

## **NOTE TO USERS**

**This reproduction is the best copy available.**

UMI<sup>®</sup>





uOttawa

L'Université canadienne  
Canada's university

**FACULTÉ DES ÉTUDES SUPÉRIEURES  
ET POSTDOCTORALES**



**uOttawa**

L'Université canadienne  
Canada's university

**FACULTY OF GRADUATE AND  
POSTDOCTORAL STUDIES**

**Robin D. Tierney**

-----  
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

**Ph.D. (Education)**

-----  
GRADE / DEGREE

**Faculty of Education**

-----  
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

**Insights into Fairness in Classroom Assessment:  
Experienced English Teachers Share their Practical Wisdom**

-----  
TITRE DE LA THÈSE / TITLE OF THESIS

**Marielle Simon**

-----  
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

-----  
CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

**Martin Barlosky**

**Marie-Josée Berger**

**Catherine Brighton (University of  
Virginia)**

**Christine Suurtamm**

**Gary W. Slater**

-----  
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

**Insights into Fairness in Classroom Assessment:  
Experienced English Teachers Share their Practical Wisdom**

Robin D. Tierney

Thesis submitted to the  
Faculty of Graduate and Postdoctoral Studies  
In partial fulfillment of the requirements  
For the Ph.D. degree in Teaching, Learning & Evaluation,

Faculty of Education  
University of Ottawa

© Robin D. Tierney, Ottawa, Canada, 2010



Library and Archives  
Canada

Published Heritage  
Branch

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque et  
Archives Canada

Direction du  
Patrimoine de l'édition

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file* *Votre référence*  
ISBN: 978-0-494-69109-0  
*Our file* *Notre référence*  
ISBN: 978-0-494-69109-0

**NOTICE:**

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

**AVIS:**

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

■❖■  
**Canada**

## Acknowledgments

Although I am but one voice in writing this dissertation, I have benefitted from the generosity of many. I owe special thanks to the teachers who shared their practical wisdom with me. Ethical protocol dictates that I cannot name you, but your participation made my study possible. Thanks to my thesis supervisor, Marielle Simon, for her flexibility in allowing me to pursue my interest in the ethics of classroom assessment, and for encouraging me to participate in the educational research community. Thanks to my committee members, Martin Barlosky, Marie Josée Berger, and Christine Suurtamm, for their valuable feedback. My work is stronger as a result. Thanks to the researchers, Bronwen Cowie, Susan Green, and Jim McMillan, who took time to corresponded from afar. Thanks to my colleagues, Doug Archibald, Julie Charland, Pat Holloway and Paula Kolodzie Moffat, for the feedback on my questionnaires. A big thanks to my Pips friends for the support along the way. Our conversations made perseverance so much more fun. Thanks to the Social Sciences and Humanities Research Council of Canada, the Ontario Graduate Scholarship Program, and the Faculty of Graduate and Postdoctoral Studies at the University of Ottawa for the financial support that facilitated my studies. Thanks most of all to my husband, David Allan, for your love and care. The next road trip is your pick. In the end, thanks to my family because in the end it is about us.

## **Dedication**

This work is dedicated to those who  
question  
the ethics of our daily doings.

mono no aware

jaba

## **Abstract**

Fairness is a perennial ideal in educational assessment. For most of the 20<sup>th</sup> century, objective tests were viewed as a means of bypassing teachers' inconsistencies in grading, and fairness was defined in technical terms for the development and administration of tests. As the 21<sup>st</sup> century opened, the limitations of standardized tests for supporting student learning had become apparent, and interest in the pedagogical potential of classroom assessment was renewed. Although teachers' assessment decisions can have profound, long-term effects for students, most of the guidance for fair assessment still focuses on standardized testing. This study was motivated by the need for a better understanding of fairness in the classroom environment where the primary purpose of assessment is ideally to support learning. I approached my research from the perspective of critical pragmatism, which assumes that diverse forms of knowledge are valuable, including ethical and practical knowledge. I drew on the practical wisdom (phronesis) of experienced teachers to understand fairness specifically in the assessment of writing in English Language Arts in Grades 7 to 12 in Ontario. Practical wisdom is a responsive network that combines theoretical and technical knowledge, moral beliefs and ethical knowledge, experience, personal characteristics, and an understanding of the particulars of practice. I began my research by immersing myself in the literature, investigating the educational context, tailoring my research tools, and purposefully selecting teachers to participate based on their education and experience. Six participants responded in writing to vignettes that illustrated fairness dilemmas, and we discussed their recommendations for fair assessment during follow-up interviews. Insight that emerged from this process is that some strategies contribute to the fairness of classroom assessment, such as

developing a constructive learning environment, sharing expectations and criteria, and offering students relevant opportunities to learn and demonstrate their learning. At the crux, however, fairness rests in a teacher's ability to understand students and their learning, and to reflect about the ongoing interactions and decisions involved in the use of assessment for learning.

## Table of Contents

Abstract .....	iv
Table of Contents .....	vi
List of Tables .....	xi
List of Figures .....	xii
CHAPTER ONE: INTRODUCTION .....	1
Background of the Problem .....	1
Rationale for Turning to Teachers .....	7
Research Purpose and Questions .....	8
Researcher's Perspective and Goals .....	10
Definition of Terms .....	13
Organization of the Thesis .....	16
CHAPTER TWO: ASSESSMENT IN THE CLASSROOM ENVIRONMENT .....	17
Evolution of a Continuum of Assessment Purposes .....	17
The Concept of Assessment for Learning .....	20
Issues in the Use of Classroom Assessment for Multiple Purposes .....	26
Influence of External Factors on Classroom Assessment .....	31
Classroom Level Factors .....	31
School Level Factors .....	36
System Level Factors .....	38
Chapter Summary .....	48
CHAPTER THREE: FAIRNESS IN CLASSROOM ASSESSMENT .....	50
Conceptual Work on Fairness in Classroom Assessment .....	51
Principles and Standards for Fair Classroom Assessment .....	51
Fairness in Classroom Assessment Theory .....	56
Empirical Studies Relating to Fair Classroom Assessment .....	64
Fair Classroom Assessment in Studies on the Ethics of Teaching .....	64
Fairness as an Emerging Issue in Classroom Assessment Research .....	69
Research Targeting Fairness, Equity or Ethics in Classroom Assessment .....	74
Fairness in Writing Assessment in Ontario .....	83
Chapter Summary .....	87

CHAPTER FOUR: TURNING TO TEACHERS' PRACTICAL WISDOMS .....	89
Critical Pragmatism .....	89
Teachers' Thinking .....	91
Interests in Teachers' Practical Thinking .....	92
Conceptualizations of Teachers' Thinking .....	93
The Concept of Practical Wisdom .....	99
Teachers' Practical Wisdom in Ontario .....	106
Concepts and Questions Situated in the Framework .....	111
Fairness in an Assessment for Learning Interaction .....	111
A Teacher's Practical Wisdom about a Fairness Dilemma .....	113
External Factors Within an Educational Context .....	114
 CHAPTER FIVE: METHODOLOGY .....	 116
Research Design .....	116
Information Sources .....	117
Large-Scale Survey Data .....	117
Existing Documents .....	118
Teacher Participants .....	119
Tools for Gathering Information .....	120
Background Questionnaire .....	120
Vignette Questionnaire .....	122
Interview Protocols .....	129
Research Process .....	129
Gaining Access and Ethical Approval .....	129
Analyzing Secondary Data and Relevant Documents .....	130
Recruiting and Selecting Participants .....	132
Working with Participants .....	133
Analyzing and Interpreting the Information Gathered .....	135
Role of Narrative .....	135
Research Quality .....	138
 CHAPTER SIX: CASE BY CASE .....	 143
Kevin: The Same Hoops .....	143
Kevin's Portrait .....	143
Recommendations for Fair Classroom Assessment .....	149
Elements of Practical Wisdom .....	152

External Factors .....	153
Summary of Kevin's Case .....	155
Lorena: Who the Learner Is .....	156
Lorena's Portrait .....	156
Recommendations for Fair Classroom Assessment .....	161
Elements of Practical Wisdom .....	163
External Factors .....	165
Summary of Lorena's Case .....	167
York: As Much of the Picture as Possible .....	168
York's Portrait .....	168
Recommendations for Fair Classroom Assessment .....	172
Elements of Practical Wisdom .....	175
External Factors .....	176
Summary of York's Case .....	179
Tanar: Walk with the Student .....	179
Tanar's Portrait .....	179
Recommendations for Fair Classroom Assessment .....	183
Elements of Practical Wisdom .....	186
External Factors .....	188
Summary of Tanar's Case .....	189
Lucy: They've Got to Earn It .....	190
Lucy's Portrait .....	190
Recommendations for Fair Classroom Assessment .....	194
Elements of Practical Wisdom .....	196
External Factors .....	199
Summary of Lucy's Case .....	201
Amada: Never Settle for Mediocrity .....	202
Amada's Portrait .....	202
Recommendations for Fair Classroom Assessment .....	206
Elements of Practical Wisdom .....	209
External Factors .....	212
Summary of Amada's Case .....	213
CHAPTER SEVEN: CROSS-CASE RESULTS .....	215
Practical Wisdom about Fairness in Classroom Assessment .....	215
Offering Students Opportunities .....	216

Reflecting about Assessment Interactions .....	218
Sharing Learning Expectations and Assessment Criteria .....	219
Establishing a Constructive Environment .....	223
Maintaining Consistency in the Assessment Process .....	226
Ensuring the Appropriateness of an Assessment for Individual Learners .....	228
Elements of Practical Wisdom at Play .....	231
Knowledge about Learners, Content and Pedagogy .....	232
Teaching Experience .....	235
Personal Characteristics .....	237
Moral Beliefs and Ethical Knowledge .....	239
Understanding of Particular Contexts, Situations and Students .....	241
External Factors of Influence .....	244
Classroom Level Factors .....	244
School Level Factors .....	248
System Level Factors .....	250
 CHAPTER EIGHT: DISCUSSION .....	 256
Fairness in Classroom Assessment .....	256
Constructive Learning Environment .....	257
Relevant Opportunities with Purpose in Mind .....	258
Transparent Communication about Assessment .....	262
Appropriate Assessment for Individual Learners .....	265
Consistent Treatment for a Group of Students .....	269
Reflective Interaction in the Assessment for Learning Cycle .....	272
Teachers' Practical Wisdom about Fairness .....	277
Types of Knowledge at Play .....	277
The Particulars of Practice .....	279
The Moral Dimension .....	280
The Circularity of Experience .....	284
Elucidating Practical Wisdom about Fairness .....	285
The Influence of External Factors .....	287
Framework Factors .....	288
People Factors .....	291
 CHAPTER NINE: CONCLUSION .....	 299
Limitations and Benefits of the Research Project .....	299

Using Secondary Data .....	299
Using Teachers' Practical Wisdoms .....	300
Eliciting Information for Research .....	302
Contributing to a Dialogue .....	304
Practical Implications .....	304
Fair Classroom Assessment for Learning .....	304
Teacher Education and Professional Development .....	307
Recommendations for Further Research .....	310
Fairness as a Quality of Classroom Assessment .....	310
Development of Teachers' Practical Wisdom about Fair Assessment .....	313
External factors Impacting Fair Classroom Assessment .....	315
Closing Thoughts .....	316
REFERENCES .....	318
APPENDICES .....	343
Appendix A: Learning English in Ontario's Educational System: An Overview .....	344
Appendix B: Ensuring Fairness in <i>The Student Evaluation Standards</i> (2003) .....	345
Appendix C: Background Questionnaire .....	347
Appendix D. Blueprint for the Development of the Background Questionnaire .....	351
Appendix E: Vignette Questionnaire .....	354
Appendix F: Blueprint for the Development of the Vignette Questionnaire .....	357
Appendix G: Template for Individualized Interview Protocols .....	360
Appendix H: Email Recruitment Text .....	363
Appendix I: Information Letter .....	364
Appendix J: Consent Form .....	366

## List of Tables

Table 1. Purposes of Classroom Assessment on a Continuum .....	19
Table 2. Aspects, Components, Strategies for Fair Classroom Assessment .....	61
Table 3. Empirical Research on Fairness, Equity and Ethics in Classroom Assessment ..	75
Table 4. Initial Responses on the Ethics of Classroom Assessment .....	147

## List of Figures

Figure 1. Process model of formative assessment as a dual cycle .....	23
Figure 2. Visual illustration of the conceptual framework for this study .....	112

## CHAPTER ONE: INTRODUCTION

*The variability in the marks given for the same subject and to the same pupils by different instructors is so great as frequently to work real injustice to the students.*

Isidor Finklestein (1913, p.6)

*There are potential sources of inequity deeply embedded in traditional and reform processes for formative and summative assessment.*

Candia Morgan and AnneWatson (2002, p.103)

Concern has long been expressed about the fairness of educational assessment.

