average grade to combat grade inflation has been a source of tremendous pressure. I strongly believe that this artificial grade ceiling is unethical and harmful for learning and for the community. Indeed, when I tell my students that my goal is to help students do well on the exam, I am not completely honest with them nor with myself. I know that in every class there are students who get lower grades, and there are various reasons for that: lower motivation, different priorities, high

levels of stress, heavy course loads, ineffective learning strategies, you name it.

Usually, these students are aware of their problems, and they do not blame the teacher or the course design for their low performance. Though it sounds horrible, I should confess that I cherish these students because they lower the class average and allow me to assign well-deserved, high grades to those who perform well. And I know that I am not alone in feeling that way. Isn't it awful? Does it have to be this way that a teacher wants her students to underperform?

In addition to being anti-pedagogical, a “recommended” average grade is harmful to the learning community. The lower the grade ceiling, the more apparent the harm becomes. The average grades in lower-level courses are under stricter scrutiny, and student academic competitiveness and detachment is particularly strong there. Students express a lot of interest in their scores in relation to the class average, and instead of focusing on their own progress, they want to know whose audio or poetry assignment or presentation delivery is numerically “better.” Students’ eagerness to compete is a serious obstacle for peer assessment and peer support in my beginners' classes. From a feminist perspective, teachers should strive for refocusing students to enhance each other learning.

High average grades are considered “problematic,” and when I argue with the administration about this unethical and anti-pedagogical policy, the response is always the same: "The administration combats grade inflation." In the journal, I described my worries about the high average grade I expected to have in one of my upper-level courses:

I calculated the grade average I expected in my RUSS 400, and I realized that even without the final paper score that tended to be high, the class average was

around 90%. After I informed [the administration] about high grades that I expected in one of my upper-level courses, I was told: “Leave everything as is, just make the exams tougher next time around. A very high average is like a very low teaching score: Nobody really cares if it happens once, but you want to avoid repetitions” (11.16.2016).

Thus, being in the middle of the hierarchy under the pressure from the administration from above and students who compete for higher grades from below, I felt caught between university requirements, students’ expectations, and my values and conceptions.

### **My First-Year Student Anna**

Interestingly, students do not seem to question grading as a method of evaluation, nor do they complain about the class grade average. At least there were no negative comments about grading in class or student evaluations. My so-called “A” students who were planning to apply for competitive graduate programs considered grades to be important. They told me that getting good grades for my courses would increase their chances of being accepted because high grades for such a difficult language as Russian would be an indicator of high academic ability and a capacity to overcome challenges. Moreover, even “unclassified” students who took Russian for self-interest without seeking a credit preferred to be graded as well so that they “knew where they were.” Furthermore, although all my courses allow the “Credit/D/Fail” grading option in which “Credit” (a grade of 55% or higher), “D” (at least 50 but less than 55%), or “Fail” (less than 50%), students who need credits do not use this option. When I promote “Credit/D/Fail” in class, students respond that they are interested in grades or that the

number of credits eligible for this option is very limited, and they want to use it with other courses.

The value of grades in our educational system seems to be so high that scores have become more important than knowledge, and sometimes learning is entirely forgotten. This focus is in opposition to feminist pedagogy because it values an arbitrary score rather than fostering community and the process of learning. I experienced an intense feeling of frustration with students' tendency to put so much value on grades. To reduce my frustration, I have created a narrative of a first-year student, whom I called Anna. She represents an extreme case of the perception of learning at university as an accumulation of scores.

Anna demanded 100% for all tests and graded assignments, which I interpreted as an obsession with grades. A 99% grade made her upset and even ill. Unlike students described by the participants and included in the category Students' Values and Attitude, Anna worked very hard for her grades, and she did not look for "shortcuts." None of my attempts to shift her values to learning worked. In class, I told students a few real-life stories about my former so-called "B students" who achieved incredible fluency in the Russian language. We also talked about the skills students obtained rather than grades. However, Anna did not give up her self-imposed unrealistically high standards. Over the semester, I kept receiving emails from her access advisor with requests to allow Anna to write quizzes later, each time with different reasons.

Usually, I appreciate it when students attend my office hours because this is our one-on-one time when I can help with the material or get to know the student better. During that particular semester, however, I did not enjoy my office hours because I had a

feeling that I held them solely for grade negotiations with Anna. She came to see me every week, which would have been glorious if she had expressed the least interest in learning the Russian language. Instead, when Anna needed help with Russian, it always returned to grades she did not accept or grades for upcoming tests she was worried about. Feminist concern for relationship, caring and the sharing of power was destroyed with this focus on grades.

Anna monopolized my office hours to pressure me to assign her higher grades rather than using this time for help or for discussions about the Russian language and Russian culture. Thus, in addition to the pressure from Anna, I felt dissatisfaction with myself because I felt that I was unable to support a student such as Anna who experienced these types of difficulties. It was evident to me that Anna did not cope with the demands of the university, and the system perpetuated Anna's obsession with grades, but I was not sure how much I could or should help her with her studies.

It is worth noting that I added a story about Anna later, after I reviewed my journal during one of my revisions of this dissertation. At first, I did not think that a case of severe obsession with grades was relevant to my story describing my attempt to implement feminist assessment. Moreover, since that student told me about her disability due to some mental health condition, I believed that feeling frustration with a person who suffered some psychological disorder was unprofessional and unethical. However, while I was reading and rereading my notes, I realized that the story needed some attention. This extreme case is important and relevant to my story because of the object of Anna's obsession. Her excessive concern with her grades resembles the anxiety produced by eating disorders such as anorexia, which individuals (mostly women) develop under the

social pressure to be thin. Why wasn't Anna obsessed with her pronunciation skills or deep knowledge of Russian grammar? Anna's unrealistic standards and desire for high grades were only partially self-imposed. The student was reacting to pressures from the university, parents, social circles, and potential future employers. The fact that her self-worth was wrapped up in her GPA number reflects the values of a society that privileges grades.

Being part of the system that grades and labels students, I ask myself what a feminist educator can do to help students to thrive at university? After looking back at my failure to help Anna, I came to the conclusion that my arguments were unsuccessful because grades were still the focus of my attempts to persuade her. A much better approach to deemphasizing grades would be grading less and providing narrative feedback that tracks student progress. So instead of putting my efforts into discussions about the unimportance of grades, I should have striven to engage Anna in learning the subject I was teaching. Indeed, in order to stir students' interest in the subject, it is essential to be passionate, creative, engaging, and inspirational because pedagogy is more than methodology: "As a craft and art, pedagogy is seduction and performance: we cajole, humor, invite, persuade, and convince in efforts to 'seduce' students into the knowledge we embody" (Luke, 1996, p. 288). Engaging instructional practices empowers students with knowledge, promotes participatory classroom community and puts the emphasis on the process rather than on the end results.

### **Challenges of Feminist Assessment in Large Classes**

As I have already mentioned in the introduction to this chapter, my lower level courses had high student enrolments. For example, I had 33 students (maximum capacity

for the rooms) in two classes, and for a foreign language course, these are considered big classes. There are a few entries in my teacher journal in which I expressed my dissatisfaction with my assessment practices in those large classes. In one of the classes, I admitted that I was not able to remember every student's name after almost a month of teaching:

This class is very large for a language course. There are still 33 students in the class. Two students dropped the course, but two other students registered. I have problems not only giving enough attention to every student but even memorizing their names. Following students' progress and providing quality formative feedback seemed almost impossible in these classes.

Feminist pedagogy focuses primarily on individual growth and learning” (Clifford, 2002, p. 109), and not knowing my students' progress prevented my teaching in large classes from being feminist. Also, not knowing students' interests and backgrounds considerably limited my opportunities of being caring just because I was not aware of each student's needs.

By the middle of the term, I managed to adjust my assessment strategies by making more notes about every student's progress so that students could get feedback on their written work. However, the quality of my immediate feedback during instruction suffered simply because there were so many students in these classes. Furthermore, formative feedback is more time- and energy-consuming than summative assessment, and I felt overloaded with work when there were too many students. I realized that I was not prepared for such large classes, and I made some attempts to implement assessment strategies that would work equally well in both small and large classes.

I implemented a combination of my own feedback and peer-assessment in my first-year large classes to improve assessment of students' pronunciation. First, we reviewed pronunciation rules and discussed common difficulties in Russian pronunciation. Then students read aloud in groups and provide feedback to each other. At the end of this activity, each group reported on their progress and most frequent problems to the class. Students seemed very involved in this activity, and they told me that they found it useful. Yet the following comment from a student evaluation survey indicates that formative assessment was not as efficient as I thought in a class with a large number of students:

If anything, I think Veta just needs to be a bit "meaner" about our pronunciation when we speak as a whole class, because there were definitely things we mumbled through that she wouldn't pronounce herself and would leave me unaware of how a word really sounded...

I interpreted this comment as a student's reluctance to hurt my feelings with a harsh critique of my assessment, and at the same time, there is dissatisfaction with a lack of individualized feedback on every student's pronunciation.

As in debates on the possibility of implementing feminist pedagogy in large classes, discussions on the possibility of feminist assessment seem to be useful for the development of feminist assessment theory. I concluded from my experience that feminist assessment is possible in large classes if instructors use personalized feedback in conjunction with peer- and self-assessed assignments as a solution for balancing their workload and assessment needs. However, if other factors increase instructor workloads

(e.g., large course-work, teaching a new course, etc.), the quality of assessment drops and it ceases to be feminist because it is not personal, not caring and not empowering.

### **Internal Barriers for Feminist Assessment**

In addition to the pressures and discontent I have been experiencing while working under the constraints of the existing education system, I feel internal pressures related to the mismatch between my explicit beliefs as a feminist educator and my actual behaviour. As I have already mentioned in the previous chapter, students are not the only ones affected by unhealthy competitiveness. It appears that being a product of the system I have been living and working in for over twenty years, I became prone to competition, particularly when it concerns students' evaluations of teaching (SET). In the journal, I confessed:

I check student response rate in the course evaluation survey a few times per day. There is plenty of time before the evaluation is closed, but I check. And it is not that I worry that my evaluations will be low. I know that it will be high. But how high? It is a spirit of competition and a desire to achieve. And I know that the rating does not necessary mean that I am a good teacher. It means that I am a popular teacher. However, this number is important to me.

Indeed, recent studies show that the widespread belief that students learn more from instructors with high SET rating is wrong (Uttl, White, & Gonzalez, 2017), and SETs score is influenced by such factors as instructor gender (MacNell, Driscoll, & Hunt, 2015), SETs design, course, student variables (McClain, Gulbis, & Hays, 2018), and a grade ceiling policy which is required class grade average (Gorry, 2017).

Biased assessment was another personal limitation that I discovered while I was analyzing my pedagogical practices. According to the journal, I became aware of my gender bias while I was grading essays anonymously in one of my upper-level Russian classes. I noticed that it was difficult for me to write comments without knowing the authors' names. I then suspected that my comments would be different depending on the students' gender. I realized I would have written more encouraging comments to support female students because I perceived them to be less confident in themselves. At the same time, I expected more effort from female students while I was more forgiving of the lack of effort from male students. After that discovery of my own bias, I examined their roots, which came from my perception of myself when I was a student. After this incident, I became vigilant in writing equitable feedback in assessments and evaluations. Meanwhile, it is not always clear when equitable feedback should be abundant in favour of caring for students who need my support. The same word has a different meaning to different students based on what the word is used for in the context of the student's work.

### **Promising Assessment Practices for a Feminist Educator**

In addition to challenges causing feelings of frustration, disappointment and tension described in the previous sections, I experienced joy and satisfaction with my assessment practices that I consider promising feminist practices, namely, formative, peer and self-assessment, and complete/incomplete grading. Unfortunately, in the literature review of my dissertation, these types of assessment fall into the category of "less common types of assessment" (see p. 37), and my own pedagogical practices are no exception. For various reasons, summative assessment was still predominant during my journey to become a feminist pedagogue.

## **Formative Assessment**

Formative assessment has been in the centre of my attention since I started to express some interest in feminist pedagogy, and I consider it the most promising for feminist assessment. Since I was trying to determine what a feminist formative assessment should be and what its purpose is, formative assessment has the highest number of references in my journal.