During most of the 20<sup>th</sup> century, fairness was defined in technical terms by measurement specialists for the development and use of psychological and educational tests. Although it has been recognized that a technical definition is no longer sufficient, very little research has focused on understanding fairness as a quality of the ongoing interactions involved in classroom assessment. With this problem in mind, I turned to six experienced teachers to understand fairness in the use of assessment to support learning, specifically in the context of assessing student writing in English Language Arts classrooms in Ontario.

### Background of the Problem

Early in the 20<sup>th</sup> century, teachers' judgements about student work were found to be highly varied and often inaccurate (Finklestein, 1913; Monroe, 1923; Rinsland, 1937). As objective tests seemed to bypass teachers' inconsistencies, they were hailed as a superior means of measuring student achievement (Monroe; Rinsland). In the subsequent decades, modern measurement specialists concentrated mainly on issues relating to the development of tests that would produce valid and reliable results under standardized conditions (Giordano, 2005; Shepard, 2006). By the 1970's, though, new insights in learning theory,

shifting social ideals, and political upheaval began to impact the field of education (Bredo, 2006; Giordano; Gipps, 1999; Shepard), and many of the psychometric assumptions and technical preoccupations of modern measurement were eventually questioned from both inside (e.g., Popham, 1993; Stiggins, 1991) and outside the field (e.g., Dunne, 1993; Willie, 1985). As the 20<sup>th</sup> century closed, it had become evident that standardized tests were limited as a means of supporting student learning, and interest in the possibilities of classroom assessment as an alternative had grown to the degree that it emerged as a distinct area of research (Brookhart, 2001, 2003; Stiggins).

The dramatic rise in the amount of research on classroom assessment brought forth new understandings for the beginning of a new century, three of which are of particular note here. First, the pedagogical potential of classroom assessment became apparent (Black & Wiliam, 1998; Shepard, 1995), fueling considerable interest in the use of formative assessment to support student learning (e.g., Bell & Cowie, 2001; Black, Harrison, Lee, Marshall & Wiliam, 2003; Torrance & Pryor, 1998, 2001). This purpose of classroom assessment is often referred to as *assessment for learning*. Second, the dynamic and contextual nature of classroom assessment was often noted, suggesting a need for diverse better, rather than universal best assessment practices (Black & Wiliam, 2005; Gipps, 1999, 2005; Moss, 2003; Raveaud, 2004; Watson, 1999). And third, despite repeated recommendations from mid-century on for the improvement of teachers' competence in assessment (e.g., Barnes, 1985; Feldt, 1962; Joint Committee on Standards for Teacher Competence in Educational Assessment of Students, 1990; Scriven Carter, 1952), their practices were found to be persistently at odds with established measurement theory

(Brookhart, 1994; Cross & Frary, 1996). During the last decade of the 20<sup>th</sup> century, urgent concern was expressed within the measurement community about the lack of assessment literacy among teachers (Hills, 1991; Plake & Impara, 1993; Stiggins, 1991). However, much of the research at this time was approached from a conventional measurement perspective based on psychometric theory. As Brookhart (2004) points out, “someone or something is found wanting” (p.447) when classroom assessment is studied from this perspective.

Faced with a persistent gap between measurement theory and classroom practice, some researchers rejected the assumption that the cause was necessarily located in practice. As measurement theory had evolved through the 20<sup>th</sup> century to guide the development of tests and the use of test results, they questioned its relevance for the dynamic nature of classroom assessments (e.g., Brookhart,1993; Stiggins,1991; Whittington,1999, Wiggins, 1993). Wiliam (1994) recommended that educational assessment theory find “its own voice” (p.18), and Gipps (1994) suggested that it move “beyond testing and its technology” (p.158). Current empirical work continues to echo concerns about the use of technical qualities, such as validity and reliability, to judge teachers’ classroom assessment practices (e.g., Stokking, van der Schaaf, Jaspers & Erkens, 2004). Delandshere (2001) has been particularly critical of the concentration on technique and procedure that has dominated research in this area:

Ethical questions in educational assessment have primarily addressed the test developer’s competence, administration procedures, and the right for privacy and confidentiality of the person being assessed. Morality, intellectual coercion, and violation of students’ rights . . . have not been extensively addressed in the context of educational assessment, for which they are equally relevant. (pp.121-122)

By the beginning of the 21<sup>st</sup> century, the growing dissatisfaction with existing measurement theory (e.g., Bulterman-Bos, Terwel, Verloop & Wardeker, 2002; Popham, 2003) gave

impetus to considerable discussion regarding theoretical alternatives for classroom assessment (Brookhart, 2003, 2004; Delandshere, 2001; Gipps, 2005; McMillan, 2003; Moss, Pullin, Gee & Haertel, 2005; Shepard, 2000, 2006). Within this discussion, the definitions of measurement qualities, such as validity and reliability, have been adjusted to be more attuned to classroom dynamics (e.g., Moss, 2003; Smith, 2003), and the social nature of classroom assessment has been widely acknowledged in the literature (Allal & Pelgrims Ducrey, 2000; Bell & Cowie, 2001; Gipps, 1999; Shepard, 2005; Torrance & Pryor, 2001; Yorke, 2003). In the process of rethinking assessment theory for the classroom, it has sometimes been assumed, as Elwood (2006) and Gipps (2005) point out, that these changes make assessment more equitable. They argue, though, that more attention must be paid to issues of fairness for the ideals of assessment for learning to be realized in practice.

Concerns about fairness are viewed differently now than when they were expressed a century ago. As part of the theoretical evolution in psychological and educational measurement, the central concept of validity was expanded to include consequential evidence, which effectively pushed assessment discourse beyond technical matters into the ethical dimension (Messick, 1989; Moss, 1998). The role of values in educational assessment is now recognized, and fairness is seen as an important quality that is distinct from, but intertwined with validity (Messick, 2000; Stobart, 2005). Measurement specialists have broadened the definition of fairness (e.g., Crocker, 2003; Helms, 2006), and testing organizations have responded to these changes by revising their standards for fair tests (e.g., Educational Testing Service, 2002). In the last revision of the *Standards for Educational and Psychological Testing* (Joint Committee on Standards for Educational and Psychological

Testing [JCSEPT], 1999), a new section with four chapters of standards relating to fairness was included, and the document now emphasizes “the importance of fairness in all aspects of testing and assessment” (p.73). It also acknowledges that “varying views of fairness” (p.74) currently exist, and it discusses four common interpretations. Two of these, which are an absence of instrument bias and equitable treatment in the testing process, are determined to be the most appropriate for fairness in testing. The other two interpretations, which are opportunity to learn and equality of outcome, are rejected for the context of standardized testing. The standards that follow focus on issues of bias and equity in the development and administration of standardized tests and in the interpretation of their results. These standards are particularly concerned with “subgroups” (p.80) in the testing population, which are based on “age, gender, racial/ethnic, cultural, disability, and/or linguistic groups” (p.81), and the nature of any differences in the testing process or results for these groups.

Although social, ethical and legal issues impact the development and use of standardized tests and the interpretation of their results, test fairness is still determined largely through the technical analysis of group differences. Delanshere (2001) has noted that fairness issues in educational assessment are “often reduced to discussions of the technical aspects of test scores and selection bias, and rarely address matters of philosophical and social theory” (p.116). McMillan (2003) and Whittington (1999) have stressed that these technical aspects of measurement theory lack relevance for the ongoing decision-making involved in classroom assessment, and Camilli (2006) has pointed out that the statistical techniques used to determine fairness in testing are “impractical” and “of limited application” (p.246) for establishing fairness in classroom assessment. At present, however, fairness has

not been the target of much investigation outside the confines of standardized testing, and it is not well understood as a quality of the dynamics of classroom assessment.

To date, most of the research relating to fairness has been concerned with large-scale or standardized testing. Although fairness is often referred to in the literature on classroom assessment, explicit definitions are much less frequent. This omission is illustrated by a textbook for teachers that uses the word fair prominently in its title, but does not index, describe, or define fairness (Wormeli, 2006). The desirability of fairness as a quality of classroom assessment seems to be generally assumed, perhaps because fairness is considered an essential virtue in teaching practice (Campbell, 2003). Teachers do voice concern about assessing students fairly (Brookhart, 1993, 1994; McMillan, 2003; Rex, 2005; Ryan, 2001; Shepard, 2000; Stiggins, 1991; Yip & Cheung, 2005), and they also mediate principles and policies based on their beliefs about what is fair for their students (Brookhart, 1994; Delandshere & Jones, 1999; Morgan & Watson, 2002; Nagy, 2000; Rex). However, they not only interpret the concept of fairness differently (Yung, 2001b; Zoeckler, 2005), they often have little guidance in making assessment decisions, and they frequently face dilemmas “where competing personal, professional and practical demands” make it “particularly difficult to determine the most appropriate action” (Black & Halliwell, 2000, p.105; also Barnes, 1985; Bell & Cowie, 2001; Black, 2001; Brookhart, 1994; Delandshere & Jones; Eggen, 2004; Whittington, 1999). Yet, the personal, social, and educational effects of teachers’ assessment decisions can be significant and long-lasting for students (Brookfield, 2001; Brookhart, 1994; Delandshere & Jones; Gipps, 1999; Hargreaves, Earl & Schmidt, 2002; Watson, 1999). With the emphasis that is currently placed on the use of assessment for

learning in many educational systems, and with the recognition that all forms of educational assessment are value-laden, fairness can no longer be approached from a purely technical perspective. There is now a substantive need to understand fairness as a quality of the ongoing interactions involved in classroom assessment.

#### Rationale for Turning to Teachers

Teachers have had difficulty explaining their assessment decisions in past research (e.g., Dixon & Williams, 2003; McMillan, 2003). This may be because practical knowledge is tacit (Berliner, 2004; Hodgkinson, 1991; Meijer, Verloop & Beijaard, 1999) or latent (Torrance & Pryor, 1998). It could also be that teachers do not think of their assessment practices in technical terms, but focus on practical or moral aspects, such as the needs of particular students, or the consequences of an assessment decision. Larabee (2003) notes that teaching is “irreducibly moral” in that it “puts a premium on doing what is best for the student” (p.17), and teachers do seem to be more expressive about their assessment practices when discussions are framed in practical and moral, rather than technical terms (e.g., Blanchard, 2005; Brookhart, 1993; Campbell, 2003; Eggen, 2004; Zoeckler, 2005).

Critics of 20<sup>th</sup> century “technical ascendancy” (Dunne, 1993, p.16) argue that practical and moral knowledge has been undervalued in education, and they suggest that greater consideration should be given to the wisdom that guides practice (Bernstein, 1985; Biesta, 2007; Dunne; Hodgkinson; Kessels & Korthagen, 1996). In classical terms, practical wisdom is a “a reasoned and true state of capacity to act with regard to human goods”(Aristotle, trans. 1925/1975, p.143). Essentially, it is the ability “to judge correctly, all things considered, the best action to perform in any given situation” (Warne, 2006, p.15). Practical wisdom draws

on a mental network that includes technical knowledge, theoretical knowledge, moral beliefs and professional ethics, personal characteristics, experience, and understanding of particulars (Aristotle; Connolly, 2005; Dunne). In using practical wisdom in the classroom, a teacher might use different combinations of the elements in this network. For example, a teacher might draw on classroom experience and theoretical knowledge about learning to determine the degree of scaffolding needed by a class in certain circumstances, whereas deciding how to accommodate a particular student might be based on technical knowledge, moral beliefs, and an understanding of that particular student. Research on practical wisdom has explored issues and dilemmas encountered in teachers' practices (e.g., Connolly; Salloum, 2006), and it has focused on how new teachers develop practical wisdom (Birmingham, 2003; Field & Latta, 2001; Phelan, 2005). Teachers' practical wisdoms are a rich source of information that has not been explored for understanding the theory and practice of classroom assessment.

#### Research Purpose and Questions

The purpose of this multi-case study is to understand and describe the concept of fairness in classroom assessment by drawing on the practical wisdoms of six teachers, particularly in relation to dilemmas commonly found in practice. Three questions were developed to guide the research process:

1. What are teachers' practical wisdoms about fairness in classroom assessment?
2. How do these practical wisdoms about fairness inform teachers' decisions in classroom assessment dilemmas?
3. Which external factors facilitate or constrain teachers' practical wisdoms about fairness in classroom assessment dilemmas?

While the first question addresses the purpose of the study directly, the other two questions were designed to provide a fuller understanding of fairness in classroom assessment by looking at internal and external influences. Thus, the second question focuses on the elements of practical wisdom that teachers might use, and the third question looks at the contextual factors that might come into play. Given that the primary purpose of classroom assessment is to support learning, my focus in responding to these questions is on fairness in assessment for learning. My work is exploratory and descriptive in nature, which means that I do not aim to debate the fairness of teachers' assessment practices, nor do I wish to prescribe how fair assessment should be practiced universally by all teachers.

This research is situated in Writing in English Language Arts in Grades 7 to 12 in Ontario's publicly funded educational system. A specific subject was chosen because several studies have attributed differences in teachers' assessment practices to the subject matter (Duncan & Noonan, 2007; Hodgen & Marshall, 2005; McMillan, 2001). I chose the particular subject, Writing in English Language Arts, because it was of interest to me as a former elementary English-to-French Immersion teacher. In undertaking this work, I felt that my professional qualifications in this subject area would infuse a degree of credibility in my relationship with the participants, help me develop appropriate research tools, and also improve my ability to analyse participants' responses (Fenstermacher, 1994). Writing assessment is an area that teachers find particularly challenging (Blanchard, 2005; Graham, 2005; Zoeckler, 2005), and fair assessment to support the development and appreciation of writing is neither straightforward, nor well understood. Social and educational imperative also exists for this choice in that writing is generally considered a "gateway skill" (Frey,

Fischer & Hernandez, 2003, p.48) for further academic studies and success in many areas of employment. My justification for choosing Grades 7 to 12 is similar as the early adolescent years are considered a critical period for academic learning (Wigfield & Eccles, 2002; National Middle School Association, 2003). This research was undertaken in Ontario's educational system as a matter of personal interest as well as geographical convenience.

### Researcher's Perspective and Goals

My research perspective and goals for this work evolved through the interweaving of my interests, experiences, and beliefs about knowledge. Pursuing graduate studies has enabled me to elucidate my beliefs about knowledge more clearly, and my experiences as a teacher, a parent, and a learner continue to spark my interests in educational research.

While I was an elementary English Language Arts teacher, I enjoyed sharing my passion for texts of all kinds with young readers, and I found that helping students develop communication skills and cultivate enjoyment in the art of expression was immensely rewarding. When it came to classroom assessment, though, I was not as satisfied. I believed that it was my professional responsibility to assess and evaluate student learning as sensitively and accurately as possible, but it was not a process I enjoyed. I felt that factors beyond my control, such as report card requirements, large class sizes, and limited preparation time, hindered the process. With little knowledge about classroom assessment, I did not understand the potential of assessment for learning. I was not at ease as a teacher when I recognized that the impact of educational assessment was often less than desirable for students, nor was I satisfied as a parent with the murky reflection of learning that my children's report cards presented. Experiencing the portfolio assessment process as a student

in one of my graduate courses drew my attention to the potential of classroom assessment for learning. However, in reading the research on classroom assessment, I was surprised by the frequent focus on teachers' deficits. I also became increasingly aware that the technical approach to classroom assessment, which dominates much of the literature, meant that underlying assumptions often went unexamined in both research and practice. Through my personal and professional experiences and my formal studies, then, I came to see that the ethical aspects of classroom assessment were not only those which concerned me the most, but also that they were not well understood within the educational community. I undertook this study with the belief that the ethics of classroom assessment in general, and the fairness of assessment for learning in particular, would benefit from inquiry that valued teachers' practical wisdom.

My philosophical approach to research is best described as a critical form of pragmatism. Pragmatism is well suited to educational research because it focuses on practice (action) and interaction (Biesta & Burbules, 2003). Cherryholmes (1988) writes that "critical pragmatism results when a sense of crisis is brought to our choices, when it is accepted that our standards, beliefs, values, guiding texts, and discourses-practices themselves require evaluation and reappraisal" (p.151). In my dissertation, the reappraisal, or rethinking, is focused on the concept of fairness. From the perspective of critical pragmatism, knowledge is both situated and transient, and this has several implications for my research. Because I view knowledge and knowing as situated in time and place, and I agree with Cherryholmes that "we make decisions about beliefs and actions against this background" (p.185), I believe that a generic classroom is inadequate as the context for understanding the concept of fairness in

assessment. I have, therefore, situated my inquiry both in terms of a general classroom environment, and more specifically in the time and place of a specific educational context, and I pose my research questions against that backdrop. A second implication for my research relates to the transient nature of knowledge. Rather than maintaining a fixed lens on educational practice, such that certain values remain constant, critical pragmatism assumes that what is thought to be desirable should always be open for review (Cherryholmes; De Waal, 2005; Hartrick Doane & Varcoe, 2009). As such, I view research as part of an ongoing conversation (interaction) within an educational community, and I see myself, in undertaking this work, as a participant in a conversation about fairness in classroom assessment with other teachers and researchers.

With this research project, I wanted to bring teachers' practical wisdoms to a "more conscious and sophisticated level" (Hodgkinson, 1991, p.110; also Fenstermacher, 1994) in a conversation with teachers and researchers that goes beyond identifying effective techniques to consider what is both "possible" and "educationally desirable" (Biesta, 2007, p.10) in classroom assessment. The initial idea for this work was inspired by Shepard's (2000) call for classroom assessment research to be "embedded in the dilemmas of practice" (p.13). I was also inspired by those who have advocated change in our assessment systems (Brookhart, 1994; Stiggins, 2002; Wiggins, 1993), and by the researchers who have worked with teachers in practice to reconceptualize classroom assessment theory, particularly Bell and Cowie (2001), Black and Wiliam (2006b), and Torrance and Pryor (2001). Ultimately, I hope that this dissertation contributes to the growing body of literature on classroom assessment theory from which I have learned, particularly the literature that illuminates

assumptions about assessment in education, such as that by Bulterman-Bos and colleagues (2002), Delandshere (2001), Moss and colleagues (2005), and Schendel and O'Neill (1999).

### Definition of Terms

In this section, I briefly define five key terms that appear in my research questions to provide an understanding of how they are used in this text.

1. *Classroom Assessment*: Educational assessment is broadly encompasses all forms of assessment in education, including large-scale, standardized testing administered by external organizations, and the assessment and evaluation done by teachers and students in their classrooms. Classroom assessment refers more specifically to the latter. A distinction is often made between the phases of classroom assessment, where assessment is limited to the gathering of information about student learning, and evaluation involves a subsequent judgment (e.g., Ontario Ministry Education [OME], 2008). However, the compound term *classroom assessment* encompasses more than the act of gathering information in a classroom. In most of the current literature, the term classroom assessment is used broadly to include all phases of the process. For example, McMillan (2007) defines classroom assessment as the “collection, evaluation, and use of information to help teachers make better decisions” (p.8). Classroom assessment involves an interaction between individual knowledge and beliefs, the educational context, and professional action within a specific period. In Ontario’s English Language Arts curriculum, it is stated that the “primary purpose of assessment and evaluation is to improve student learning” (OME, 2007a, p.20). This assessment purpose is often referred to in current educational texts as assessment for learning. However, teachers are also required to evaluate student learning regularly, and the

results that are formally reported to students and their parents may also be used to rank, certify, and place students in educational programs. This assessment purpose is often referred to as *assessment of learning*. The multiple purposes of classroom assessment are discussed further in Chapter Two.

2. *Dilemma*: In the Ancient Greek language, the word *dilemma* means a double proposition (Harper, 2001). By current dictionary definition, a dilemma requires a choice between two equally undesirable outcomes. Usage is less strict in educational literature where a problematic situation involving a difficult choice is often described as a dilemma, and some even suggest the possibility of positive resolution. Buzzelli & Johnston (2002), for example, define a moral dilemma as that “in which the options involve consequences that are both good and bad” (p.3). A caveat in defining this word is that a dilemma can be a matter of perception. What one teacher with a particular combination of beliefs, knowledge and experience in one context might perceive as a dilemma, another teacher might not. Dilemmas may lead to practices that are less than ideal (Windschitl, 2002), but they may also be “assets” (Eggen, 2004, p.103) for further professional learning (Black & Halliwell, 2000; Connolly, 2005; Ross, 2005), especially if teachers’ practical wisdoms about dilemmas are shared with other teachers and researchers (Shulman, 1997/2004). Shepard (2000) suggests that research situated in the dilemmas of practice may be the most useful for understanding the complexities of classroom assessment. In this study, I focus on dilemmas relating to fairness that teachers face in classroom assessment for learning.

3. *External Factors*: Many factors that are external to teachers’ thinking can influence the practice of classroom assessment. Some that have been found in current classroom-based

research include: assessment policies, class size, collegial relationships, curriculum, grade level, parents, professional development opportunities, the range of student abilities, report card requirements, subject matter, school leadership, student behaviour, teaching traditions, and time constraints. These are discussed in further detail in Chapter Two. My interest in this study is in understanding which factors are relevant for fairness in classroom assessment.

4. *Fairness*: The word *fair* has several meanings in the English language, but in the field of education it is generally considered a desirable quality. Fair practice is “marked by impartiality or honesty, and free from self-interest, prejudice, or favouritism” (Meriam-Webster, 2006). In current assessment literature, fairness is understood as a social value as well as a technical issue (Messick, 2000; Stobart, 2005), but it is not consistently defined. The *Standards for Educational and Psychological Testing* (JCSEPT, 1999) acknowledge that differing interpretations exist, but emphasize that “many testing professionals would agree that if a test is free of bias and examinees have received fair treatment in the testing process, then the conditions of fairness have been met” (p.76). Fairness has also been subject to diverse interpretations in classroom assessment literature, but it has not been the focus of sufficient attention within the educational research community for any form of consensus to have emerged. The call to reconceptualize fairness specifically and clearly for the purposes of classroom assessment is at the heart of this study. The various interpretations of fairness, along with related terms (e.g., equity, equality, bias and ethics) used in classroom assessment literature, are discussed further in Chapter Three.

5. *Practical Wisdom*: This term was defined earlier in this chapter (see pp.7-8), and the concept is discussed further in Chapter Four.

## Organization of the Thesis

This dissertation is organized into nine chapters. I opened the first chapter with the problem and its background, and I stated the purpose and the questions for my study, defined key terms, and briefly elucidated my research perspective and goals. In the second chapter, I provide information that is relevant for understanding fair assessment in the classroom environment. This includes a discussion about the purposes of classroom assessment, and a description of the external factors that may be influential in Ontario's educational system. In the third chapter, I focus on literature relating to fairness in classroom assessment, including a small body of research relating to fair writing assessment in Ontario. In the fourth chapter, I elaborate further on critical pragmatism, and I explain why I chose to use the Aristotlean concept of practical wisdom rather than other conceptualizations of teachers' knowledge. I then summarize my discussion in relation to a visual illustration to clarify how the key concepts and questions in this study are related. The fifth chapter contains an explanation of the methodology, with sections on the research design, sources of information, research tools and processes, as well as the strategies I used to enhance the quality of my research. I present the individual cases in the sixth chapter, cross-case results in the seventh chapter, and my discussion of the most salient aspects of the results in the eighth chapter. I wanted the results and my discussion to respond clearly to my three research questions, so I have organized these three chapters accordingly. In the ninth and final chapter, I account for the limitations of this project, discuss implications for classroom practice and teacher education, and I give my recommendations for future research before concluding with some final thoughts.