The primary purpose of feminist formative assessment is to motivate students to compete against their own personal best rather than against each other. I personally feel deep satisfaction when my feedback is caring, respectful, and encouraging. I noted, “Having a rather fragile self-esteem myself, I try to be very careful when giving feedback to students. Insensitive negative comments undermine students and damage their self-confidence.” Therefore, I try to use extensive formative assessment while bearing in mind that the assessor tends to overpower the person to be assessed, and even formative feedback can contribute to the power imbalance in class if it is inconsiderate or insensitive and disrespectful. Therefore, I try to ensure that my comments are sensitive, specific, include both strong and weak elements of students’ performance, and inform students on their improvements. Actually, this kind of feedback is time consuming and takes considerable effort because it requires tracking students’ progress and developing a relationship with them so that I know that my feedback will correspond to students’ needs, interests and personalities.

Among successful formative assessment strategies that focus on learning, relationships with students, and helping them gain confidence rather than on grading and ranking, I described collecting home assignments every class and providing students with

brief comments. This practice continues to be one of the most important forms of feedback in my classes to this day. At the beginning of each class, I pass around students' home assignments with my feedback. Then we start a homework discussion. I create a PowerPoint slide with common problems without answers, and students solve these problems. I use animation effects to make information more memorable. After students provide their answers, the correct answer with an explanation appears with a sound. If there are problems without one single answer, it is even more exciting because situations with multiple answers create discussions.

In my journal, I also recalled my practice of providing detailed feedback on students' home compositions along with grades. This practice did not seem to be as successful as I expected. Indeed, the next test results indicated that students did not pay attention to my feedback, which often took hours to write. Since 2016, students have been writing essays that I do not grade, but instead provide personal feedback on the strong elements of their writing and suggestions about what material to review. The results of the subsequent quiz demonstrate to me that this feedback is more effective than feedback accompanied by a grade. I suspect students are mostly concerned with their grades which then consumes their whole attention. This observation is consistent with that described in the literature review studies of Butler (1988) and Lipnevich and Smith (2009) that indicate that detailed, descriptive feedback in the absence of a grade is more effective than similar feedback accompanied by a grade.

One of my journal references that related to formative feedback on home assignments described successful use of formative feedback. From students' home

assignments, I found out that a few students did not understand material critical for learning Russian grammar.

I provided them with my feedback: “Please learn all the conjugations of the verbs “understand” and "speak." If you know their endings, you will be able to conjugate almost every Russian verb. You will be asked to conjugate them at the beginning of the next class." At the beginning of the next class, I asked these students to conjugate a few verbs in writing, and all students did well. No grading was needed!

After this revelation, I started to analyze students’ difficulties and write them my comments on what to review. I also informed students that I would ask them to complete a few tasks on this material at the beginning of the next class. Interestingly, this approach always works without any grading being involved. I think this practice deserves more attention than simply being mentioned in this dissertation as an assessment strategy that supports learning, and I am planning a qualitative study of this type of assessment in the future.

In the beginners’ classes, I considered the practice of assessing students’ reading skills while they were reading texts to each other in groups to be very useful. I walked around the classroom listening to their reading and pronunciation and taking notes on their development/progress. After classes, I sent emails to students with my feedback, which included recommendations on the required improvements. At the next reading activity, I found noticeable improvements in students’ reading skills. However, this form of assessment had a disadvantage as well—it was an extremely time-consuming task, and

I modified it by combining my feedback with peer-assessment, which I described in the section “Challenges of Feminist Assessment in Large Classes.”

In the upper-level courses, I used a reflective journal with formative written feedback as one of the effective assessment tools. The journals were focused on the topics covered in class. While reading students’ journals, I assessed their grammar, vocabulary, and writing skills. Interestingly, gradually students started to depart from the assigned topics. I did not mind those departures because I appreciated students’ decisions to write on the topics of their choice. Feminism is about “the honoring of perspective, the encouragement for reflection that makes us more aware of ourselves and our actions” (Gravett & Bernhagen, 2018, p. 18). I therefore consider an ungraded journal to be a feminist assessment tool. From a pedagogical perspective, writing skills in a foreign language become long-lasting if the content of students’ writing is personally relevant. I also felt honoured that students entrusted me with their memories, thoughts, and desires. After all, the purpose of the journal assignment was to motivate students to write in Russian and for me to assess their language use.

I thus found formative assessment to be the most effective in all my classes. In addition, I consider formative assessment containing sensitive and respectful feedback to be aligned with feminist pedagogy. However, the personalized feedback that I used in my classes was infrequent because it was very time-consuming. When I provided it more often in high enrolment language classes, I unavoidably felt overloaded.

### **Diagnostic Assessment**

Similarly to formative assessment, I consider diagnostic assessment to be a feminist practice because it supports learning by identifying students’ strengths and

weaknesses, and it allows the instructor to provide students with recommendations on what level of Russian to choose. I conducted a diagnostic assessment in the first class, but only in my upper-level Russian language courses. In these multi-level classes, diagnostic assessment is essential so that students can be assigned to an appropriate group. First, I ask students to answer a few questions about their background in writing, and then they write a few paragraphs about their interests and their experience with the Russian language. Students describe their language background by answering some specific questions, such as, “Do you speak Russian at home?” or “Did you ever attend a Russian school?” Depending on the students’ experience with the language, I suggest a particular group. Still, students can decide on their own what group they think would be more beneficial for their learning.

### **Complete/incomplete Grading**

In my classes, a complete/incomplete grading system worked well. This form of evaluation creates less stress for the teacher and fewer complaints from students. In my journal, I described how I used the complete/incomplete grading in the evaluation of students’ presentations in a fourth-year Russian language class. Students had to send their presentation scripts and visual materials to me so that I could write some suggestions to improve their scripts and correct possible grammar errors. By doing so, I achieved the following:

1. Students learned how to write presentation scripts;
2. They memorized grammatically correct structures rather than their own mistakes;
3. Complete/incomplete grading and corrections of errors prior to the presentation dramatically reduced anxiety in those who have public speaking challenges.

To eliminate pressure from students, I implemented a complete/incomplete grading system by assigning a full mark to all students who made their presentation. In case if students did not pass, they had another chance to present on a different topic or present on the same topic in my office. As a result, in this class, all students got a full mark for their presentations and feedback from their peers and me. To ensure students' learning, I established high standards of achievement for their presentations, which were described in the syllabus and discussed in class. Students were asked to demonstrate knowledge of their words (not reading from a paper or screen) and clear delivery (the audience should understand the presenter) and the ability to generate meaningful questions for the audience. Students who failed to achieve these standards had an opportunity to present for the second time using my own and students' feedback either on the same topic in my office or on another topic in class. The purpose of the complete/incomplete grading system for the presentations was to support the collaborative environment by decreasing competition among students.

I enjoyed teaching this class the most because together with students we created a friendly and stress-free learning environment by shifting our focus from grades. However, after giving another thought to my positive experience in using the complete/incomplete grading with presentations, I realized that I was too optimistic in implying that this form of evaluation was a sufficient condition for creating a collaborative environment. A feminist pedagogue should be cautious when establishing “high” standards that may cause a student to fail. Also, assigning “pass” or “fail” grades for presentations to students who prepared well or were not prepared at all is not a difficult task. I was lucky in that particular class that all students were well prepared. But

what if there were to be a borderline presentation when it would be problematic to assign either a “pass” or “fail” grade? Finally, only while I was writing this chapter did I realize that the students did not have any input into the standards, which does not make my use of complete/incomplete grading feminist! Now, I would involve students in standard development that would give them a voice in assessment.

### **Peer and Self-Assessment**

I recorded the implementation of self-assessment in the first-year class by giving home assignments in the form of tests in which each question had a reference to the material the question tested. In the beginners’ classes, tests with only one answer prevailed due to limited knowledge of vocabulary and grammar. These tests were easy to design and implement, and students knew what they needed to review if they had problems answering a particular question. I also provided students with answer keys so that they could grade themselves. Students submitted that test as a regular home assignment, which had the complete/incomplete value regardless of performance. Many students found this kind of assessment useful because after each module, they asked me to give them similar tests.

In my grammar classes, I gave students answer keys for their homework exercises so that they had an opportunity to study more efficiently by correcting their errors right away and not waiting while I marked their home assignments. I asked students to make corrections using different coloured ink so that I could see their problem areas.

I believe self-assessment in my first-year classes enables students to take more control of their learning and therefore it supports students’ learning. However, this assessment is limited to the form that corrects grammar rather than a deeper level of self-

assessment that enables students to explore. Furthermore, I think using only one form of self-assessment in beginners' courses is not sufficient. I should develop more self-assessment opportunities for all four levels of Russian classes, which will be formative and will not involve grades. Self-assessment can be in the form of self-rating checklists with statements like "I can identify the characters and settings in a story" and "I can infer the meaning of unknown words from the context of the text." Students can rate themselves on a scale from Strongly Disagree to Strongly Agree. Students can self-assess their pronunciation skills by analyzing their audio files.

In my teacher journal, I wrote two entries related to peer assessment in two upper-level classes. In both classes, I had discussions with students on what feedback we considered to be constructive. Then I made a short presentation, and all together we provided constructive feedback, which was supposed to be positive, specific, and clear. Usually, third- and fourth-year classes consist of students of native and non-native speakers of Russian. For non-native speakers, their listening comprehension of native speakers was particularly challenging. Therefore, I considered a peer evaluation of presentations in this class as a great success because for non-native speakers it was a great listening comprehension experience and they also enjoyed listening to presentations made by native speakers so much that they even applauded:

I wrote a very positive feedback to the presenters. Students also provided them with their feedback. I allocated 10 minutes for this activity; however, it took students less than 5 minutes to write their feedback.

Students seemed to be experienced in providing feedback. During presentations, students took notes, which was likely helpful for writing their feedback in a short period

of time. One student shared with me the notes she received from students with their feedback, and I was amazed by the insightfulness of their comments. Due to years of learning, students appeared to have become accomplished assessors who provided their peers with valuable feedback. Peer assessment of presentations in that class facilitated engagement in class and created a collaborative and friendly atmosphere. Students benefited from peer assessment because it gave the learners greater feedback from various perspectives. I noticed that students were interested to learn their peers' perspectives. Those are positive features of peer-feedback. On the other hand, I must admit that I saw students' feedback once, and it is very possible that this feedback was not always useful and appropriate. So there were drawbacks in this assessment practices as well, and peer-assessment needs more scrutiny from a feminist perspective. Peer-assessment also needs to be aligned with feminist pedagogy's principles of being caring and empowering, and I have reflected in my journal that I should discuss these essential qualities of assessment with students more.

### **Teacher's Self-Assessment**

Feminism “pushes us to do better, with the full knowledge that, in a world of differences and attending inequalities, perfection is not possible” (Gravett & Bernhagen, 2018, p. 18). Assessing my teaching through self-reflection for the purpose of improvement of pedagogy and assessment is essential. In the journal, I described my frustration when I realized that students did not do well on a quiz due to some flaws in my instruction:

This time, I do not need to write my feedback to every student. Everybody had the same difficulty. All students had problems using words within their context.

This means something is wrong with my pedagogy or with the textbook or with both. When we assess students' understanding of new material, we assess our work and our ability to explain. It should be one of the principles of feminist assessment.

Thus, students' performance can be considered as feedback to our teaching. But of course, it is not always the case. Sometimes, no matter how well instructors explain the material, students have different priorities or demands that prevent them from learning.

In our university, the administration conducts student evaluation surveys at the end of each term and peer evaluations through the faculty's mentors or mentoring committee regularly. Unfortunately, these forms of feedback are not sufficient for assessment of teaching due to their considerable limitations. Instructors get feedback from student evaluations after the submission of the final grade, when it is too late to make any changes in teaching. Although teachers can consider student feedback and adjust their pedagogy in the future, there is a drawback, because each class is different, and future groups of students may have different needs. Also, the questions contained in the evaluation survey are very general and are not tailored to particular courses. Peer evaluation also is not sufficient for assessment of teaching because observation of merely one class may not be representative of teaching performance. I must confess that my observed classes were very different-- or as I call them, "artificial"—when my students and I were being observed. In comparison to summative peer evaluation, formative peer observation, which one of my colleagues conducted for me on my request, was very useful. The class was authentic because students nor I paid attention to a person who was sitting in the back row, and the suggestions I received from my colleagues after that

observation were precious for me professional development. Moreover, my colleague told me that she found this observation beneficial for her teaching although she taught a different language. Occasionally, all instructors experience various problems with audio-visual equipment, and she liked the way I handled technical difficulty and remained "composed." However, I should admit that I found a solution so quickly not because of superior technical skills, but because of the positive atmosphere in class and the support my students provided in this situation.