## CHAPTER TWO: ASSESSMENT IN THE CLASSROOM ENVIRONMENT

*To understand the elephant . . . one must see it in context and as part of an ecological system in relation to other flora and fauna, in its natural environment.*

Michael Quinn Patton (2002, p.62)

This research aims to shed light on fairness as a quality of classroom assessment. For this it is necessary to understand the purposes of, and influences on, assessment in the classroom environment. In this chapter, I discuss the purposes of assessment, elaborate on the concept of assessment for learning, and discuss several issues in related to the use of classroom assessment for multiple purposes. In a final section I describe external factors that have been found to influence classroom assessment, and I identify those that seem most pertinent for fair assessment in Ontario's educational system.

### The Evolution of a Continuum of Assessment Purposes

Hodgkinson (1991) writes that education serves a “constellation of purposes” (p.23), and the same can be said more specifically about assessment in education. The purposes that assessment serves have fluctuated as educational needs and ideals have evolved over time (Delandshere, 2001; Gipps, 1999). According to Bell and Cowie (2001), explicit discussion regarding different purposes of assessment began in the 1990's when politicians and other stakeholders “looked to assessment to provide the information required for the accountability process” (p.537). Although assessment had been used for centuries for the purpose of selecting students or candidates (Delandshere; Gipps), the increased demand for accountability sparked debate within the educational research community about the ideal purposes of assessment (Bell & Cowie).

There is currently little consensus regarding the number of purposes that educational assessment should serve. Often only two main purposes are identified (e.g., Brookhart, 2003; Butterfield, Williams & Marr, 1999; Delandshere & Jones, 1999; Eggen, 2004), but other authors identify purposes ranging from three (e.g., Nagy, 2000; Rea-Dickens, 2001) to as many as nine (Fenwick & Parsons, 2000). This diversity of opinion supports Harlen's (2007) argument that there is no "hard and fast dividing line" (p.121) between assessment purposes, and that they should be considered as a continuum. Drawing on work by Earl (2003), Harlen (2006, 2007), Horn (2002) and Katz, Earl and Olson (2001), Table 1 shows various aspects of educational assessment in relation to a continuum of purposes. While thinking of assessment purposes as a continuum suggests a linear progression that does not reflect the dynamics of classroom practice, it does help illustrate the complexities of assessment, and facilitate discussion about issues relating to assessment purposes.

In his writing about reform in curriculum, instruction and assessment, Horn (2002) describes a range of assessment purposes. He explains that one purpose is to provide evidence of accountability for educational stakeholders based on the results of standardized tests. This purpose of assessment is the least supportive of individual learning because it does not provide regular, timely or detailed feedback to students, and it aims instead to "rank and sort" (Horn, p.219) students or schools in an educational system. In Table 1, this purpose of assessment is shown at one end of the continuum as a form of assessment of learning. Harlen (2007) explains that while assessment of learning may be administered as a one-time assessment, such as a standardized test, it can also involve "summarizing achievement across a period of time up to the reporting date" (p.16). In the classroom environment, assessment of

Table 1

*Purposes of Educational Assessment on a Continuum*

Anchor	LEARNING	↔	ACCOUNTABILITY
Prepositional Distinction in Assessment Literature	Assessment for Learning (as per Black & William, 2006; Stiggins & Chappuis, 2005)		Assessment of Learning
	Assessment as Learning	Assessment for Learning (as per Earl, 2003)	
Assessment Function	Formative and Ipsative	Formative	Formative or Diagnostic
Specific Assessment Purpose (as Points on Continuum)	Developing Self as Learner	Guiding Learning	Determining Instructional Needs
	Learner	Formative and Guiding Learning	Formative and Adjusting Teaching
Frame of Reference	Learner	Learner and Criteria	Learner and/or Criteria and/or Norms
Evaluator	Student	Student Teacher	Teacher Psychologist
			Teacher Administrator Psychologist External Assessor
Teacher's Role	Mentor and Coach	Facilitator and Planner	Judge and Reporter
			Teacher Administrator External Assessor Politician
			Summative
			Summative
			Checking and Reporting on Achievement
			Checking System Effectiveness
			Criteria and/or Norms
			Criteria and/or Norms
			Criteria and/or Norms

learning can include a single major assessment (e.g. final examination) as well as a summarization of assessment results accumulated over the duration of a course or year.

Classroom assessment of learning usually requires teachers to report on student achievement to students and parents, and these reports are used by teachers and administrators for a variety of educational decisions. Rea-Dickins (2001) suggests that this is the “bureaucratic face” (p.451) of classroom assessment. Earl (2003) notes that there “will always be milestones and junctures” where assessment of learning is needed to measure student achievement.

However, interest in assessment for learning has increased dramatically as learning theories, and thinking about the role of assessment in learning have evolved.

#### The Concept of Assessment for Learning

Carr & Claxton (2002) suggest that the purpose of education in the 21<sup>st</sup> century has shifted from the transmission of knowledge to the encouragement of lifelong learning. In a seminal speech on assessment in a learning culture, Shepard (2000) described this as a shift from a traditional framework based on behaviourist learning theory and modern measurement to an emergent framework involving social constructivist learning theory, and she argued for a reconceptualization of classroom assessment that recognized this shift. Research on classroom assessment is now frequently approached from perspectives that emphasize the social nature of learning (e.g., sociocultural, social constructivist). This shift in thinking about classroom assessment has illuminated the link between teaching, learning and assessment, and focused attention on the potential of assessment for learning.

In principle, it is widely accepted that the primary purpose of classroom assessment in elementary and secondary education is to support learning. Not only is this purpose

championed by specialists in the field (e.g., Black & Wiliam, 2006a; Stiggins & Chappuis, 2005), it seems intuitive for some teachers (Black & Wiliam, 2003; Hayward, Priestly and Young, 2004; Torrance & Pryor, 1998), and it is now commonly endorsed within educational systems around the world. For example, the Ontario Ministry of Education (OME) clearly states that “the primary purpose of assessment and evaluation is to improve student learning” (2008, p.1-I). The term *assessment for learning* can be traced back to the questions raised by Gipps (1994) about the purpose of educational assessment (Pryor & Crossouard, 2008; Threlfall, 2005). The roots of the concept go further back, though, to early definitions of formative assessment by Bloom, Hastings and Madaus (1971), Ramaprasad (1983), and Sadler (1989). At present, the terms formative assessment and assessment for learning are often used synonymously (e.g., Earl, 2003; Gardner, 2006; Harlen, 2007; Weeden, Winter & Broadfoot, 2002), and they tend to be considered “conceptually identical” (Threlfall, 2005, p.54). One exception to this is in the distinction that Earl (2003) makes in using the term *assessment as learning* to describe assessment actively used by students to adjust their learning (see Table 1). When assessment as learning is thus distinguished, assessment for learning is defined more narrowly as being used by the teacher to adjust instruction. Dann (2002) uses the term assessment as learning in a book that focuses on understanding learning and the role of the learner in assessment, and it has also been taken up by several educational organizations (e.g., see the Internet sites of Learning and Teaching Scotland, 2008; British Columbia Ministry of Education, 2007). However, the term assessment as learning has yet to be widely adopted by the research community. Assessment for learning continues to be used as an umbrella term for all assessment that functions formatively, either through the direct

use of assessment information by students, or through use by teachers to adjust instruction and support learning (e.g., Black & Wiliam, 2006a; Stiggins, 2005). While I have tried to respect each author's terminology, my use of the term assessment for learning in this dissertation is consistent with common usage in classroom assessment literature in that it refers to formative assessment and includes assessment as learning.

The early definitions of formative assessment fit well with behaviourist and early cognitive theories of learning (Allal & Pelgrims Ducrey, 2000; Yorke, 2003), and formative assessment is still defined cognitively as a linear sequence, involving teacher-directed instruction, feedback and "correctives" (Guskey, 2005, p.4). Yorke (2003) notes, however, that "formative assessment is a concept that is more complex than it might appear at first sight" (p.478). This notion is supported by Pryor and Torrance (1998) who found in their analysis of formative assessment events in elementary classrooms in England that a "purely cognitive" approach minimized the "complexity of the situation" (p.170). The social nature of assessment for learning is now commonly recognized (e.g., Cowie, 2005; Gipps, 1999; Shepard, 2005; Torrance & Pryor, 2001; Yorke, 2003), and assessment for learning is understood as an interaction between the teacher, the students and the students' work (Allal & Pelgrims Ducrey, 2000; Pryor & Crossouard, 2008).

According to Cowie and Bell (1999) formative assessment can be both planned and interactive. Cowie and Bell conducted a study for the New Zealand Ministry of Education where they collaborated with 10 science teachers and their students to model formative assessment, and they concluded that the process involves a dual cycle (see Figure 1). With planned formative assessments the teachers aimed to "obtain information from the whole

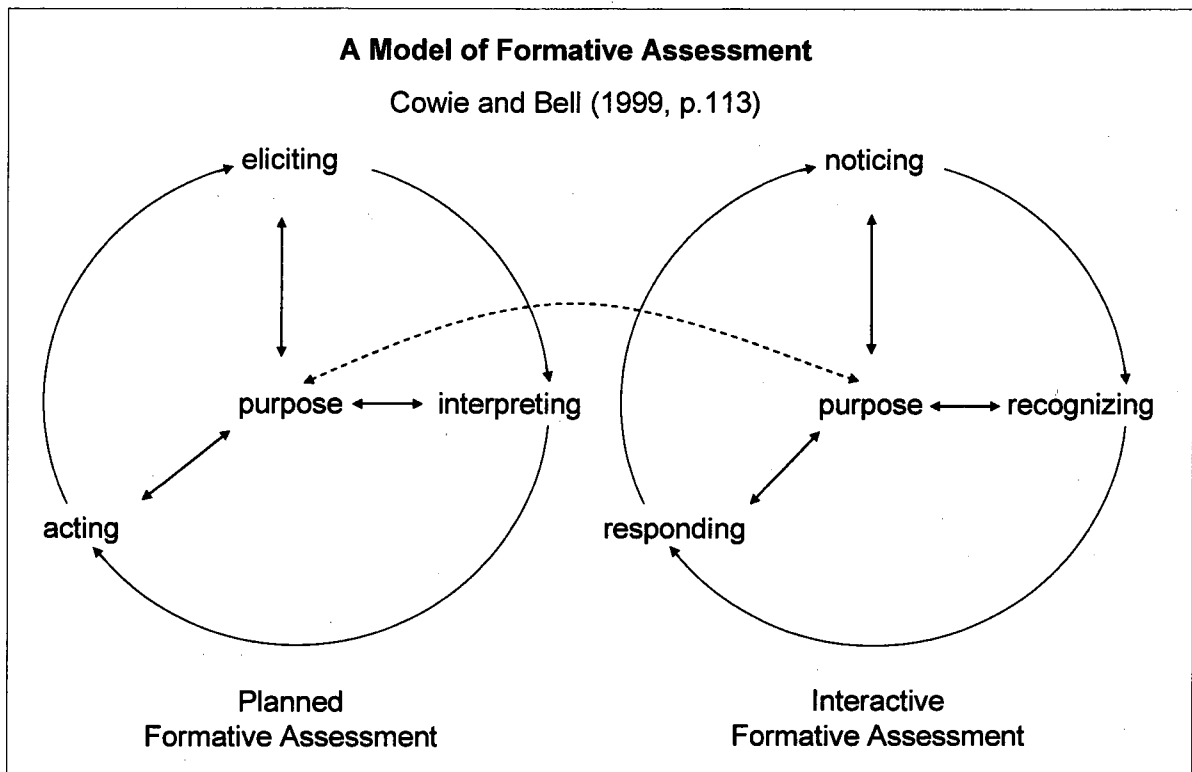


Figure 1. Process model of formative assessment as a dual cycle.

class” (p.103) about learning in relation to the curriculum. They used a variety of strategies in a cycle of eliciting, interpreting and acting on information with the intention of informing instruction and supporting group learning. The teachers also engaged in interactive formative assessment that occurred “during student-teacher interactions” when a “specific assessment activity was not planned” (p.107). This use of the term *interactive formative assessment* is somewhat different from its use in the French-language literature of the 1980's and early 1990's. As Allal and Pelgrims Ducrey (2000) explain, interactive formative assessment in the earlier French-language literature involves assessment by teachers in the ongoing interactions of classrooms for the purpose of informing further instruction, which became known in the English-language literature of the 1990's simply as formative assessment. As research on

formative assessment progressed during this period, its ability to inform learning as well as instruction was emphasized. For Cowie and Bell, interactive formative assessment is a cycle of noticing, recognizing, and responding to student learning that occurs spontaneously during classroom activities, and is thus “embedded in and strongly linked to learning and teaching activities” (p.108). Interactive formative assessment is more immediate and responsive to individual students than planned formative assessment, and it may occur whenever students and teachers interact. While the teachers in Cowie and Bell’s study found that interpretation in planned formative assessment drew heavily on their teaching experience and knowledge base (Shulman’s model), they found the interactive process even more challenging because it required an alertness to students’ disclosure of learning during classroom activities. Based on Cowie and Bell’s work, I understand assessment for learning as being either planned or interactive (spontaneous) processes that depend upon a teacher’s ability to use assessment information to adjust teaching and support student learning in an ongoing cycle.

Over the past decade there has been a concerted effort to describe assessment for learning, and a number of recent works have identified practices within its ongoing cycles. Tierney and Charland (2007) synthesized descriptions given by the Assessment Reform Group (ARG, 2002), Bell and Cowie (2001), the Centre for Educational Research and Innovation [CERI] (2005), Crooks (2001), Leahy, Lyon, Thompson and Wiliam (2005), Shepard (2005, 2006), Stiggins (2002), and Torrance and Pryor (2001), and they model assessment for learning as a composite of five core practices: a) clear learning expectations and assessment criteria, b) varied approaches to elicit learning, c) balanced and descriptive feedback for students, d) the adjustment of teaching as a result of the assessment, and e) the

active involvement of students in the learning-assessment process. Tierney and Charland (2007) note that it is through these core practices that teaching, learning and assessment are integrated in a constructive learning environment.

While the need for a constructive learning environment is often stated or implied in the literature on assessment for learning (e.g., ARG, 2002; CERI, 2005; Stiggins, 2002), ideas about what that entails differ. For example, Shepard (2006) emphasizes that teachers should “establish a climate of trust and develop classroom norms that enable constructive criticism” (p.632), whereas Pryor and Crossouard (2008) suggest that the “traditional socio-affective-cognitive dichotomy of formal learning contexts” needs to be disrupted to permit an “affective environment that is more supportive of learning” (p.14). However, they note that this involves sociocultural changes beyond the classroom. It may be that the exact features of a constructive learning environment vary across educational contexts.

My understanding of assessment for learning combines Cowie and Bell’s (1999) cyclical model and Tierney and Charland’s (2007) composite model. I see the five core practices as part of the actions and interactions in the ongoing cycles. Generally, the process begins with teachers planning tasks or activities (approaches) and sharing the learning expectations and assessment criteria with students. Whether the assessment interaction is planned or spontaneous may change features of assessment for learning practices, such as how expectations and criteria are shared, or how feedback is used. While these practices are part of assessment for learning, their exact use by students and teachers depends on the nature of the assessment activity, the particular circumstances of the classroom, and the characteristics of the larger educational context.

## Issues in the Use of Classroom Assessment for Multiple Purposes

Although assessment for learning has been widely accepted in principle, empirical research in different educational contexts suggests that assessment of learning still dominates classroom practice (e.g., Lee, 2007; McNair, Bhargava, Adams, Edgerton & Kypos, 2003; Volante & Fazio, 2007). Studies with teachers who aim to change their classroom practices and use assessment for learning have shown that the process is not without challenges (e.g., Borko, Davinroy, Bliem, & Cumbo, 2000; Briscoe & Wells, 2002; Lock & Munby, 2000; Schmidt, 2001). Given that classroom assessment is not usually practiced, at this point, solely to serve learning, I identify and discuss several issues that arise in the use of classroom assessment for multiple purposes as they may have implications for studying fairness.

Teachers and researchers have used a wide variety of labels and metaphors to represent teachers' roles, including butcher, caretaker, coach, collaborator, companion, editor, entertainer, examiner, judge, police officer, shepherd, tutor (Brighton, 2001; Eggen, 2004; Hand & Prain, 2002; Markus, 2002; Schmidt, 2001; Speck, 1998; Yung, 2001a). Going back to Table 1, it can be seen how some of these roles fall along the continuum of assessment purposes. Teachers who are involved in research on changing assessment practices often comment on how their roles change, and shifting between roles can be problematic for some teachers for two reasons. First, as Delandshere and Jones (1999) note, the intellectual demands are different. In assessment of learning, the teacher functions as judge and reporter, checking student achievement against an external frame of reference (e.g., criteria or norms) in order to communicate information to stakeholders. In assessment for learning, the teacher may play a number of supportive roles in an ongoing dialogue with the

learner. Ruddiman (2004), for example, found that the teacher in her case study shifted between being a friend, mentor, leader, and taskmaster as she wrote comments in her students' reading response journals. The second reason that changing assessment roles can be challenging occurs when students are involved in assessment, and they become increasingly responsible for their own learning. In theory, this is one of the benefits of assessment for learning, but it alters the teachers' position in the classroom, and accords power to the students. For example, one veteran science teacher in a four-year collaborative professional development project in Australia stated that he became less of a "wisdom-giver" and more of a "facilitator" (Hand & Prain, 2002, p.750) as he embraced assessment for learning. This is consistent with current texts on classroom assessment that encourage a high degree of student involvement and portray teachers more as learning coaches than classroom authorities (e.g., McMillan, 2007; Stiggins & Chappuis, 2005). However, questions about sharing responsibility, control, and the nature of teacher-student relationships are frequently raised in relation to the practice of assessment for learning (e.g., Black & Harrison, 2001; Borko et al., 2000; Brighton, 2001; Clark, Chow-Hoy, Herter & Moss, 2001; Hand & Prain, 2002; Lock & Munby, 2000; Schmidt, 2001). Several authors have suggested that the underlying power dynamics of assessment should be given further consideration (Elwood, 2006; Gipps, 1999; Hargreaves et al., 2002; Speck, 1998), and this may be particularly important for understanding fairness in classroom environments where teachers and students shift between different roles.

Another issue that arises when classroom assessment serves multiple purposes involves the frame of reference that teachers use as assessors. A multi-site case study by

Markus (2002) provides a good example. She looked at the practices of three experienced Art teachers in Ontario, and she found notable differences in how they assessed learning. One teacher encouraged her students to assess their own growth using an ipsative framework, whereas another used a normative framework to compare her students and prepare them for an external exam (International Baccalaureate). The third teacher paired flexibility with a criterion-referenced assessment system, and Markus concludes that this approach had a more positive impact than the other two. What is most important here, though, is that while all three teachers complied to some degree with mandated policies, they all developed their own strategies (e.g., taking attitude and effort into account, or maintaining flexible due dates), which allowed them to reconcile policy requirements with their own views about assessing student art work. Markus' findings are similar to what has been seen in other educational contexts with criterion-referenced assessment systems. For example, some of the teachers who collaborated with Torrance and Pryor (1998) in England explained that they used student-referenced assessments to support learning, and some of the teachers who shared their stories with Bulterman-Bos and colleagues (2002) in the Netherlands related that they made judgements about students using a practical form of norm-referencing based on knowledge of learners accumulated through experience. McMillan (2001) also found that many of the teachers he surveyed in Virginia used "some kind of norm-referencing" (p.30), and he questioned whether these types of comparisons are fair. While some teachers may shift their reference point from the student to the criteria when they move from assessment for learning to assessment of learning, others seem to maintain the same frame of reference throughout. Consequent differences in how teachers assess, either within or across

classrooms, raise questions about the validity and fairness of classroom assessment results.

A final issue relating to the multiple purposes of classroom assessment concerns the degree to which they can be functionally distinguished. Several studies have shown that despite the different purposes and qualities of assessment for learning and assessment of learning, they are not always divisible in practice. For example, the teachers who worked with Torrance and Pryor (2001) appreciated the pedagogical power of divergent assessment (i.e., variation of assessment for learning), but they still felt that it was important to practice convergent assessment (i.e., repeated assessment of learning). It seems that the “conceptual distinctions” in classroom assessment theory were blurred in practice as the “processes were often embedded one within another or occurred in linked sequences or progressions” (p.622). There is not always a clear demarcation of when an assessment process stops being formative, and takes on a summative function. In another classroom-based study in England, the teachers were encouraged to develop their own methods of assessment for learning, and several combined formative strategies with summative events, such as reflective reviewing with students before and after classroom tests (Black et al., 2003). In theory, the formative function precedes the summative function, but these teachers sometimes reversed that order. In this case, students were being taught to use summative results for learning. Brookhart (2001) argues that some students do this without formal instruction. She interviewed high achieving students at a suburban secondary school in the United States, and she found that they use grades, which theoretically have a summative function, as information to direct further learning. These studies raise questions for fairness because some students may be better equipped to understand when assessment functions for different purposes, and how to

use all assessment information to their own advantage. This may be compounded further when assessment purposes shift as learning expectations and classroom relationships evolve during a course or school year. Zoeckler (2005) studied the grading practices of a group of teachers in one school department over a full school year, and noted that despite “its apparent stability of structure . . . one might think of a grading system as a kind of organic entity or living thing, constantly changing and yet always the same in some deeper sense” (p.223). In summary, it seems that assessment is a dynamic process in the classroom environment where the boundaries between purposes are often blurred.

Despite the blurring of purposes in classroom assessment, neglecting to consider how assessment for and of learning are different may have implications for fairness. In theory, the processes and characteristics of assessment for and of learning are distinct. Harlen (2007) models assessment for learning as a cycle where teachers and students share information, and assessment of learning is a sequence where students are a “source of evidence” (p.122). Earl (2003) also argues that “different purposes require vastly different approaches, and mixing the purposes is likely to ensure that none of them will be well served” (p.12). At this time, what might make both of these processes more or less fair has not been clearly established for the classroom environment. Camilli (2006) suggests, though, that establishing a “clear purpose” (p.247) may be a prerequisite for fairness. As the interest in assessment for learning is sufficiently recent, much of the discussion about fairness to date has been limited to assessment of learning. Although I recognize that assessment purposes are blurred in practice, I maintain the distinction in this dissertation to focus attention on assessment for learning, where fairness issues have been the least discussed.

## Influence of External Factors on Classroom Assessment

In a review of the influences on educational assessment, Gipps (1999) argued that because the purposes of assessment are “driven largely by social, political, and economic forces,” assessment can be understood only within the “contexts in which it operates”(p.355). Classroom-based research in different educational context has since shown that many factors do influence teachers’ assessment practices. In my study, I am specifically interested in factors that might influence fair Writing assessment in English Language Arts in Grades 7 to 12 in Ontario’s educational system. In this section, I discuss factors that have been found to influence classroom assessment in relation to the particulars of Ontario’s educational system. This provides the background for understanding the information on the influence of external factors given by the participants in this study. I have organized this section to move outward from classroom level factors, which are the closest to the practice of assessment for learning, to factors at the school and system levels.

### *Classroom Level Factors*

Based on the existing literature, factors that have been shown to influence assessment at the classroom level include class size, students’ characteristics, and parents.

*Class Size.* Class size has been the subject of considerable study, but contradictory results have been reported, and debate on the benefits of small class sizes is ongoing (Gilman & Kiger, 2003). Although most of the research on class size is concerned with the effects on student achievement, Duncan and Noonan (2007) surveyed secondary teachers (n = 513) in Saskatchewan about class size effects on assessment practices. While their analysis showed no significant statistical evidence to associate class size and teachers’ assessment strategies or

grading practices, they concluded that a relationship between class size and the use of assessment for learning could not be ruled out. Case study research has suggested that class size can influence the use of assessment for learning. One example is Jeranyama's (2001) work with three middle school science teachers in Michigan, where she noted that their ability to integrate teaching and assessment was impacted by the size of their classes.