An anonymous survey is a good alternative to student and peer evaluation. Approximately a month after the beginning of the term, I conduct a survey with questions like: "Do you find this course manageable?" or "What activities do you not find useful?" (See Appendix G). The survey is a valuable source of information for how I could adjust the course for a particular group of students.

Self-assessment of teaching may have different forms. I found my teacher journal to be a very useful self-assessment tool for examining and assessing my own pedagogy from a feminist perspective. Writing the journal has helped me reflect on my professional strengths and weaknesses and develop teaching strategies for the upcoming classes while reading the journal after classes allowed me to think critically about my pedagogy and to find solutions to recurring problems.

### **Some Outcomes of my Journey**

#### **Were my Attempts Successful?**

Looking back at my attempt to become a feminist pedagogue and the measures I took to establish a more balanced relationship of power between students and myself, I

believe my assessment methods and pedagogy have improved in general. This observation confirms Hutchings' (1992) claim that feminist assessment shares similar characteristics with educationally meaningful assessment. Since there is no measurement scale of power balance in relationships, I used student evaluations as a measuring instrument despite the link between grade inflation and high student evaluation. Indeed, as I mentioned earlier in this dissertation, the reliability of student evaluations of teaching effectiveness is a controversial issue. From my experience, although students are focused on grades, they prefer to be challenged rather than get "easy" grades and be perceived by their instructors as incapable of meeting high expectations.

Overall, my attempts to become a feminist educator resulted in receiving positive feedback from the students. Students wrote about my assessment and evaluation practices in their comments as "clear expectations," "helpful," "receptive," "respectful," "rigorous, with assignments due each lecture, but is also very effective," "a good level of feedback throughout the course" etc. I consider the following characteristics of assessment that I promoted in my pedagogical practices to be aligned with feminist pedagogy: providing respectful and caring feedback, giving appropriate levels of difficulty in assignments, and involving students in developing evaluation rubrics. I believe my approach to assessment created an enjoyable and low-stress learning environment.

However, feelings of extreme frustration persist despite some progress in assessment practices, and it was not my heavy workload that caused it. The constraints of traditional structures with pervasive graded summative assessment make me question whether feminist assessment and feminist pedagogy is possible in our contemporary educational structure. I agree with Luke (1996) that feminist pedagogy cannot be reduced

to merely some aspects of pedagogy without its political dimensions. Indeed, before making an attempt at restructuring the curriculum, the political background of the current educational policies should be considered. Educators who are proposing to make changes in educational policies should know answers to the following questions “Who holds the power at the institution?” and “Whose interests do the current educational policy serve?”

### **What Can Be Done?**

I stopped writing my filed notes at the end of the term; however, my journey to become a feminist pedagogue did not end... Feminist pedagogy is an ongoing journey, and I continue doing my best as a feminist instructor. Looking back at my records, I think there are certain things that can be done, such as grading less by using instructor’s feedback, peer and self-assessment, ungraded diagnostic assessment and complete/incomplete grading. I hope to use feminist feedback, which would be caring, respectful of various personal experiences, and motivational. That feedback would encourage students to compete against their own personal best rather than against each other, and to achieve it, I would apply Isaptive assessment by keeping track of how students improved. The purpose of assessment would not be grading or rating students’ abilities but providing students with practical skills that enhance learning, facilitate knowledge and empower them because “Knowledge is power. Information is liberating” (Kofi Annan, 1999).

It is important to raise community awareness by doing more research and sharing findings on alternative assessment and evaluation systems. Moreover, it is crucial to persuade the university administration to increase the number of courses allowed to be taken as “Credit/D/Fail,” to remove the class average grade from students’ transcripts,

and to abandon “recommended” class average policy as an artificial rigour and create an environment in which students will appreciate learning rigour because students get supported when they get challenging tasks and move to higher level of learning .

## CHAPTER 8: PULLING IT ALL TOGETHER

In this chapter, I will continue the preliminary analysis I started in my finding chapters, keeping the focus on the themes related to the research questions:

1. What forms of assessment do instructors in the Department of Modern Languages and Literatures use that comply with principles of critical feminist pedagogy?
2. What tensions exist for instructors related to assessment in higher education?
3. How can these tensions be addressed using a feminist reconceptualization of assessment?

### **Assessment Aligned with Principles of Critical Feminist Pedagogy**

According to the findings of this research, participants used various assessment methods, but only their ungraded formative types of assessment are aligned with critical feminist pedagogy. Therefore, the following types of assessment are compatible with feminist principles: formative instructor feedback, peer- and self-assessment, diagnostic assessment and formative assessment of teaching. These findings are consistent with some suggestions of feminist scholars who advocate participatory assessment, self-assessment, and participatory evaluation as feminist methods of assessment (Accardi, 2013; Hutchings, 1992; Soeiro et al., 2011).

Formative instructor feedback, diagnostic and peer- and self-assessment are *feminist* assessment because they reduce competition between students and restore power balance between teachers and students. All these forms of assessment, however, can be considered feminist under the condition that they are utilized in an ungraded class. They cease to be feminist in a graded class because grading disempowers students, creates

competition between them, and is a major source of the power imbalance between teachers and students. Another necessary condition for feminist assessment is the use of *feminist feedback*, which is sensitive, motivational, caring, and respectful of various personal experiences.

Out of all graded forms of assessment and evaluation, only the complete/incomplete grading has feminist potential. In comparison to grading involving either letter grades or numerical marks, the complete/incomplete grading reduces competition in class and provides opportunities for instructors to be caring by decreasing stress from students. Also, in contrast to the arbitrary letter or numeric grades that rank students, the complete/incomplete grading is less damaging for power balance between teachers and students. However, feminist educators need to be cautious when establishing standards for this type of grading. The complete/incomplete grading can be empowering if instructors set clear standards for "incomplete," and if students have an opportunity to participate in creating these standards. Also, it is essential that students have a chance for more attempts if they could not complete an assignment.

Assessing teaching by using (1) early evaluation of instructor's pedagogy by students, (2) a teacher journal as a self-assessment tool, and (3) formative peer-review of teaching are all aligned with feminist pedagogy. Early evaluation of pedagogy of instructors is empowering for students because they are involved in designing their learning. It is also empowering for instructors because this form of assessment is a rich source for useful information that may improve pedagogy. A teacher journal is a feminist self-assessment tool; it provides educators with an opportunity to reflect on their pedagogy and on the power dynamic in their classes. Finally, formative peer-review of

teaching can be a useful resource to improve teaching that can be beneficial for both the observer and the observed. Under the condition that the observer uses feminist feedback described above, this form of assessment aligns with principles of critical feminist pedagogy.

### **Tensions Related to Assessment**

This study collected rich data on the second research question, “What tensions exist for instructors related to assessment in higher education?” The findings of this study indicate that faculty experience significant tensions related to assessment. The preliminary analysis of the data described in the previous chapters has revealed the following main concerns and tensions participants shared:

- 1) discontent with students’ values and attitudes;
- 2) being under external pressure from students from below and the administration from above;
- 3) experiencing internal pressure from facing a moral dilemma between contradictory goals: helping students learn and the need for some students to underperform;
- 4) feeling discontent with the university setting as a place of intense competition.

All these overarching clusters are interconnected and interrelated, and they have one source for the tensions—discontent with the current assessment and evaluation policy. Indeed, feeling discontent with students’ values, students’ focus on grades rather than on learning as well as pressure from students who compete for higher grades would cease to exist if grading is abolished. Furthermore, an internal pressure that participants experience due to the moral dilemma, the mismatch of their goal to educate students and to comply with an arbitrary class average grade policy also would not exist without

grades. Finally, grading is the primary source for competition between students because students compete for a higher place in the bell-curve created by the university.

Participants try different things to modify their assessment practices so that they promote learning, but the university policy, the system itself that fosters grading, does not allow much flexibility in assessment. No matter how much effort instructors make to modify their assessment practices so that their pedagogy facilitate students' learning, even their most innovative practices do not work in graded classes. The constraints of the university assessment policy and the lack of results despite the efforts of instructors create confusion and discontent among faculty.

This study also revealed that the current graded summative assessment does not facilitate learning a foreign language or literature. This finding is consistent with a large body of literature on summative assessment according to which grading does not support learning. Moreover, there is widespread evidence that grading is problematic because it lacks the theoretical framework, has fussy criteria, and is prone to bias. As a result, grades are not a reliable reflection of student learning and critical thinking. Moreover, grading on a curve designed as a solution for some of these problems, made grading even more arbitrary.

### **Institutional Resistance to Alternative Assessment Approaches**

Despite the abundant evidence for problems with grading as an assessment practice, educational institutions continue using it as measurements of student performance, and some institutions even impose recommended grade averages on instructors to prevent grade inflation. Furthermore, institutions discipline instructors who

disobey their grading policies. For example, the administration of the University of Ottawa accused its faculty, Dr. Denis Rancourt, of committing academic fraud for assigning A-plus grades to all student in his fourth-year physics class after his futile attempts to replace the university's alphanumeric grading with the complete/incomplete grading scheme (Catano, 2017). Dr. Denis Rancourt suggests that the conflict had a political basis, and the university used his grading method as a pretext to fire him (Birchard & Lewington, 2013). However, it was determined that Dr. Rancourt "was not fired for his ideas or beliefs, but rather for persisting in violating the Collective Agreement by not grading on an objective basis after being warned on several occasions to do so" (Catano, 2017, p. 17). This conflict is an illustration of the importance for a university to have the perception of "objective" grading. Indeed, even if there are courses with complete/incomplete grading option, universities tend to limit this option for students. For example, at my university, there are seven restrictions for complete/incomplete courses. Allowing students to take no more than four courses with this option and limiting complete/incomplete to elective courses create significant obstacles for students to use complete/incomplete grading. The abovementioned examples demonstrate a strong resistance of universities to abandoning numeric or letter grading.

### **Grades as a Market-Driven Commodity**

Higher education institutions justify grades as the means for accountability expectations for higher education. Indeed, grades are a convenient numeric tool for being accountable to the public and to business. However, taking into consideration all the recourses of universities, they could develop and implement alternative assessment tools

that would serve the purpose of accountability for student learning. Therefore, accountability does not explain the predominance of grades in the context of higher education.

Some scholars see the causes of the promotion of grades at universities in the marketization of education (Lynch & Hennessy, 2017; Wall, 2014). Wall et al. (2014) also claim that the assessment practices that quantify learning reflect a neoliberal market orientation of universities, and quantified assessment can make universities a better market product in their competition for students and funding. Lynch and Hennessy (2017) explain the role of grades at the contemporary neoliberal higher education context as “a desirable form of capital” that measures success, which causes students to shift their focus from inquiry towards investment in a high grade point average (p. 1753). Influenced by market ideology, universities use grades in the same way as the market uses objective economic justifications for dividing people into winners and losers (Lynch & Hennessy, 2017; Wall et al., 2014 ). Raaper (2019) claims that the neoliberal institutional assessment policy enforces consumerism in higher education and produces increasing pressures on both students and faculty.

### **Power, Authority and Hierarchy**

Universities keep using the current assessment and evaluation system because it is a tremendous source of power over students. Instructors have the authority to assign grades, and therefore, they have power over students. The more arbitrary the grades, the more power instructors have. A student may dispute their grades only within the institution and not in court because grades are the academic matter in nature and outside

the court's jurisdiction (Schuetze, 2013). Moreover, when a student enrolls at a higher education institution, "it is understood that the student agrees to be subject to the institution's discretion in resolving academic matters, including the assessment of the quality of the student's work" (*Murrey v. Lakehead Univ.*, 2011 as cited in Schuetze, 2013, p. 76).

Grades allow the administration to establish a policy of recommended class average grades that encourages students to compete. This competition creates social alienation among students. *Social alienation* refers to "detachment, separation, disengagement, or disconnection from the social world," which causes powerlessness in people (TenHouten, 2017, p. 1). For the administration, it is easier to control alienated and disempowered individuals than members of a consolidated student community.

The current assessment policy creates an unequal distribution of power not only for students but also for instructors. The administration considers student evaluation scores when it offers a contract renewal, and the lack of job-security makes the policy particularly disempowering for non-tenured faculty. Sessional instructors and lecturers are the ones who experience more pressure from the administration's requirement to keep grades low. Their student evaluation scores may be negatively affected if instructors do not conform to students' expectations of higher grades. Indeed, it is not a coincidence that all three tenured participants in this study reported that they had no problems with high class average grades in their classes. I suspect that their class average grades were low not only because they taught large classes in which grades tend to be lower in our department but also because low student evaluation scores do not affect tenured faculty's job security.

The current assessment policy demonstrates unequal hierarchal and patriarchal relationships of subordination at university. Instructors experience tension and discontent because, despite their efforts, assessment is counterproductive and flawed. It does not facilitate learning, nor it accurately reflects students' knowledge, but it serves to a patriarchal, hierarchical top-down institutional structure that controls both students and instructors. In this institutional structure, students are on the bottom, while the patriarchal administration is on the top. Instructors who have low job security positions are somewhere in the middle, and they are under tremendous pressure of both the student body and the administration.

### **Reconceptualization of University Assessment through a Feminist Lens**

This study responded to the third research question: "How can these tensions be addressed using a feminist reconceptualization of assessment?" There is one clear-cut answer to this question: there is an urgent need to replace the current university policy with a progressive and liberating alternative policy. In contrast to the dominant oppressive assessment policy, an alternative policy should implement assessment aligned with feminist pedagogy and encourage students to learn so that they become educated people.

Substantial research is necessary on alternative assessment and evaluation policies in which the following kinds of assessment would prevail: formative instructor feedback, peer- and self-assessment, diagnostic assessment and formative assessment of teaching. Assessment policies based on formative assessment while courses are in progress in conjunction with complete/incomplete grading after the completion of the courses have liberating potential. Also, formative assessment in conjunction with a comprehensive

narrative evaluation after completion of each course is another meaningful student-focused alternative assessment policy that complies with feminist pedagogy.

One might think that progressive assessment practices are not possible to implement because the dominant Western educational system reflects the values and principles of an individualistic hierarchical society where competition and self-interest prevail over collective or *communal interests*. For example, Kohn (2013) argues that we can see competition in every aspect of our lives:

in the pervasiveness of competitive activities and the fervor with which we approach them. Our economic system is predicated on competition, while our schooling, from the earliest grades, trains us not only to triumph over others but to regard them as obstacles to our own success. (p. 2)

However, despite individualistic market-driven values that dominate our society, there are successful examples of more progressive alternative assessment policies implemented in the United States. For example, Brown University has a unique evaluation system that is more student-focused in comparison to traditional assessment and evaluation policies. In Brown University, students have two options to get “full-letter grades A, B, or C (without pluses and minuses) or S (for Satisfactory). There is no grade of D, and failing grades are not recorded” (Brown University, n.d., para. 1). Brown University does not calculate grade point averages, and online portfolios created by students serve as qualitative evidence of students’ knowledge and skills (Brown University, n.d.).

A comprehensive narrative evaluation after completion of each course is another meaningful student-focused alternative assessment policy. One of the participants of this

study reported his positive experience of being a student of a college that used this form of evaluation. Currently, several American colleges, such as Hampshire College, Evergreen College, and New College of Florida, replaced grades with narrative evaluations for completed course (Evergreen College, n.d.; Hampshire College, n.d., New College of Florida, n.d.).

This kind of assessment and evaluation is more beneficial to student learning than a letter or numeric grade because it provides detailed and informative feedback that reflects individual strengths and weaknesses and contains suggestions for improvement. Administration of New College of Florida claims that this assessment provides faculty with an opportunity to look at student work “holistically, rather than assigning an arbitrary grade that doesn’t tell the whole story” (New College of Florida, n.d., para. 3). As a result, in New College of Florida, 80 % of students who got narrative evaluation instead of grades were accepted to various graduate schools within six years of their graduating, and 100 % of alumni who applied to law school got accepted (New College of Florida, n.d.).

Evergreen College conducts faculty evaluation of student progress, which faculty share with students at a one-on-one evaluation conference, and students have 30 calendar days to dispute it (Evergreen College, n.d.). The college combines faculty evaluation with student self-evaluation, which students submit at the end of each course (Evergreen College, n.d.). In their self-evaluations, students are encouraged to summarize their accomplishments and set new goals for the future (Evergreen College, n.d.).

Formative assessment and narrative evaluation are meaningful and beneficial for students but more time consuming for instructors. Therefore, it does not seem to be

possible to implement this alternative assessment policy in large classes with over 80. Discussion-oriented seminars offered in conjunction with large-enrolment courses may be a solution for the implementation of alternative methods of assessment.

It has to be acknowledged that assessment, even if there are no grades involved, is a source of power. Therefore, assessment is a political and an ethical matter, in which students are the most vulnerable and insecure party merely because they are the ones who are assessed. There is always a danger in imposing the instructor's values and judgments that seem independent of the form of assessment used. To protect students from possible oppressive practices, assessment should be transparent, and it should actively involve students.

### **Conclusion**

This study presents persuasive evidence that formative instructor feedback, peer- and self-assessment, diagnostic assessment and formative assessment of teaching are aligned with critical feminist pedagogy and can be considered *feminist* assessment. Instructors experience tension and discontent in assessment because their pedagogy is restricted by a university assessment policy that serves to a patriarchal, hierarchical institutional structure rather than facilitates learning.

The purpose of assessment should serve to students and the public good rather than to those who are at the top of the institutional hierarchy. Instructors' feedback should be empowering, which is caring, respectful of various personal experiences, and motivational. In combination with empowering feedback and student self-evaluation, alternative forms of assessment, such as Brown University assessment policy and

narrative assessment implemented in a few other USA educational institutions, have potential to be feminist.

More research is needed to investigate these alternative assessment policies, but according to some preliminary examination, these policies challenge the traditional pedagogical notion of graded summative assessment, support student community by eliminating competition between students, and create bases for a more balanced relationship between professor and student. Alternative forms of assessment are aligned with feminist pedagogy if they shift away from grading, provide students with academic freedom and allow their voices to be heard.

## Bibliography

- Accardi, M. T. (2013). *Feminist pedagogy for library instruction*. Sacramento, CA: Library Juice Press.
- Akyea, S. G., & Sandoval, P. (2005). A feminist perspective on student assessment: An epistemology of caring and concern. *Radical Pedagogy*, 6(2). Retrieved from [http://radicalpedagogy.icaap.org/content/issue6\\_2/akyea-sandoval.html](http://radicalpedagogy.icaap.org/content/issue6_2/akyea-sandoval.html)
- Alderson, J. C. (2017). The challenge of (diagnostic) testing: Do we know what we are measuring? In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 21–39). Ottawa: University of Ottawa Press.
- Aljohani, M. (2017). Principles of “constructivism” in foreign language teaching. *Journal of Literature and Art Studies*, 7(1), 97–107.
- Allen, K. R. (2009). Keeping the feminist in our teaching: Daring to make a difference. In S. A. Lloyd, A. L. Few, & K. R. Allen (Eds.), *Handbook of feminist family studies* (pp. 351–359). Thousand Oaks, CA: SAGE.
- Annan, K. (1997, June 23). “*If information and knowledge are central to democracy, they are conditions for development,*” says Secretary General [press release]. Retrieved from <https://www.un.org/press/en/1997/19970623.sgsm6268.html>
- Apple, M. W. (2012). *Can education change society?* New York and London: Routledge.
- Arnold, C. W. (1992). Seasoning your own spaghetti sauce: An overview of methods and models. In C. M. Musil (Ed.), *Students at the center: Feminist assessment* (pp. 51–64). Washington, DC: Association of American Colleges.

- Aronson, P. (2003). Feminists or “postfeminists”? Young women’s attitudes toward feminism and gender relations. *Gender and Society, 17*(6), 903–922.
- Assessment and Reporting Unit, Office of Learning and Teaching (2005). Current perspectives on assessment. Retrieved from <http://dl.icdst.org/pdfs/files1/21c18c23b5b9b12d00f58c43a13fca14.pdf>
- Astin, A. W., & Antonio, A. L. (2012). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. Lanham, MD: Rowman & Littlefield.
- Atkins, L., & Wallace, S. (2012). *Qualitative research in education*. London, UK: SAGE.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford, UK: Oxford University Press.
- Barnett, R. (2007). Assessment in higher education: An impossible mission? In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 29–40). London, UK: Routledge.
- Beenstock, M., & Feldman, D. (2018). Decomposing university grades: A longitudinal study of students and their instructors. *Studies in Higher Education, 43*(1), 114–133.
- Behling, K. C., Gentile, M. M., & Lopez, O. J. (2017). The effect of graded assessment on medical student performance in TBL exercises. *Medical Science Educator, 27*(3), 451–455.
- Behlol, M. G., & Dad, H. (2010). Concept of learning. *International Journal of Psychological Studies, 2*(2), 231–239.

- Birchard, K., & Lewington, J. (2013). A self-proclaimed dissident riles up Canadian academy. (Denis Rancourt). *The Chronicle of Higher Education*, 59(45).
- Black, P., & Wiliam, D. (1998a). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–144.
- Black, P., & Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 0(0), 1–25.  
<https://doi.org/10.1080/0969594X.2018.1441807>
- Bloom, P., & Taylor & Francis eBooks A-Z. (2017). *The ethics of neoliberalism: The business of making capitalism moral*. New York: Routledge, Taylor & Francis Group.
- Bloxham, S., & Boyd, P. (2007). *Developing effective assessment in higher education: A practical guide*. Maidenhead, UK: Open University Press.
- Boucouvalas, M., & Lawrence, R. L. (2010). Adult learning. In C. E. Kasworm, A. D. Rose, & J. M. Ross-Gordon (Eds.), *Handbook of adult and continuing education* (pp. 35-48). Thousand Oaks, CA: SAGE.
- Brantmeier, C. (2006). Advanced L2 learners and reading placement: Self-assessment, CBT, and subsequent performance. *System*, 34(1), 15–35.
- Brna, P., Self, J. A., Bull, S., & Pain, H. (1999). Negotiated collaborative assessment through collaborative student modelling. In R. Morales, H. Pain, S. Bull, & J. Kay

(Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 35-42). Le Mans, France: IOS.

- Brown, D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices*. New York: Pearson Education.
- Brown, J. (1992). Theory or practice—What exactly is feminist pedagogy? *The Journal of General Education*, 41, 51–63.
- Brown, V., Clarke, V., & Weate, P. (2016). Using thematic analysis in sport and exercise research. In B. Smith, & A.C. Sparkes (Eds.), *Routledge handbook of qualitative research in sport and exercise* (pp. 191-205). London, UK: Routledge.
- Brown University. (n.d.). Brown's grading system. Retrieved from <https://www.brown.edu/campus-life/support/careerlab/employers/employer-resources/browns-grading-system>
- Boud, D., & Falchikov, N. (2007) *Rethinking assessment in higher education: Learning for the longer term*. London, UK: Routledge.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40.
- Boyle, E., & Rothstein, H. (2006). *Effective college and university teaching: A practical guide* (2nd rev. ed.). Vancouver, BC: Granville Island Publishing.
- Buchanan, I. (2010). *A dictionary of critical theory*. New York, NY: Oxford University Press.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58, 1–14.