Many teachers in Ontario report that large class sizes are problematic. In a survey conducted for the Ontario College of Teachers (OCT) in 2005 (n = 1000), teachers were asked to identify major problems in Ontario's schools from a list of eight options, and large class size was selected the most frequently (32%), superceding all other options (i.e., special students, discipline, curriculum and technology, government, teacher shortage, salaries). In the same survey, many respondent (65%) identified small class sizes as an "extremely desirable" (p.11) government initiative, and almost all (91%) rated class size as very important for student success. The eight teachers who participated in Schmidt's (2001) dissertation on the challenges they faced when implementing new classroom assessment practices in Ontario gives further insight on the influence class size. Most felt that their assessment practices, either in the amount of one-on-one interaction, the variety of assessment choices offered, or their depth of knowledge about students, was affected by the number of students in their classes. Although there has been an effort in Ontario to reduce class sizes (Naylor, 2007), most secondary teachers still had fairly large groups (23 or more students) during the 2006-2007 school year (People for Education, 2007). As the number of students affects the quality of the information that teachers gather, there may be implications for the fairness of classroom assessment related to the size of classes in Ontario.

*Student Characteristics.* When student characteristics are identified as a factor of influence in classroom assessment, the focus is often on students' abilities and attitudes. One of the teachers in Hand and Prain's (2002) study, for example, explained that the degree of choice he offered in his assessment tasks depended on the students' maturity. The English and Mathematics teachers (n = 24) who discussed their classroom assessment practices with McMillan and Nash (2000) were even more vocal on this point. They "repeatedly expressed frustration" about the impact of "classroom realities" (p.12), which related primarily to students' abilities, attitudes and behaviour, on their assessment practices. In most of the research on factors that facilitate or constrain classroom assessment, fairness is not explicitly discussed. An exception is seen in Morgan and Watson's study (2002), where they looked at the influence of students' writing skills on teachers' assessment decisions in mathematics, and concluded that students' abilities could affect the fairness (equity) of an assessment.

Students in Ontario tend to demonstrate high levels of achievement in Language Arts. Of the 43 countries that participated in the 2006 Progress in International Reading Literacy Study (PIRLS), which measures reading achievement at the Grade 4 level, students in only two countries (Singapore and Russia) scored higher than students in Ontario (International Association for the Evaluation of Educational Achievement [IEA], 2007). The results of the 2006 Programme for International Student Assessment (PISA), which measures the achievement of 15-year old students, were similar. Of the 57 participating countries, students in only three (Korea, Finland, Hong Kong-China) scored higher in reading than students in Ontario (Education Quality and Accountability Office [EQAO], 2007). This strong performance in English Language Arts is particularly noteworthy because of the diverse

nature of the student population in Ontario. According to the IEA (2007), the number of students in Ontario (38%) who speak a language at home other than their primary language of instruction (i.e., English or French) is the highest in Canada, and also higher than the international average (29%). A 2006 census of Grades 7 to 12 students (n = 105,440) in the Toronto District School Board (TDSB), which is the largest school board in Ontario, showed that the majority of TDSB students have at least one parent who was born outside of Canada (80%). Students in Ontario also have a wide range of learning needs. Many school boards provide advanced learning programs, such as French-Immersion, Concentrated Arts, Advanced Placement, and International Baccalaureate programs, as well as special education programs for students who require program accommodation and modifications. The teachers who responded to the 2006 OCT survey (n = 1000) estimated that 26% of their students had “special needs or required extra accommodation or compensating strategies” (p.31). In brief, Ontario’s student population is culturally and linguistically diverse, and fairly heterogeneous in terms of learning needs.

Teachers in Ontario have indicated that their classroom practices are affected by the characteristics of the student population. Almost all (96.9%, n = 415) of the English teachers in Ontario who participated in the 2002 School Achievement Indicators Program (SAIP) indicated that how they teach Language Arts is limited to some extent by the range of students’ abilities (Council of Ministers of Education, Canada [CMEC], 2002a). Despite this, teachers in Ontario do not identify student characteristics as one of their significant challenges. When the teachers who responded to the 2006 OCT survey rated the stressfulness of 16 aspects of teaching, for example, none that topped the list were associated with

students' characteristics (e.g., teaching English language learners or special needs students, maintaining discipline, etc.). It seems that while teachers do recognize the impact of student characteristics on classroom practices, there are more challenging aspects of teaching in Ontario. Given that the diversity of Ontario's student population limits practices in English Language Arts, it is possible that the fairness of classroom assessment is also affected, but existing information suggests that teachers are more concerned about other factors in Ontario's educational system.

*Parents.* Parents are the final factor to consider at the classroom level. Although parents are not usually present in Grades 7 to 12 classrooms, I have included them at this level because they are their children's legal representatives, and assessment results are reported to them through the secondary years. McMillan and Nash (2000) found that parents in Virginia influenced grading practices because the teachers in their study sought "sufficient justification for grades to avoid parental conflicts" (p.18). In a similar vein, Schmidt (2001) noted that reporting to parents in Ontario had become "forbidding" for the teachers in her study "because they had to defend why students received the mark they did" (p.116). Teachers in Ontario have reported that communicating with parents about student learning is sometimes stressful (OCT, 2006), but pressure from parents does not seem to strongly influence their classroom practices. On the contrary, most of the English teachers (70.5%) who responded to the 2002 SAIP questionnaire indicated that pressure from parents did not limit how they taught (CMEC, 2002a). This suggests that teachers' practices might be influenced by the knowledge that they must report to parents, but it is not evident that parents, in themselves, are a strong influence the fairness of classroom assessment in Ontario.

### *School Level Factors*

Beyond the confines of the classroom, two factors are seen in the literature that influence teachers' assessment decisions at the school level: leadership and collegial support.

*School Leadership.* In research on changing classroom assessment practices, the role of school leaders appears to be critical, especially in studies that look across schools. While there are several examples of this (e.g., Hall & Harding, 2002; Taylor, Pearson, Peterson & Rodriguez, 2005), the study by Hayward and colleagues (2004) is most pertinent to assessment for learning. Hayward and colleagues analyzed the project evaluation reports from 33 elementary and secondary schools that participated in the initial implementation of the *Assessment is for Learning* program in Scotland, and they found that the head teacher (principal) strongly impacted the change process. In schools where the head teacher did not endorse the program, the teachers felt powerless, isolated, and afraid of "rocking the boat" (p.409). Reports from these schools tended to focus on hindrances, and those that could have been resolved by the head teacher, such as an inflexibility in the timetable or a lack of staff meetings, became more obvious as the implementation of the program progressed.

Very little information about the influence of principals on classroom assessment in Ontario is available at this time. Officially, a principal's responsibility for classroom assessment involves decision-making about special cases, specifically in relation to assessment accommodations or modifications and credit recovery (OME, 2008, pp.7-ii, 9-ii). However, the practice of teacher moderation, which is essentially a form of collaborative assessment, also seems to be encouraged (e.g., OME, 2007d). As a result, some principals may be participating more actively with teachers in assessing student learning. There is

evidence that many teachers in Ontario value principals as a source of information (OCT, 2003), but none to indicate the degree to which principals and teachers collaborate in classroom assessment. As such, it seems that it would be possible for principals in Ontario, as decision-makers for special cases, to impact the fairness of an assessment, but little is known otherwise about their influence in the process of classroom assessment.

*Collegial Support.* In most of the research on system-wide assessment reform, the need for teachers to have opportunities to discuss assessment with their colleagues is strongly emphasized (e.g., Borko et al., 2000; Cheung, 2002; Hayward et al., 2004; Taylor et al., 2005). Hall and Harding (2002) are particularly emphatic on this point. They studied six schools in England over two years, and they concluded that the “quality of teaching and learning inside the classroom is strongly influenced by the quality of the professional relationships teachers have with their colleagues outside the classroom” (p.13). Studies on changing assessment practices through professional development programs also highlight the importance of collegial support (e.g., Briscoe & Wells, 2002; Lovett & Gilmore, 2003; Treagust, Jacobwitz, Gallagher & Parker, 2001). All of these studies are based on the premise that the quality of classroom assessment can be improved, but none focus on fairness.

A connection between collegial support and fairer classroom assessment is made in a report on an international research program. CERI (2005) conducted 19 case studies in 9 countries (i.e., Australia, Canada, Denmark, England, Finland, Italy, New Zealand, Scotland), and aimed to “address the barriers to suggest ways forward” (p.5) for assessment policy. Consequently, the report is solutions-oriented and focused on successful strategies used by the case study teachers. The Danish teachers found that assessing student work as a team

helped them “bring potential biases to light” (p.77). Although the CERI report offers little discussion on this point, it does suggest that fairness in classroom assessment is facilitated by collegial collaboration. Brighton (2001) tempers this conclusion, however, with the observation that even though school culture was important in her research on changing classroom practices in the eastern United States, its influence was not “significant enough to overshadow individual teachers’ responses” (p.184) in the professional development process.

Of the teachers who responded to the 2006 OCT survey, the majority (89%) indicated that they find “working collaboratively with colleagues” (p.7) to be a satisfying aspect of teaching (4-5 on 5-point scale). At present, teacher collaboration seems to be encouraged in Ontario. One document in a new set of materials released for teachers (Capacity Building Series) suggests that the quality and fairness of classroom assessment decisions can be improved by engaging in teacher moderation (OME, 2007d), and one study with 10 literacy resource teachers in Ontario does seem to support this assertion (Engemann & Gallagher, 2006). The results of the 2002 SAIP questionnaire indicate that many English teachers (61.5%) do meet on a weekly basis with colleagues to “plan lessons, units, tests, or discuss other program matters” (CMEC, 2002a, p.9). However, the degree to which these meetings focus on classroom assessment is not clear. At this point, the benefits and drawbacks of teacher collaboration, especially as method for improving the fairness of classroom assessment in Ontario, could be studied further.

### *System Level Factors*

Moving outside the school, factors at the system level that have been found in the existing literature to influence classroom assessment in different contexts include time,

assessment policies, standardized testing, professional development opportunities, and teaching traditions.

*Time.* Time is the most ubiquitous issue, and it appears to be a facilitator in classroom assessment literature only when it is accorded to teachers in an effort to change practices. In a two-year collaborative professional development-research project that focused on classroom assessment, Borko and colleagues (2000) found that the change process was facilitated by teachers having time for self-reflection, common planning, and team meetings. However, they also noted the reverse; a shortage of time hindered change. Most often, time is viewed as a factor that constrains how classroom assessment should ideally be practiced. Teachers complain about a lack of time to implement new assessment practices (e.g., Cheung, 2002; Hargreaves et al., 2002; Lock & Munby, 2000; Torrance & Pryor, 2001), or to fully assess the learning of each student (e.g., Morgan & Watson, 2002). Time is also cited as a problem in relation to the breadth of the learning expectations mandated within a system, which results in a tension between prioritizing learning and covering curriculum. One of the veteran teachers in Borko and colleagues' study worried that she would not have time to teach all of the curriculum if she used assessment for learning because she would have to re-teach what students did not understand the first time (p.291). All three of the teachers in Jeranyama's (2001) study experienced a similar situation as they attempted to integrate teaching and assessment. Because many of the students "lacked basic skills" (p.226), the teachers had to choose between moving on with the curriculum, and stopping to help students develop skills.

Teachers in Ontario also report that they face time related challenges. Almost all of the English teachers in Ontario (90.5%) who completed the 2002 SAIP questionnaire

indicated that they spend at least three hours per week, outside or regular school hours, marking student work, and they devote more time to marking than any other professional activity, including planning lessons and working with students (CMEC, 2002a). When teachers were asked on the 2004 OCT survey (n = 1000) to state their “greatest challenge as a teacher” (p.12), the most frequent response was time constraints. This is also reflected in the stress ratings that teachers gave on the 2006 OCT survey, where time constraints topped the list with 61% of the teachers rating this aspect of teaching as stressful (4-5 on 5 point scale). Given that time constraints have been found to influence assessment practices, and that teachers in Ontario report facing time constraints while devoting more time to assessing student work than any other aspect of teaching, it is possible that time constraints affect the quality of their assessments.

*Assessment Policies.* The policies established within an educational system have been found to influence classroom assessment in different ways. McMillan and Nash (2000) noted the effects of certain policies (e.g., inclusion) on teachers’ assessment decisions, but they also found that in “some cases teachers completely ignored division policies” (p.18). In the literature on assessment reform, it is clear that teachers do function as mediators in the implementation of new policies because they interpret and adapt them to mesh with their own beliefs and practices (Cheung, 2002; Tunstall, 2001). While this may result in internal consonance for individual teachers, it can be problematic for students when policies are interpreted differently and assessment practices become inconsistent across an educational system. In some studies on classroom assessment, issues arise when teachers attempt to follow conflicting policies that have been established within their educational system. In

studying teachers' interpretations of policy in England, for example, Watson (1999) noted that they faced a tension embedded in the system, specifically that the value placed on student discussion and thinking in the revised national curriculum conflicted with the emphasis on system accountability and the use of students' assessment results for that purpose.