- Calfee, R. C. (1994). Cognitive assessment of classroom learning. *Education and Urban Society*, 26(4), 340–351.
- Campbell, K. (2002). Power, voice and democratization: Feminist pedagogy and assessment in CMC. *Educational Technology & Society*, 5(3), 27–39.
- Carless, D. (2015). *Excellence in university assessment: Learning from award-winning practice*. London: Routledge. Retrieved from <https://doi.org/10.4324/9781315740621>
- Catano, V., N. (2017). *Report of the independent committee of inquiry into the situations of Dr. Denis Rancourt at the University of Ottawa*. Retrieved from Canadian Association of University Teachers <https://www.caut.ca/sites/default/files/final-independent-committee-of-inquiry-rancourt-university-of-ottawa-2017-12.pdf>
- Cephe, P. T., & Toprak, T. E. (2014). The common European framework of reference for languages: Insights for language testing. *Journal of Language and Linguistic Studies*, 10(1), 79–88.
- Chang, H. (2008). *Autoethnography as method*. New York, NY: Routledge. <https://doi.org/10.4324/9781315433370>
- Child, J. (2019). *Hierarchy: A key idea for business and society*. London: Routledge. <https://doi.org/10.4324/9781315172378>
- Clifford, V. (2002). Does the use of journals as a form of assessment put into practice principles of feminist pedagogy? *Gender and Education*, 14(2), 109–121.
- Coeyman, B. (1996). Applications of feminist pedagogy to the college music major curriculum: An introduction to the issues. *College Music Symposium*, 36, 73–90.

- Cohen, L., Manion, L., Morrison, K. (2000). *Research methods in education* (5th ed.). New York, NY: Routledge.
- Collins, J. W., & O'Brien, N. P. (2011). *The Greenwood dictionary of education* (2nd ed.). Santa Barbara, Calif: Greenwood.
- Con. (2003). In *Merriam-Webster's Dictionary* (11th ed.). Springfield, MA: Merriam-Webster.
- Cook, J. A., & Fonow, M. M. (1986). Knowledge and women's interests: Issues of epistemology and methodology in feminist sociological research. *Sociological Inquiry*, 56(1), 2–29.
- Council of Europe (n.d.). *Global Scale - Table 1 (CEFR 3.3): Common Reference Levels*. Retrieved from <http://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale>
- Coutts, R., Gilleard, W., & Baglin, R. (2011). Evidence for the impact of assessment on mood and motivation in first-year students. *Studies in Higher Education*, 36(3), 291–300.
- Crabtree, R., Sapp, D., & Licona, A. (2009). Introduction: The passion and the praxis of feminist pedagogy. In R. Crabtree, D. Sapp, & A. Licona (Eds.), *Feminist pedagogy: Looking back to move forward* (pp. 1–20). Baltimore, MD: The Johns Hopkins University Press.
- Crawley, S. L., Lewis, J. E., Mayberry, M. (2008). Introduction—Feminist pedagogies in action: Teaching beyond disciplines. *Feminist Teacher*, 19(1), 1–12.

- Crombie, G., Pyke, S. W., Silverthorn, N., Jones, A., & Piccinin, S. (2003). Students' perceptions of their classroom participation and instructor as a function of gender and context. *The Journal of Higher Education*, 74(1), 51–76.
- Cronjé, J., (2006). Educational technology research and development: Paradigms regained: Toward integrating objectivism and constructivism in instructional design and the learning sciences. *Education Technology Research and Development* 54(4), 387–416.
- Crowe, S., Cresswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC Medical Research Methodology*, 11, 1–9.
- Cullen, R., & Harris, M. (2009). Assessing learner-centredness through course syllabi. *Assessment & Evaluation in Higher Education*, 34(1), 115–125.
- Curtis, R. S., & Wu, W. (2012). Learning and assessment: The application of ePortfolios. *Journal of Higher Education Theory and Practice*, 12(3), 66-74.
- Dalziel, J. (1998). Using marks to assess student performance: Some problems and alternatives. *Assessment & Evaluation in Higher Education*, 23(4), 351–356.
- Darder, A. (2019).“Foreword.” In D. Bottrell & C. Manathunga (Eds.) *Resisting neoliberalism in higher education volume I* (pp. v-x). Cham: Springer International Publishing.
- Darnon, C., Jury, M., & Aelenei, C. (2018). Who benefits from mastery-approach and performance-approach goals in college? Students’ social class as a moderator of the link between goals and grade. *European Journal of Psychology of Education*, 33(4), 713–726.

- David, M. (2015). Women and gender equality in higher education? *Education Sciences*, 5(1), 10-25.
- Davidson, E. J. (2005). What is evaluation? In E. J. Davidson, *Evaluation methodology basics: The nuts and bolts of sound evaluation* (pp. 1–12). Thousand Oaks, CA: SAGE.
- Davies, A. & Le Mahieu, P. (2003). Assessment for learning: Reconsidering portfolios and research evidence. In M. Segers, F. Dochy, & E. Cascallar (Eds.). *Optimising new modes of assessment: In search of qualities and standards* (pp. 141–169). Dordrecht, Netherlands: Kluwer Academic.
- Day, I. N. Z., van Blankenstein, F. M., Westenberg, P. M., & Admiraal, W. F. (2018). Explaining individual student success using continuous assessment types and student characteristics. *Higher Education Research & Development*, 37(5), 937–951.
- De Grez, L., Valcke, M., & Roozen, I. (2012). How effective are self- and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education*, 13(2), 129-142.
- de Laine, M. (2000). *Fieldwork, participation and practice: ethics and dilemmas in qualitative research*. London, UK: SAGE.
- Delmar, R. (1986). What is Feminism? In J. Mitchell & A. Oakley (Eds.), *What is feminism?* (pp. 8–33). New York, NY: Pantheon Books.
- deNoyelles, A., Milanés, C. R., & Dunlap, K. (2016). Ms/Use of technology: Reflections on feminist pedagogy from the technological front line. *Pedagogy*, 16(3), 481–509.

- Devault, M. L. (1990). Talking and listening from women's standpoint: Feminist strategies for interviewing and analysis. *Social Problems*, 37(1), 96–116.
- Disch, L. J., & Hawkesworth, M. (2016). Feminist theory: Transforming the known world. In L. J. Disch & M. Hawkesworth (Eds.). *The Oxford handbook of feminist theory* (pp. 2–18). New York: Oxford University Press.
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331-350.
- Doe, C. (2014). *Diagnostic English language needs assessment (DELNA)*. London, UK: SAGE.
- Durand-Bush, N., McNeill, K., Harding, M., & Dobransky, J. (2015). Investigating stress, psychological well-being, mental health functioning, and self-regulation capacity among university undergraduate students: Is this population optimally functioning? *Canadian Journal of Counselling and Psychotherapy*, 49(3), 253–274.
- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- Easthope, C., & Easthope, G. (2000). Intensification, extension and complexity of teachers' workload. *British Journal of Sociology of Education*, 21(1), 43–58.
- Effective. (2003) In *Merriam-Webster's Dictionary* (11th ed.). Springfield, MA: Merriam-Webster.
- El-Alayli, A., Hansen-Brown, A. A., & Ceynar, M. (2018). Dancing backwards in high heels: Female professors experience more work demands and special favor requests, particularly from academically entitled students. *Sex Roles*, 79(3), 136–150.

- Elbow, P. (1993). Ranking, evaluating, and liking: Sorting out three forms of judgement. *College English*, 55(2), 187–206.
- Ellis, C. (1991). Sociological introspection and emotional experience. *Symbolic Interaction*, 14(1), 23–50. <https://doi.org/10.1525/si.1991.14.1.23>
- Ellis, C. (2008). Systematic sociological introspection. In L. M. Given (Ed.), *The SAGE encyclopedia of qualitative research methods* (pp. 854–855). Thousand Oaks, CA: SAGE.
- Ellis, C., Adams, T. E., & Bochner A. (2011). Autoethnography: An overview. *Historical Social Research*, 36(4), 273–290.
- Ellsworth, E. (1992). Why doesn't this feel empowering? Working through the repressive myths of critical pedagogy. In C. Luke & J. Gore (Eds.), *Feminism and critical pedagogy* (pp. 90–119). New York, NY: Routledge.
- Engels, F. (2004). *The origin of the family, private property and the state*. Chippendale, Australia: Resistance Books. (Original work published 1884)
- English, L. M. (2006). Women, knowing, and authenticity: Living with contradictions. *New Directions for Adult and Continuing Education*, 2006(111), 17–25.
- English L. M., & Irving, C. J. (2015). Critical feminist pedagogy. In *Feminism in community: Adult education for transformation* (pp. 103–113). Rotterdam, Netherlands: SensePublishers.
- Etherington, K. (2004). *Becoming a reflexive researcher: Using our selves in research*. London: Jessica Kingley.
- Evergreen College. (n.d.). Evaluations. Retrieved from <https://www.evergreen.edu/evaluations>

- Fair. (2003) In *Merriam-Webster's Dictionary* (11th ed.). Springfield, MA: Merriam-Webster.
- Finlay, L., Gough, B. (Eds.) (2003). *Reflexivity: A practical guide for researchers in health and social sciences*. Oxford, UK: Blackwell.  
<https://doi.org/10.1002/9780470776094>
- Fisher, M. B. (2001). *No angel in the classroom: Teaching through feminist discourse*. New York, NY: Rowman & Littlefield.
- Flores, J. S. (2015). Feminist Pedagogy in the UK University Classroom: Limitations, Challenges, and Possibilities. In T. P. Light, J. Nicholas, & R. Bondy (Eds.). *Feminist pedagogy in higher education: Critical theory and practice* (pp. 33–56). Waterloo, ON: Wilfrid Laurier University Press.
- Freedman, E. (2009). Small group pedagogy: Consciousness-raising in conservative times. In R. Crabtree, D. Sapp, & A. Licona, (Eds.), *Feminist pedagogy: Looking back to move forward* (pp. 117–137). Baltimore, MD: The Johns Hopkins University Press.
- Freire, P. (1995). *Pedagogy of the oppressed* (New rev. 20th-Anniversary ed.). New York: Continuum.
- Fu, H., Hopper, T., & Sanford, K. (2018). New BC Curriculum and Communicating Student Learning in an Age of Assessment for Learning. *Alberta Journal of Educational Research*, 64(3), 264–286.
- Gaukroger, S. (2012). *Objectivity: A very short introduction*. Oxford, UK: Oxford University Press.

- Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1(1), 3–31.
- Giroux, H. A. (1988). *Teachers as intellectuals: Toward a critical pedagogy of learning*. Massachusetts: Bergin and Garvey Publishers, Inc.
- Given, L. M., Gale Virtual Reference Library EBook Collection, & Gale Group. (2008). *The sage encyclopedia of qualitative research methods*. Thousand Oaks, Calif: Sage Publications.
- Glassman, M. (2001). Dewey and Vygotsky: Society, experience, and inquiry in educational practice. *Educational Researcher*, 30(4), 3-14.
- Glick, P. (2007). Sexism. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 858–859). Thousand Oaks, CA: SAGE.
- Gorry, D. (2017). The impact of grade ceilings on student grades and course evaluations: Evidence from a policy change. *Economics of Education Review*, 56, 133–140.
- Grant, A., Short, N., & Turner, L. (2013). Introduction: Storying life and lives. In I. Goodson, N. Short, & L. Turner (Eds.), *Contemporary British Autoethnography* (pp. 1–16). Rotterdam, Netherlands: Sense Publishers.
- Gravett, E. O., & Bernhagen, L. (2018). Ways of doing: Feminist educational development. *To Improve the Academy*, 37(1), 17–29.
- Green, K. H., & Emerson, A. (2007). A new framework for grading. *Assessment and Evaluation in Higher Education*, 32(4), 495–511.
- Griffin, G. (2017a). Inequality. In (Ed.), *A Dictionary of Gender Studies*. Oxford, UK: Oxford University Press. <https://doi.org/10.1093/acref/9780191834837.001.0001>

- Griffin, G. (2017b). Reflexivity. In G. Griffin (Ed.), *A Dictionary of Gender Studies*. Oxford, UK: Oxford University Press.
- <https://doi.org/10.1093/acref/9780191834837.001.0001>
- Gronlund, N. E., & Cameron, I. J. (2004). *Assessment of student achievement*. Toronto: Pearson Canada.
- Gur-Ze'ev, I. (2005). Feminist critical pedagogy and critical theory today. *Journal of Thought, 40*(2), 55–72.
- Hamilton, L., & Corbett-Whittier, C. (2013). *Using case study in education research*. London, UK: SAGE.
- Hamp-Lyons, L., & Condon, W. (1993). Questioning assumptions about portfolio-based assessment. *College Composition and Communication, 44*(2), 176–190.
- Hampshire College. (n.d.) Our academy philosophy. Retrieved from <https://www.hampshire.edu/discover-hampshire/our-academic-philosophy>
- Harackiewicz, J., Canning, E., & Tibbetts, Y. (2015). Academic motivation and performance: Task value interventions. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 37–42). London, UK: Elsevier.
- Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies, 14*(3), 575–599.
- Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education, 36*, 101–111.

- Harvey, L., & Green, D. (1993). Defining quality. *Assessment & Evaluation in Higher Education*, 18(1), 9–34.
- Hassan, O. A. (2011). Learning theories and assessment methodologies—An engineering educational perspective. *European Journal of Engineering Education*, 36(4), 327–339.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hennink, M., Hutter, I., & Bailey, A. (2011). *Qualitative research methods*. Thousand Oaks, CA: SAGE.
- Hershberg, R. M. (2014). Constructivism. In D. Coghlan & M. Brydon-Miller (Eds.), *The SAGE encyclopedia of action research* (pp. 183–187). London, UK: SAGE.
- Heywood, L. L. (2006). Introduction: A Fifteen-Year History of Third-Wave Feminism. In L. L. Heywood (Ed.), *The women's movement today: An encyclopedia of third-wave feminism*. (Vol. 1, pp. xv–xvii). Westport, CT: Greenwood Press.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53–64.
- Hiles, D., & Čermák, I. (2007, July). Qualitative research: Transparency and narrative oriented inquiry. Paper presented at the 10<sup>th</sup> European Congress of Psychology, Prague, CZ.
- Hodges, L. C., Anderson, E. C., Carpenter, T. S., Cui, L., Gierasch, T. M., Leupen, S., Nanes, K. M., & Wagner, C. R. (2015). Using reading quizzes in STEM

- Classes—The what, why, and how. *Journal of College Science Teaching*, 45(1), 49–55.
- Hoffmann, F. L., & Stake, J. E. (1998). Feminist pedagogy in theory and practice: An empirical investigation. *NWSA Journal*, 10(1), 79–97.
- Holroyd, J., Scaife, R., & Stafford, T. (2017). Responsibility for implicit bias. *Philosophy Compass*, 12(3), p. e12410. <https://doi.org/10.1111/phc3.12410>
- hooks, b. (1994). *Teaching to transgress: Education as the practice of freedom*. New York, NY: Routledge.
- hooks, b.. (2014). *Ain't I a woman: Black women and feminism*. New York: Routledge. <https://doi.org/10.4324/9781315743264>
- Hopkins, A. H. (2009). Women's studies on television? It's time for distance learning. In R. Crabtree, D. Sapp & A. Licona (Eds.), *Feminist pedagogy: Looking back to move forward* (pp. 176–192). Baltimore, MD: The Johns Hopkins University Press.
- Hughes, G. (Ed.) (2017). *Ipsative assessment and learning gain: Exploring international case studies*. London, UK: Palgrave Macmillan.
- Hughes, I. (2001). But isn't this what you're paid for? the pros and cons of peer and self-assessment. *Planet*, 3(1), 20–23.
- Hutchings, P. (1992). The assessment movement and feminism: Connection or collision? In C. M. Musil (Ed.), *Students at the center: Feminist assessment* (pp. 17–28). Washington, DC: Association of American Colleges.
- Hyland, K., & Hyland, F. (2006). *Feedback in second language writing: Contexts and issues*. New York, NY: Cambridge University Press.

- Jacobs, L. C., & Chase, C. I. (1992). *Developing and using tests effectively: A guide for faculty*. San Francisco, CA: Jossey-Bass Publishers.
- James, M. E. (2013). *Educational assessment, evaluation and research: The selected works of Mary E. James*. London: Routledge.
- Jimoyiannis, A., & Tsiotakis, P. (2016). Self-directed learning in e-portfolios: Analysing students' performance and learning presence. *EAI Endorsed Transactions on e-Learning*, 3(9), 1-9.
- Joint Advisory Committee, Centre for Research in Applied Measurement and Evaluation (1993). *Principles for fair student assessment practices for education in Canada*. Edmonton, AB: Joint Advisory Committee. Retrieved from: <https://www.wcdsb.ca/wp-content/uploads/sites/36/2017/03/fairstudent.pdf>
- Joint Committee on Standards for Educational Evaluation (A. R. Gullickson, Ed.) (2003). *The student evaluation standards: How to improve evaluations of students*. Thousand Oaks, CA: Corwin Press.
- Kaufman, D. (2004). Constructivist issues in language learning and teaching. *Annual Review of Applied Linguistics*, 24, 303–319.
- Kenway, J. & Modra, H. (1992) Feminist pedagogy and emancipatory possibilities. In C. Luke & J. Gore (Eds.), *Feminisms and critical pedagogy* (pp. 138–166). New York, NY: Routledge.
- Keutzer, C. S. (1993). Midterm evaluation of teaching provides helpful feedback to instructors. *Teaching of Psychology*, 20(4), 238–240.

- Keyssar, H. (1985). Staging the feminist classroom: A theatrical model. In M. Culley & C. Portuges (Eds.), *Gendered subjects: The dynamics of feminist teaching* (pp. 108–123). Boston, MA: Routledge.
- Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28–37.
- Klenowski, V., Askew, S., & Carnell, E. (2006). Portfolios for learning, assessment and professional development in higher education. *Assessment & Evaluation in Higher Education*, 31(3), 267–287.
- Knight, P. (2002). The Achilles' heel of quality: The assessment of student learning. *Quality in Higher Education*, 8(1), 107–115.
- Knight, P. (2007). Grading, classifying and future learning. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 72–86). London, UK: Routledge.
- Knight, P., & Yorke, M. (2003). *Assessment, learning and employability*. Maidenhead, UK: Open University Press.
- Kocadere, S. A., & Ozgen, D. (2012). Assessment of basic design course in terms of constructivist learning theory. *Procedia—Social and Behavioral Sciences*, 51, 115–119.
- Koh, A., (2017). Feminist pedagogy in the digital age: Experimenting between MOOCs and DOCCs. In E. Losh (Ed.), *MOOCs and their afterlives: Experiments in scale and access in higher education* (pp.123–134). Chicago: University of Chicago Press.

- Kohn, A. (2002). The Dangerous Myth of Grade Inflation. *Chronicle of Higher Education* 49(11),1–10.
- Kohn, A. (2013). *No contest: The case against competition*. Boston and New York: Houghton Mifflin Company.
- Jezak, M., & Piccardo, E. (2017). Introduction: The Canadian Language Benchmarks and Niveaux de compétence linguistique canadiens – Canadian Language Framework in the Era of Glocalization. In M. Jezak (Ed.), *Language is the key: A Canadian language benchmarks model*. Ottawa: University of Ottawa Press.
- Kvale, S. (2007). Contradictions of assessment for learning in institutions of higher learning. In D. Boud, & N. Falchikov (Eds.), *Rethinking assessment in higher education: learning for the longer term* (pp. 57–71). London, UK: Routledge.
- Lenin, V. I. (1965). Party organization and party works. In *Lenin Collected Works: Vol. 10* (pp. 44-49). Moscow: Progress Publishers. (Original work published 1905)
- Lew, M. D. N., Alwis, W. A. M., & Schmidt, H. G. (2010). Accuracy of students' self-assessment and their beliefs about its utility. *Assessment & Evaluation in Higher Education*, 35(2), 135-156.
- Lewis, H. R. (2006). *Excellence without a soul: How a great university forgot education*. New York, NY: PublicAffairs.
- Lewis, M. (1992). Interrupting patriarchy: Politics, resistance and transformation in the feminist classroom. In C. Luke & J. Gore (Eds.), *Feminisms and critical pedagogy* (pp. 167–191). New York, NY: Routledge.
- Lipnevich, A. A., & Smith, J.K. (2009). Effects of differential feedback on students' examination performance, *Journal of Experimental Psychology*, 15(4), 319–333.

- Llewellyn K. R., & Llewellyn, J. (2015). A restorative approach to learning: Relational theory as feminist pedagogy in universities. In T. P. Light, J. Nicholas & R. Bondy (Eds.), *Feminist pedagogy in higher education: Critical theory and practice* (pp.11–31). Waterloo, ON: Wilfrid Laurier University Press.
- Lloyd, S., Warner, R., & Baber, K., & Sollie, D. (2009). Activism in the academy: Constructing/negotiating feminist leadership. In S. A. Lloyd & A. L. Few, & K. R. Allen (Eds.), *Handbook of feminist family studies* (pp. 292–303). Thousand Oaks, CA: SAGE.
- Luke, C. (1992). Feminist politics in radical pedagogy. In C. Luke & J. Gore (Eds.), *Feminisms and critical pedagogy* (pp. 25–53). New York, NY: Routledge.
- Luke, C. (1996). Feminist pedagogy theory: Reflections on power and authority. *Educational Theory*, 46(3), 283–302.
- Luke, C., & Gore, J. (1996) Women in the academy: Strategy, struggle, survival. In C. Luke & J. Gore (Eds.), *Feminisms and critical pedagogy* (pp. 192–210). New York, NY: Routledge.
- Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 351–372.
- Lynch, R., & Hennessy, J. (2017). Learning to earn? The role of performance grades in higher education. *Studies in Higher Education*, 42(9), 1750-1763.
- MacNell, L., Driscoll, A., & Hunt, A. N., (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, 40, 291–303.
- Magnusson, E., & Marecek, J. (2015). *Doing interview-based qualitative research: A learner's guide*. New York, NY: Cambridge University Press.

- Maher, F. A., & Tetreault, M. K. (1994). *The feminist classroom: An inside look at how professors and students are transforming higher education for a diverse society*. New York, NY: Basic Books.
- Malouff, J. M., & Thorsteinsson, E. B. (2016). Bias in grading: A meta-analysis of experimental research findings. *Australian Journal of Education, 60*(3), 245–256.
- Manicom, A. (1992). Feminist pedagogy: Transformations, standpoints, and politics. *Canadian Journal of Education, 17*(3), 365–389.
- Manning, D. H., Ference, K. A., Welch, A. C., & Holt-Macey, M. (2016). Development and implementation of pass/fail grading system for advanced pharmacy practice experiences. *Currents in Pharmacy Teaching and Learning, 8*(1), 59-68.
- Marx, K. (2011). *Capital, A critique of political economy: Volume 1* (S. Moore & E. Aveling, Trans.). New York, NY: Dover Publications. (Original work published 1867)
- Maslow, A. H. (1970). *Motivation and personality* (2nd ed.). New York, NY: Harper & Row.
- Matejka, K., & Kurke, L. B. (1994). Designing a great syllabus. *College Teaching, 42*(3), 115–117.
- McCabe, D. L., Butterfield, K. D., & Trevino, L. K. (2012). *Cheating in college: Why students do it and what educators can do about it*. Baltimore, MD: The Johns Hopkins University Press.
- McClain, L., Gulbis, A., & Hays, D. (2018). Honesty on student evaluations of teaching: Effectiveness, purpose, and timing matter. *Assessment & Evaluation in Higher Education, 43*(3), 369-385.