In step with many other educational systems around the world, a standards-based approach was adopted in Ontario in the 1990's to increase the accountability of the educational system. As part of this reform, a standardized report card and curriculum for all school subjects were developed and gradually implemented across the province. Initially, these documents provided very little guidance for fair classroom assessment. An achievement chart, which is a rubric with four categories (i.e., knowledge or skills) and generic descriptors across four levels, was included in the curriculum documents, but the accompanying text was limited. Provided with vague instructions for using the achievement chart "among a number of tools" (OME, 1997, p.4), teachers were left to "work out" (p.7) the details of classroom assessment for themselves. The first reference to fair assessment was seen in 1999, in a document on program and diploma requirements where it was noted that special education students should have a "fair and equal opportunity to complete the secondary school literacy test" (OME, 1999, p.10). Eleven principles for assessment and evaluation were also listed, one of which stated that teachers' assessment strategies must be "fair to all students" (p.31). A group of researchers who were studying classroom assessment in Ontario at that time noted that teachers were faced with conflicting policies, and they argued as Watson (1999) in England had, that differing purposes for educational assessment were at the heart of the conflict (Hargreaves et al., 2002; Katz et al., 2001; Schmidt, 2001).

With a change of government in 2003, a program of ongoing curriculum revision was initiated in Ontario, which resulted in the release of new English Language curriculum for grades 1 to 8 in 2006, and for grades 9 to 12 in 2007. For an overview of how English is taught in Ontario's educational system, see Appendix A. My interest here in these documents pertains to the guidance they provide for assessing student writing. For an overview of how English is taught in Ontario's educational system, see Appendix A. The revised English curriculum reflects newer ideals for teaching and assessing writing, and many of the specific expectations given are consistent with the recommendations of established writing specialists. For example, Bright (2007), a specialist in language education who has published several resources on writing for teachers, explains that "the portfolio is the heart of the writing instruction program" (p.74). A learning expectation for writing across the intermediate and senior years in Ontario states that students will "select" and "reflect" on pieces of writing from their portfolios (OME, 2006, pp.132, 146; OME 2007a, pp.51, 65, 79, 94, 108; OME 2007b, pp.54, 70, 87, 104, 122, 139). Another practice that is strongly recommended in current texts on writing pedagogy is the involvement of students in the assessment process, which can be accomplished through collaborative development of rubrics, goal setting, or peer and self-assessment (Bright, 2007; Bromley, 2007). Again, these types of practices are frequently found in the OME's revised writing curriculum. Peer editing, for example, is recommended for students in Grade 7 and 8 (2006, pp.131, 145), and in advanced classes, such as the Grade 12 Writer's Craft, a "class-constructed rubric" is recommended for collaborative writing (2007b, p.194). Despite this updated approach, the revised English curriculum continues to say little about fairness. All three documents contain the same

singular statement about fair assessment, embedded in set of 11 principles, that was found in earlier OME policy documents. The only additional reference to fair assessment is for special education students, where the statement “fairness is not sameness” (2007a p.27; 2007b, p.28) was added to the revised secondary curriculum.

In 2008, the OME “launched an initiative to update, consolidate, coordinate and clarify policy for assessing, evaluating and reporting student achievement” (Durst & Plourde, 2008). This resulted in a new document, *Growing Success* (OME, 2008), that showed a increased attention to fair classroom assessment. The same 11 principles that were found in previous documents are repeated, but *Growing Success* also contains references to fairness in several other sections, including those on late or missing assignments, and accommodations or modifications for both special education students and English language learners. Elaboration on the concept of fairness is taken from a magazine for schools (not peer-reviewed) where fairness is interpreted in terms of group differences (Volante, 2006). Some strategies for the classroom are offered, most of which deal with tools (tests) and grading, rather than student-teacher interactions (p.2-ii). In other sections of *Growing Success*, fairness is associated with consistency. For example, the section on the achievement chart, which aims to justify the establishment of a standards-based system in Ontario, states:

In the previous system of assessment, standards varied from teacher to teacher, and school to school, and this led to results that were not always fair for all students. With criterion referenced assessment, we have well-defined standards commonly understood and consistently applied across the province. As a result, assessment of student achievement is fairer and more reliable.  
(p.3-ii)

Fair assessment is thus seen as a desirable quality in Ontario’s educational system, but the repeated association of fairness with testing and grading (e.g., pp.2-ii to 2-iii, p.4-iv) orients

the interpretation to the assessment of learning. Although it reiterates that “the primary purpose of assessment and evaluation is to improve student learning,” and it claims that this recognition . . . is one of the most important changes in policy over the past fifteen years” (p.1-I), *Growing Success* offers little guidance on fair assessment for learning.

This document is presently circulating within the educational system in draft form. Although it is available on the Internet, it has not been officially implemented, so many teachers in Ontario may not be familiar with its content. It has also become increasingly apparent that some of the OME’s assessment policies, particularly those relating to late and incomplete assignments, are problematic for teachers (Tierney, Simon & Charland, 2009; Laucius, 2009). Essentially, classroom assessment policies in Ontario are currently in the latter stages of a review process, and issues that are emerging in the implementation of some of these policies are the subject of heated discussion, not only among teachers, school leaders, researchers and policymakers, but also in the public media. As such, when this study was undertaken, policies that might impact the fairness of classroom assessment in Ontario were not only in flux, they were also viewed as conflicting by some stakeholders, and they were contested by others.

*Standardized Testing.* Research on the influence of large-scale, standardized testing on classroom assessment practices has shown mixed results. For example, teachers in Virginia and Washington state expressed concern about classroom assessments being limited to imitations of the multiple-choice format commonly found on standardized tests (McMillan & Nash, 2000; Mabry, Poole, Redmond & Schultz, 2003), whereas teachers in Alberta felt that classroom assessment practices had improved as a result of an external examination

program (McDonald, 2002).

The only system-wide assessment of students' writing in Grades 7 to 12 in Ontario is the Ontario Secondary School Literacy Test (OSSLT). The OSSLT received a negative reception when it was first administered to Grade 10 students in 2002, and in general, teachers in Ontario continue to maintain an anti-testing stance (Elementary Teachers' Federation of Ontario, 2008). For example, the 2003 OCT survey report notes that "standardized tests are a heavy duty lightning rod insofar as teachers see little to no merit in any aspect of them" (p.12). Very few teachers (11%) responding to the 2004 OCT survey agreed that provincial testing provided an accurate reflection of student learning, and most of those responding in 2005 (71%) identified standardized testing as the least effective initiative for improving student learning. Despite this, the classroom practices of most English teachers in Ontario do not seem to be greatly influenced by standardized testing. Of those who responded to the 2002 SAIP questionnaire, most (83.0%) indicated that the way they teach English Language Arts is not greatly impacted (i.e., not at all, a little) by external testing (2002a). More recently, the 2006 OCT survey results showed that only a third (31%) of the responding teachers identified "preparing students for exams / EQAO assessments" (p.12) as a stressful aspect of teaching (9<sup>th</sup> of 16 options). This suggests that although teachers in Ontario do not support large-scale, standardized testing, the impact of the OSSLT on the quality of their assessment practices may be minimal.

*Professional Development Opportunities.* The professional development opportunities that are afforded to teachers within an educational system are frequently identified as a factor that influences classroom assessment practices, especially in educational systems where new

assessment programs are being implemented. Hayward and colleagues (2004) attributed the success of the *Assessment Is for Learning* program in Scotland to its emphasis on collegial support and collaborative development among teachers. In contrast, Yip and Cheung (2005) found that a lack of professional development left many teachers ill-prepared for the “dual role” (p.161) of teacher-assessor that was expected when the Teacher Assessment Scheme was implemented in Hong Kong. Writing about the same assessment program, Yung (2001b) explicitly associates professional development with fair assessment:

There is clearly a need for teacher professional development in this vital aspect, especially on developing teachers’ analytic awareness on the issue of fairness. . . . Special attention has to be paid to how to get the philosophy and intentions behind the new form of assessment across to the teachers . . . . Otherwise, there is a danger of teachers interpreting these changes within a traditional ‘testing paradigm’ . . . . And this would be grossly unfair to all parties concerned—teachers and students alike. (p.1002)

Yung’s emphasis on the importance of professional development for fair assessment in Hong Kong suggests that professional development should be considered as a potential influence on the quality of classroom assessment in other educational contexts.

In several studies in Canada, concerns have been voiced about limitations in teachers’ assessment literacy (e.g., Hargreaves et al., 2002; McDonald, 2002). Volante and Fazio (2007) surveyed 69 pre-service teachers in Ontario, and they concluded that an increased emphasis on classroom assessment is needed to improve assessment literacy in Ontario. In comparison to other tasks, such as marking student work or planning and preparing lessons, teachers seem to devote very little time to professional development. Most of the English teachers in Ontario (84%) who responded to the 2002 SAIP questionnaire indicated that they spend less than two hours per week outside of the regular school schedule on professional

development (CMEC, 2002a). This may, however, be another area of change currently underway in Ontario. In 2005, a Working Table reviewed research literature and existing professional development practices with input from 16 educational organizations (i.e., college, councils, federations and associations), and several improvements were subsequently recommended, which included coordinating professional learning and facilitating opportunities for teachers to share their knowledge (OME, 2007c). The OCT (2008) has also established a framework for teachers' professional development. At this time, though, the degree to which teachers are engaging in professional development related to classroom assessment in Ontario is not clear. Given that assessment policies are currently under review in Ontario, and that the quality of classroom assessment is affected by teachers' assessment literacy, the opportunity for professional development in classroom assessment may be an important factor for fair assessment in this context.

*Teaching Traditions.* A final external factor discussed in the classroom assessment literature is tradition. Teaching traditions may be particular to an educational context or even a school, but they generally embody the pedagogical values of the past (historical context), and they influence how teachers view themselves and their work in the present. Field and Latta (2001) observed, based on their experiences working with pre-service teachers during field placements, that often "procedurally entrenched routine . . . forecloses on the possibility that something new might happen, that a new self might emerge" (p.891). Hayward and colleagues (2004) found that the "inherent conservatism" (p.406) of secondary school pedagogy in Scotland presented a barrier for the implementation of assessment for learning. Teaching traditions can even impede change when it is approached willingly, as Ryan (2001)

found in his action research with four secondary science teachers in Ontario. He highlights how a variety of system level factors perpetuate that traditions can limit the quality of classroom assessment practices:

This isolation appeared to narrow the scope of their teaching, making it more traditional, less creative and less satisfying. This was due in part to the system, which was controlled by time (periods, terms, semesters), and not by issues of learning. It is a system that prioritises content coverage over understanding. This need to cover content in a certain amount of time causes stress as teachers struggle to fit in the curricular content with little time for professional development, or simply talking with colleagues. This predicament results in a general lack of time or thought directed towards student achievement, assessment theory or practice. (p.388)

Although I have listed and discussed the external factors that influence classroom assessment individually and in a set order for clarity, Ryan's observations suggest that these factors are inter-related within Ontario's educational context.

In conclusion, most of the research that provides evidence about the influence of external factors on classroom assessment practices emanates from educational contexts where some type of assessment reform is underway. Often the focus is on the challenges that teachers face as they change their assessment practices. It is generally assumed that reform improves assessment practices, and with few exceptions, fairness has not been considered in the process. As such, it is not clear which external factors influence the fairness of classroom assessment, and inquiry in this area must begin with the factors that have been found to facilitate or hinder classroom assessment practices in more general terms.