- McClure, L. (2000). Feminist pedagogy and the classics. *Classical World*, 94(1), 53–55.
- McCroskey, J., & Richmond, V. (1983). Power in the classroom I: Teacher and student perceptions. *Communication Education*, 32(2), 175–184.
- McMahon, J. (2018). The pedagogy of feminist theory. *Politics & Gender*, 14(2), E10.
- Mendenhall, M. E., & Marsh, W. J. (2010). Voices from the past: Mary Parker Follett and Joseph Smith on collaborative leadership. *Journal of Management Inquiry*, 19(4), 284–303.
- Merriam, S. B. (1988). *Case study research in education: A qualitative approach*. San Francisco, CA: Jossey-Bass.
- Merriam, S. B., & Brockett, R. G. (1997). *The profession and practice of adult education: An introduction*. San Francisco, CA: Jossey-Bass.
- Mikell, G., Hawthorne, S., Moghadam, V., & Das Dasgupta, S. (2007). "Feminism." *Encyclopedia of Sex and Gender*, edited by Fedwa Malti-Douglas, vol. 2, Macmillan Reference USA, 2007, pp. 531-542.
- Miller, A., Imrie, B. W., & Cox, K. (1998). *Student assessment in higher education: A handbook for assessing performance*. London: Routledge.
- Mislevy, R. (2017) On measurement in educational assessment. In C. Secolsky & D. B. Denison (Eds.), *Handbook on measurement, assessment, and evaluation in higher education* (2nd ed., pp. 11–31). New York, NY: Routledge.  
<https://doi.org/10.4324/9781315709307>
- Monchinski, T. (2010). *Education in hope: Critical pedagogies and the ethic of care*. New York, NY: Peter Lang.

- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(41), 16474–16479.
- Murdock, T. B., Stephens, J. M., & Grotewiel, M. M. (2016). Student dishonesty in the face of assessment: Who, why and what we can do about it. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment* (pp. 186–203). New York, NY: Routledge.
- Musil, C. M. (1992). Relaxing your neck muscles: The history project. In C. M. Musil (Ed.), *Students at the center: Feminist assessment* (pp. 3–16). Washington DC: Association of American Colleges.
- Myers, S. A. (2017). A longitudinal analysis of students' motives for communicating with their instructors. *Communication Education*, *66*(4), 467–473.
- Nawratil, G. (1999). Implications of computer-conferenced learning for feminist pedagogy and women's studies: A review of the literature. *Resources for Feminist Research*, *27*(1/2), 73–107.
- Nelson L. H. (2015). Epistemological communities. In L. Alcoff & E. Potter (Eds.), *Feminist epistemologies* (pp.121–159). New York: Routledge.
- New College of Florida. (n.d.) Academics at New College. Retrieved from <https://www.ncf.edu/academics/academics-at-new-college/>
- Newson, J. (1998). The corporate-linked university: From social project to market force. *Canadian Journal of Communication*, *23*(1), 107–124.

- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.
- Noddings, N. (2013). *Caring: A feminine approach to ethics and moral education* (2nd ed.). Berkeley, CA: University of California Press.
- Noddings, N. (2015). *Philosophy of education* (4th edition). Boulder, CO: Westview Press.
- Norton, P., & Wiburg, K. M. (1998). *Teaching with technology*. Fort Worth, TX: Harcourt Brace College.
- Novek, E. (1999). Service-learning is a feminist issue: Transforming communication pedagogy. *Women's Studies in Communication, 2*(2), 230–240.
- Oakley, A. (1981). Interviewing women: A contradiction in terms? In H. Roberts (Ed.), *Doing feminist research* (pp. 30–61). London, UK: Routledge.
- Orange, A. (2016). Encouraging reflexive practices in doctoral students through research journals. *The Qualitative Report, 21*(12), 2176–2190.
- Ortlipp, M. (2008). Keeping and using reflective journals in the qualitative research process. *The Qualitative Report, 13*(4), 695–705.
- Palincsar, A. S. (2005). Social constructivist perspectives on teaching and learning. In H. Daniels (Ed.), *An introduction to Vygotsky* (2nd ed., pp. 279–308). New York, NY: Routledge.
- Palmer, P. J., & Ebooks Corporation. (2007). *The courage to teach: Exploring the inner landscape of a teacher's life* (2nd ed.). San Francisco, CA: Jossey-Bass.

- Pardjono, P. (2016). Active learning: The Dewey, Piaget, Vygotsky, and constructivist theory perspectives. *Jurnal Ilmu Pendidikan*, 9(3), 163-178.
- Panadero, E. (2016). Is it safe? Social, interpersonal, and human effects of peer assessment: A review and future directions. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment* (pp. 247–266). New York: Routledge.
- Pavlov, I. P., Gantt, W. H., & Fol'bert, G. V. (1928). *Lectures on conditioned reflexes*. New York: International Publishers.
- Personal. (2003). In *Merriam-Webster's Dictionary* (11th ed.). Springfield, MA: Merriam-Webster.
- Piaget, J., & Inhelder, B. (1973). *The psychology of the child*. London: Routledge & K. Paul.
- Ponterotto, J. G. (2006). Brief note on the origins, evolution, and meaning of the qualitative research concept thick description. *The Qualitative Report*, 11(3), 538–549.
- Pulfrey, C., Buchs, C., & Butera, F. (2011). Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology*, 103(3), 683–700.
- Raaper, R. (2019). Assessment Policy and “Pockets of Freedom” in a Neoliberal University: A Foucauldian Perspective. In: C. Manathunga & D. Bottrell (Eds.) *Resisting Neoliberalism in Higher Education Volume II*.(pp.155-175). Cham: Springer International Publishing.

- Race, P. Brown, S., & Smith, B. (2005). *500 Tips on assessment* (2nd ed.). London, UK: Routledge.
- Rauschenberg, S. (2014). How consistent are course grades? An examination of differential grading. *Education Policy Analysis Archives*, 22(92).  
<https://doi.org/10.14507/epaa.v22n92.2014>
- Reynolds, M., & Trehan, K. (2000). Assessment: A critical perspective. *Studies in Higher Education*, 25(3), 267–278.
- Rideout, C. A. (2018). Students' choices and achievement in large undergraduate classes using a novel flexible assessment approach. *Assessment & Evaluation in Higher Education*, 43(1), 68–78.
- Ron, A. (2013). Power. In G. Claeys (Ed.), *Encyclopedia of modern political thought* (p. 648). Thousand Oaks, CA: SAGE
- Ropers-Huilman, R. (2009). Scholarship on the other side: Power and caring in feminist education. In R. Crabtree, D. Sapp, & A. Licona (Eds.). *Feminist pedagogy : Looking back to move forward* (40–56). Baltimore, MD : The Johns Hopkins University Press.
- Ropers-Huilman, R., & Winters, K. T. (2011). Feminist research in higher education. *The Journal of Higher Education*, 82(6), 667–690.
- Rupiper Taggart, A., & Laughlin, M. (2017). Affect matters: When writing feedback leads to negative feeling. *International Journal for the Scholarship of Teaching and Learning*, 11(2), Article 13. <https://doi.org/10.20429/ijsotl.2017.110213>
- Sadler, R. D. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, 67(3), 273–288.

- Salkever, K. (2002) Feminist assessment. In A. M. Martínez Alemán & K. A. Renn (Eds.), *Women in higher education: An encyclopedia* (pp. 85–88). Santa Barbara, CA: ABC-CLIO.
- Sánchez-Casal, S. (2002) Unleashing the Demons of History. In A. A. Macdonald & S. Sánchez-Casal (Eds.), *Twenty-first-century feminist classrooms: Pedagogies of identity and difference* (pp. 59–85). New York, NY: Palgrave Macmillan.
- Sánchez-Casal, S., & Macdonald, A. A. (2002). Introduction. In A. A. Macdonald & S. Sánchez-Casal (Eds.), *Twenty-first-century feminist classrooms: Pedagogies of identity and difference* (pp. 1–28). New York, NY: Palgrave Macmillan.
- Sandell, R. (1991). The liberating relevance of feminist pedagogy. *Studies in Art Education*, 32(3), 178–187.
- Savin-Baden, M., & Major, C. H. (2010 2013?). The uncertainty of wisdom. In M. Savin-Baden & C. H. Major (Eds.), *Qualitative research: The essential guide to theory and practice* (pp. 1–5). New York, NY: Routledge.
- Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching*, 5(9), 9–16.
- Segers, M., Dochy, F., & Cascallar, E. (2003). The era of assessment engineering: Changing perspectives on teaching and learning and the role of new modes of assessment. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 1–12). Dordrecht, Netherlands: Kluwer Academic.

- Shackelford, J. (1992). Feminist pedagogy: A means for bringing critical thinking and creativity to the economics classroom. *American Economic Review* 82(2), 570–576.
- Shackelford, J. (2003). Feminist pedagogy and economics. In G. M. Hoyt & K. McGoldrick (Eds.), *International handbook on teaching and learning economics* (pp. 205–213). Northampton, MA: Edward Elgar.
- Shapiro, J. P. (1992). What is feminist assessment? In C. M. Musil (Ed.), *Students at the center: Feminist assessment* (pp. 29–38). Washington, DC: Association of American Colleges.
- Shih, J. C., Ing, M., & Tarr, J. E. (2013). How many classroom observations are sufficient? Empirical findings in the context of a longitudinal study. *Middle Grades Research Journal*, 8(2), 41–49.
- Scaife, J., & Wellington, J. (2010). Varying perspectives and practices in formative and diagnostic assessment: A case study. *Journal of Education for Teaching*, 36(2), 137–151.
- Schniedewind, N. (1981). Feminist values: Guidelines for teaching methodology in Women's Studies. *The Radical Teacher*, 18, 25–28.
- Schinske, J., & Tanner, K. (2014). Teaching more by grading less (or differently). *CBE Life Sciences Education*, 13(2), 159–166.
- Schuetze, H.G. (2013) Canada. In C.J Russo, & J.D. Panzer (Eds.), *Handbook of comparative higher education law* (pp. 63-86) . Portland: Ringgold, Inc.
- Shrewsbury, C. M. (1993). What is feminist pedagogy? *Women's Studies Quarterly*, 21(3/4), 8–16.

- Skinner, B. F. (1968). *The technology of teaching*. New York, NY: Appleton-Century-Crofts.
- Smith, B., & Sparkes, A. C. (2016). Interviews: Qualitative interviewing in the sport and exercise sciences. In B. Smith, & A.C. Sparkes (Eds.), *Routledge handbook of qualitative research in sport and exercise* (pp. 103-123). London, UK: Routledge.
- Smyth, J. (1989). Developing and sustaining critical reflection in teacher education. *Journal of Teacher Education*, 40(2), 2–9.
- Soeiro, D., de Figueiredo, A. D., & Ferreira, J. A. G. (2011, October). Student empowerment in higher education through participatory evaluation. Paper presented at the *Frontiers in Education Conference (FIE)*, Rapid City, SD.
- Soffer, T., Kahan, T., & Livne, E. (2016) E-assessment of online academic courses via students' activities and perceptions. *Studies in Educational Evaluation*, 54, 83-93. <https://doi.org/10.1016/j.stueduc.2016.10.001>
- Sparkes, A. C., & Smith, B. (2014). *Qualitative research methods in sport, exercise and health: From process to product*. London, UK: Routledge.
- Spradley, J. (1980). *Participant observer*. New York: Holt, Rinehart and Winston.
- Sprague, J., & Massoni, K. (2005). Student evaluations and gendered expectations: What we can't count can hurt us. *Sex Roles*, 53(11/12), 779–793.
- Stake, R. (1994). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236–247). London, UK: SAGE.
- Stanny, C., Gonzalez, M., & McGowan, B. (2015). Assessing the culture of teaching and learning through a syllabus review. *Assessment & Evaluation in Higher Education*, 40(7), 898–913.

- Sullivan, L. E. (Ed.) (2009). *The SAGE glossary of the social and behavioral sciences*. Thousand Oaks, CA: SAGE. <https://doi.org/10.4135/9781412972024>
- Swanborn, P. G. (2010). *Case study research: What, why and how?* Los Angeles, CA: SAGE.
- Tan, K. H. K. (2004). Does student self-assessment empower or discipline students? *Assessment & Evaluation in Higher Education*, 29(6), 651–662.
- Tatum, H. E., Schwartz, B. M., Schimmoeller, P. A., & Perry, N. (2013). Classroom participation and student-faculty interactions: Does gender matter? *The Journal of Higher Education*, 84(6), 745–768.
- Taylor, J. W., Atwood, C. H., & Hutchings, P. A. (2000). Why are chemists and other scientists afraid of the peer review of teaching? *Journal of Chemical Education*, 77(2), 239–243.
- TenHouten, W. D., & Taylor & Francis eBooks A-Z. (2017). *Alienation and affect*. New York: Routledge, Taylor & Francis Group.
- Therborn, G. (2004). *Between sex and power: Family in the world, 1900-2000*. London: Routledge.
- Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse and structure. *Qualitative Inquiry*, 17(6), 511–521.
- Thompson, A., & Gitlin, A. D. (1995). Creating spaces for reconstructing knowledge in feminist pedagogy. *Educational Theory*, 45(2), 125–150.
- Thompson, B. (2007). The syllabus as a communication document: Constructing and presenting the syllabus. *Communication Education*, 56(1), 54–71.

- Thomson, D. (2001). *Radical feminism today*. London: SAGE Publications Ltd.
- Topping, K. J. (2012). Peers as a source of formative and summative assessment. In J. H. McMillan (Ed.), *SAGE Handbook of research on classroom assessment* (pp. 395–412). Los Angeles, CA: SAGE.
- University of Toronto (2012). *University Assessment and Grading Practices Policy*. Retrieved from <http://www.governingcouncil.utoronto.ca/policies/uniassgpp.htm>
- Ussher, B., & Earl, K. (2010). ‘Summative’ and ‘formative’: Confused by the assessment terms? *New Zealand Journal of Teachers’ Work*, 7(1), 53–63.
- Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22-42.
- van Aalst, J., & Chan, C. K. K. (2007). Student-directed assessment of knowledge building using electronic portfolios. *Journal of the Learning Sciences*, 16(2), 175–220.
- van den Brink, M., & Benschop, Y. (2012). Slaying the seven-headed dragon: The quest for gender change in academia. *Gender, Work & Organization*, 19(1), 71–92.
- Vogel, L. (2018). Medical school admission requirements lock out many indigenous students. *Canadian Medical Association Journal*, 190(26), E809.
- Vygotsky, L. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Wagner, M. L., Suh, D. C., & Cruz, S. (2011). Peer- and self-grading compared to faculty grading. *American Journal of Pharmaceutical Education*, 75(7), 130-130.

- Wall, A. F., Hursh, D., & Rodgers, J. W., III. (2014). Assessment for whom: Repositioning higher education assessment as an ethical and value-focused social practice. *Research & Practice in Assessment, 9*(1), 5.
- Wallace, S. (Ed.) (2015). *A dictionary of education* (2nd ed.). Oxford: Oxford University Press.
- Ware, C. (1970). *Woman Power: The Movement for Women's Liberation*. New York, NY: Tower Publications.
- Watkins, C. (2003). *Learning: A Sense-Maker's Guide*. London: Association of Teachers and Lecturers .
- Watkins, S. A., Rueda, M., & Rodriguez, M. (R. Appignanesi, Ed.) (1999). *Introducing feminism*. Cambridge, UK: Icon Books.
- Webb, L. M., Allen, M. W., & Walker, K. L. (2002). Feminist pedagogy: Identifying basic principles. *Academic Exchange Quarterly, 6*(1), 67–72.
- Weiler, K. (1991). Freire and a feminist pedagogy of difference. *Harvard Education Review, 61*(4), 449–475.
- Weiler, K. (1995). Revisioning feminist pedagogy. *NWSA Journal, 7*(2), 100–106.
- Weiler, K. (2001). Rereading Paulo Freire. In K. Weiler (Ed.), *Feminist engagements: Reading, resisting, and revisioning male theorists in education and cultural studies* (pp. 67–87). New York, NY: Routledge.
- Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco, CA: Jossey-Bass.
- Woodlock, C. (1995). Looking at feminist pedagogies: What is seen in the literature and what is seen in an art education studio/classroom. *Marilyn Zurmuehlen Working*

*Papers in Art Education*, 13(1), 4–10. Retrieved from

<http://ir.uiowa.edu/mzwp/vol13/iss1/3>

Xerri, M. J., Radford, K., & Shacklock, K. (2018). Student engagement in academic activities: A social support perspective. *Higher Education*, 75(4), 589–605.

Xu, J., Lo, A., & Wu, J. (2018). Are students customers? Tourism and hospitality students' evaluation of their higher education experience. *Journal of Teaching in Travel & Tourism*, 18(3), 236–258.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45(4), 477–501

Yorke, M. (2005). Formative assessment in higher education: Its significance for employability, and steps towards its enhancement. *Tertiary Education and Management*, 11(3), 219–238.

Yorke, M. (2008). *Grading student achievement in higher education: Signals and shortcomings*. New York, NY: Routledge.

Yorke, M. (2011). Summative assessment: Dealing with the 'measurement fallacy'. *Studies in Higher Education*, 36(3), 251-273.

## Appendix A

### Letter of Initial Contact

Dear All,

For my doctoral dissertation at the University of Victoria, I am planning to conduct a study of assessment practices in university. I need six participants who are willing to take part in one individual interview, and I am writing to invite you to consider participating.

Participating in my project should not require much of your time. The interview will take approximately one hour. During the interview, I will ask you questions related to your assessment practices.

I will be happy to provide you with any information that may help you decide whether to participate or not. We can talk by phone or in person. I can provide further details via email as well. Do not hesitate to ask me any questions.

I hope to hear from you soon,

Kindest regards,

Veta Chitnev

Lecturer of Russian

Department of Modern Languages

Tel: xxx-xxx-xxxx

## **Appendix B**

### Consent Form

#### University Assessment Practices through the Lens of Feminist Pedagogy

Who is conducting the study?

Principal Investigator: Veta Chitnev

Supervisor: Dr. Kathy Sanford, UVic, Faculty of Education,

The research is part of a doctoral dissertation.

Why should you take part in this study?

This research is part of my doctoral dissertation at the University of Victoria. The study will examine some assessment practices being used in the university today, and how they are compatible with feminist pedagogy. Your participation will help fill the gap in the theory and practice of feminist pedagogy.

How is the study done?

If you say 'Yes,' here is how the study will be done. You will have one individual interview where you will be asked questions related to your assessment practices. I will email you the questions in advance, and I will make sure that your interview is scheduled at a time and location convenient to you. It is estimated that the interview will take approximately one hour. With your permission, the interview will be audio recorded, and transcriptions will be made. You will have an opportunity to review the full transcript of

the original interview and to make any changes you may feel are needed. If you do not wish to have the interview audio recorded, your request will be accommodated.

#### Dissemination of the results and future use of the data

The study is part of a doctoral dissertation in Education, and once submitted for the degree, it will exist in the public domain. The data and study results may be used by researchers for academic publications and for educational purposes. It is also anticipated that the results will be shared with other faculty members at one of the departmental colloquia.

#### Are there any potential risks to you in participating in the study?

There are no known risks and/or discomforts associated with this study. If during the interview any of the questions will seem sensitive, you do not have to answer these questions if you do not want to.

#### What are the benefits of participating?

There may or may not be direct benefits to you from taking part in this study. However, in the future others may benefit from this study if alternatives to patriarchal forms of assessment strategies are identified.

#### How will your identity be protected?

If you decide to participate in this study, your identity will be kept strictly confidential. Only Veta Chitnev will have access to the original data. The signed informed consent forms will be kept in a secure location with restricted access only to Veta Chitnev, while the recording and transcripts will be stored in her password protected computer. If the results are published, it will not be possible to match any individual response with any

data published as a pseudonym of your choice will be used instead of your real name.

In addition, the courses you teach and your personal characteristics will be masked.

Who can you contact if you have questions about the study?

If you have questions about the study, please contact me at xxxxx@xxxx.xx I expect to complete the research by the end of December, and if you are interested in the findings, do not hesitate to contact me. We would be happy to share the findings with you.

Who can you contact if you have complaints or concerns about the study?

If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, contact Research Ethics at tel.xxx-xxx-xxxx or if long distance e-mail xxxx@xxxx.xx.

#### Participant Consent and Signature Page

Taking part in this study is entirely up to you. You have the right to refuse to participate in this study. If you decide to take part, you may choose to pull out of the study at any time without giving a reason and without any negative impact on your work relationships. By submitting this questionnaire, you are consenting to participate in this research study. Your signature below indicates that you have received a copy of this consent form for your own records.

Your signature indicates that you consent to participate in this study.

---

Participant Signature	Date
-----------------------	------

---

Printed Name of the Participant signing above

## Appendix C

### Obtaining Consent Email

Dear [insert name],

I have left a consent form in your mailbox. Please read it, and if you agree to participate in the study, sign it. If you change your mind, you can refuse to participate in my study without giving a reason and without any negative impact on our work relationships. If you are still interested in participating, please leave the signed copy of the consent form in my mailbox in the envelope provided or return it to me in person, whichever is more convenient.

Kindest regards,

Veta Chitnev

Lecturer of Russian

Department of Modern Languages

Tel: xxx-xxx-xxxx

## Appendix D

### Interview Questions

1. What are your perspectives on grading?
2. Administration discourages instructors from high class average grades. This means you are supposed to keep in mind the class average grade when you are grading. Do you agree with this university approach?
3. Do you offer any assignments that are assessed as ‘complete/incomplete’ or ‘pass/fail’ method?
4. Do you consider ‘complete/incomplete’ or ‘pass/fail’ without assigning a grade value to be a fair method of assessment?
5. Can you share with me any thoughts on how a teacher can provide more feedback and do less grading?
6. Is there an assessment method in your teaching practices that you find particularly effective and interesting?
7. What kind of feedback do you find to be most useful in your practice?
8. Formative assessment is assessment performed during instruction, and it aims to form students’ competence by providing them with appropriate feedback. Usually formative assessment does not involve grading. Do you use any formative assessment tools? What formative assessment tools do you find the most effective?

9. Do you use diagnostic assessment in your practice? If yes, please elaborate.
10. Do you use peer/self/group assessment in your classes? If yes, please tell me how you use these methods of assessment.
11. How do you assess participation?
12. How do you assess presentations?
13. Does students' gender impact the ways in which you provide feedback or grades?
14. According to research, providing detailed and timely feedback in large classes can be challenging. How do you assess learning in your class? <sup>1</sup>

---

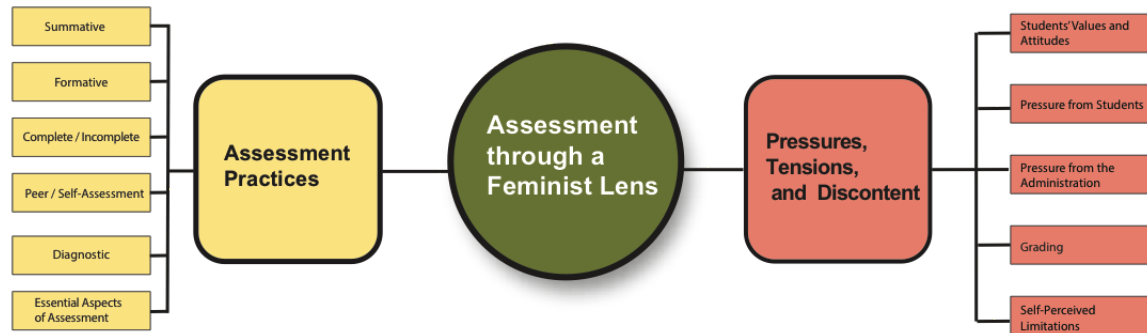
<sup>1</sup> Question N 14 is an additional question given to three instructors who were teaching large classes.

**Appendix E**

## Interview Schedule

Participants	Date	Length in minutes
Participant N 1	February 28	56
Participant N 2	March 16	29
Participant N 3	March 16	40
Participant N 4	March 22	39
Participant N 5	March 31	15
Participant N 6	May 1	22

## Appendix F



*Figure 3.1* Assessment thematic map 1

## Appendix G

### Russian Course Survey

1. Do you find this course manageable?
2. What did you do to make this course more manageable? (Please provide some advice to you fellow students and future students)
3. What activities do you find useful?
4. What activities do you not find useful?
5. What activities do you find too difficult?
6. Do you find group activities or working with your partner useful for your Russian language proficiency?
7. Your opinion on the current time allocation in class. Please choose one of the options:
  - A. The use of time is OK
  - B. I would prefer less grammar explanation in class since all the material is in the textbook, but I would prefer to use this time for speaking Russian.
  - C. I would prefer to spend less time on cultural material and to practice grammar more.
  - D. I suggest
8. Your suggestions on how we can improve this course to make it less stressful, more manageable and enjoyable.