### Chapter Summary

In this chapter I described and discussed assessment in the classroom environment because it is in this context that I aim to understand fairness. Educational assessment serves

multiple purposes which have evolved over time, particularly as thinking about learning changes. The multiple purposes of assessment, and differing beliefs about their use and emphasis are often central to conflicts within educational systems. Two main assessment purposes are distinguished prepositionally in current classroom assessment literature. In theory, assessment of learning functions summatively, and it involves practices (e.g., testing, grading, reporting) for which teachers have traditionally been responsible. Assessment for learning can function ipsatively, formatively, or diagnostically. Ideally, it involves teachers and students in planned and spontaneous cycles of interaction that focus on learning. Certain practices, or actions, such as sharing learning expectations with students and using assessment results to reinform teaching, are strongly associated with assessment for learning. Although the primary purpose of classroom assessment in Ontario is officially to support learning, teachers are also required to report regularly on assessments of learning using a standardized report card. Three issues relating to the multiple purposes of classroom assessment may be pertinent for fairness: teachers' roles and teacher-student relationships, the frame of reference used for assessing, and blurring in the function of assessment information. A host of factors that are external to teachers' thinking have also been found to affect classroom assessment. I discussed factors at the classroom, school, and system levels, and I related them to Ontario's educational context. Although some evidence suggests that the quality of classroom assessment may be influenced by these factors, very little is known about how they impact fairness. The next chapter focuses on how the concept of fairness is currently understood in different types of literature, including conceptual work, empirical research, and studies situated in the same educational context as my study.

### CHAPTER THREE: FAIRNESS IN CLASSROOM ASSESSMENT

*The horse trader, the jockey, the zoologist, and the paleontologist will all have their own experience of a horse. If their account turns out to be different, there is . . . no reason for assuming that the content of only one of them can be real and that the experiences of the others must be less accurate or real.*

Gert Biesta and Nicholas Burbules (2003, p.15)

Fairness is a perennial ideal in educational assessment. However, as educational beliefs and contexts change over time, the exact nature of fair assessment has been subject to ongoing discussion, which has resulted in diverse interpretations. For example, Airasian (2005) associates fairness broadly with ethical practice, whereas Volante (2006) understands it more narrowly as “reducing bias” (p.34). Related terms, particularly bias, ethics, equality and equity, are often used in conjunction or interchangeably with fairness. Szyrka (2001) offers a notably circular definition in her dissertation on equitable classroom practices. She writes that equity is “about fairness” and then defines fair assessment as “equitable assessment” (p.5). As precision and consistency are currently lacking in the terminology surrounding fairness, I have included literature using any of the above terms in relation to classroom assessment in this review. My main purpose in this chapter is to identify the aspects and interpretations of fairness that currently exist in the literature.

The balance of this chapter is organized into three sections. In the first, I explore how fairness in classroom assessment is understood in conceptual work, and in the second I look at interpretations found in empirical studies. I organized these two sections according to the type of literature because my initial readings suggested this might be telling differences in how fairness was interpreted, especially as different types of literature capture different pieces

of the conversation between teachers and researchers in education. In the third section of this chapter, I discuss current research relating to fair writing assessment in Ontario in order to understand how fairness has been interpreted in the specific educational context of this study.

### Conceptual Work on Fairness in Classroom Assessment

Although the two types of texts that I discuss in this section can both be described as conceptual, they differ significantly in their construction and intent. Principles and standards are developed by large committees to provide agreement across organizations and guide practice. Individual assessment specialists sometimes offer guidance for practice (e.g., classroom assessment textbooks), but their conceptual work usually aims to further discussion within the academic community. While their ideas may be highly regarded, and may have a significant influence on research and practice, the principles and standards that are adopted by the educational community inherently carry the weight of consensus.

#### *Principles and Standards for Fair Classroom Assessment*

Documents containing principles and standards present the ideals that are held by the educational community at the time of their development. Two key documents containing principles or standards for fair classroom assessment are the *Principles for Fair Assessment Practices for Education in Canada* (Joint Advisory Committee [JAC], 1993) and *The Student Evaluation Standards* (Joint Committee on Standards for Educational Evaluation [JCSEE], 2003). Both of these documents were produced by large committees to provide a “consensus of what is most important” (JCSEE, p.2) for the broader educational community. Although their basis is theoretical, they do not focus on defining fairness as an abstract concept, but aim instead to provide guidance on how fair assessment should be practiced in classrooms.

The *Principles for Fair Assessment Practices for Education in Canada* (JAC, 1993) was created to address issues in classroom and standardized assessment at a time when “no prior set of standards for fair practice” (p.4) in classroom assessment existed. Although the terms fair and fairness are not explicitly defined, the document sets forth the “guidelines generally accepted by professional organizations as indicative of fair assessment practice within the Canadian educational context” (p.3). Co-constructed by representatives from the provincial and territorial Ministries of Education and nine educational organizations in Canada (Canadian Education Association, Canadian School Boards Association, Canadian Association for School Administrators, Canadian Teachers’ Federation, Canadian Guidance and Counselling Association, Canadian Association of School Psychologists, Canadian Council for Exceptional Children, Canadian Psychological Association, and Canadian Society for the Study of Education), the document can be considered representative of the ideals within the educational community about fair assessment in the 1990s. Five general principles are given in a conventional assessment sequence:

- Assessment methods should be appropriate for and compatible with the purpose and context of the assessment (I).
- Students should be provided with a sufficient opportunity to demonstrate the knowledge, skills, attitudes, or behaviors being assessed (II).
- Procedures for judging and scoring student performance should be appropriate for the assessment method used and consistently applied and monitored (III).
- Procedures for summarizing and interpreting assessment results should yield accurate and informative representations of a student's performance in relation to the goals and objectives of instruction for the reporting period (IV).
- Assessment reports should be clear, accurate, and of practical value to the audience for whom they are intended (V). (pp.5-12)

Each of these principles is followed by a set of guidelines, which involve numerous aspects of fairness. Despite its relatively early publication date, this set of principles is comprehensive in its interpretation and discussion of fair assessment. Methods that are associated with learning, such as the involvement of students in assessment through conferences or self-assessment (Principles I.4, V.4) are mentioned, and the importance of considering assessment purpose is emphasized (e.g., Principle I.1). However, the principles are organized into a sequence that is most suitable for the assessment of learning, and there is very little elaboration to support fair practice in the dynamics of assessment for learning. As this document predates the emphasis on the benefits of assessment for learning in the English-language literature (i.e., late 1990's), it appears forward-looking in retrospect, but it still retains a measurement orientation that is consistent with its publication date.

The evolution of ideals for classroom assessment is quite evident in the more recent set of standards, *The Student Evaluation Standards* (JCSEE, 2003). Like its predecessor, this document was constructed collaboratively, but with input from over 100 individuals and representatives from 16 educational organizations in both Canada and the United States. *The Student Evaluation Standards* is concerned specifically with “evaluation practices used to guide the learning and progress of students” (p.xix) in classroom environments, and it aims to provide a “working philosophy for student evaluations” (p.xx) that are “ethical, fair, useful, feasible, and accurate” (p.3). For each of the 28 standards in this document, an explanation, rationale, caveats, guidelines, and list of common errors are provided, as well as description and analysis of classroom-based illustrative cases. One problem with this document is that its structure and terminology are more consistent with the language of program evaluation than

classroom assessment. This is seen in the use of the same four categories – propriety, utility, feasibility, and accuracy – that organize a parallel document produced for program evaluation (Patton, 2002). It is also reflected in the use of the term *formative evaluation*, which is defined as “evaluation conducted while a creative process is under way, designed and used to promote growth and improvement in a student’s performance or in a program’s development” (JCSEE, p.228). While the term formative evaluation was used in earlier literature on student assessment (e.g., Bloom et al., 1971), the term *formative assessment* has been preferred in classroom assessment literature for decades (e.g., Sadler, 1989). Although there may be similarities between evaluating a program and assessing student learning, program evaluation and classroom assessment have evolved as distinct fields over recent decades. This mismatch between the structure and terminology of *The Student Evaluation Standards* (JCSEE, 2003), and current literature is unfortunate in that it suggests that the authors were more familiar with program evaluation than classroom assessment.

Although *The Student Evaluation Standards* (JCSEE, 2003) provides more detail and discussion than the earlier *Principles* (JAC, 1993), fairness is still not defined at any point. The word fair is used interchangeably in the introductory text with related assessment qualities, such as sound and ethical (see for example, pp.2-3). The words fair or fairness appear explicitly in the statements for only three standards:

- **Appropriate Policies and Procedures:** Written policies and procedures should be developed, implemented, and made available, so that evaluations are consistent, equitable, and fair. (P2, p.33)
- **Rights of Students:** Evaluations of students should be consistent with applicable laws and basic principles of fairness and human rights, so that students’ rights and welfare are protected. (P5, p.51)

- Bias Identification and Management: Student evaluations should be free from bias, so that conclusions can be fair. (A7, p.167)

However, there are multiple references to fairness in the guidelines and cases that explain several other standards. In all, a total of 21 standards are listed under “ensuring fairness” (p.xx) in a thematic table of contents (see Appendix B), which supports the text’s claim that fairness is one of the “major components . . . and issues in student evaluation” (p.6). In essence, fairness is linked in these standards with a wide range of qualities that are currently recommended for classroom assessment tools, tasks and results (e.g., transparent, appropriate), and fair classroom assessment is clearly presented as a process that should support student learning.

In both the *Principles for Fair Assessment Practices for Education in Canada* (JAC, 1993) and *The Student Evaluation Standards* (JCSEE, 2003) there is considerable emphasis on the transparency of assessment criteria and procedures. While the newer standards dwell more on the consequences of assessment, both documents emphasize the need for sensitive interpretation, careful summarization, clear and confidential communication of results, and written assessment policies. In this light, both of these committee documents are concerned with aspects of fair classroom assessment that extend beyond teacher-student interactions and involve other stakeholders at the school and system levels. Although they were produced a decade apart, these two documents both emphasize that the process of classroom assessment should be explicit and clear. This not only suggests that it is accepted within the educational community that fair classroom assessment may be achieved through openness, but also that fairness centers on communication among stakeholders. Overall, the interpretation of fairness in the *Principles for Fair Assessment Practices for Education in Canada* and in *The Student*

*Evaluation Standards* far exceeds the breadth of the psychometric definition of fairness given in the *Standards for Educational and Psychological Testing* (JCSEPT,1999). While the fairness of a test may be firmly established through standardization and statistical analyses, fairness is presented in the principles and standards for classroom assessment as a nebulous and complex concept. Fairness issues pervade all phases of classroom assessment, and concern all stakeholders in the process.

### *Fairness in Classroom Assessment Theory*

The journal articles and book chapters on classroom assessment theory are similar to the committee documents discussed in the previous section in two ways. First, fairness is often discussed without explicit definition in classroom assessment theory. Some authors ask the reader to judge the fairness of a practice (e.g., Speck, 1998), while others associate fairness with related concepts, such as equity and ethics (e.g., Weeden et al., 2002 ), but only a handful provide operational definitions (Airasian, 2005; Camilli, 2006; McMillan, 2007; Suskie, 2002). A second similarity is that the discussion often relates to measurement qualities, which reflects the technical orientation of the *Standards for Educational and Psychological Testing* (JCSEPT, 1999). This is not surprising given, as Shepard (2000) notes in the historical overview that introduces her discussion about the role of assessment in teaching and learning, that “beliefs about the nature of evidence and principles of fairness” have been heavily influenced by the “dominance of objective tests in classroom practice” (p.5). Although it has been suggested that assessment for learning is “a process of communication guided by teaching and learning goals, rather than a process of measurement” (Allal & Pelgrims Ducrey, 2000, p.146), the reconceptualization of classroom assessment

theory that is currently underway is, nevertheless, strongly influenced by measurement theory. In many texts, fairness is associated with the two key measurement concepts of validity and reliability, and the necessity of valid information for fair assessment decisions is often emphasized (e.g., Airasian, 2005; Camilli, 2006; Volante, 2006; Weeden et al., 2002; Whittington, 1999).

Despite its roots in measurement theory, current research on classroom assessment tends to involve qualitative methods, and it is approached from a variety of theoretical perspectives (Tierney, 2006). Reflexive sociology, sociocultural theory, and critical theory have all been used to frame classroom assessment research (e.g., Ecclestone & Pryor, 2003; Pryor & Crossouard, 2008; Tierney, 1998/2005). Several researchers have drawn on post-structural philosophy to discuss the ethics or equity of classroom assessment practices, and they have raised concerns about power in assessment relationships (Gipps, 1999; Schendel & O'Neill, 1999; Torrance & Pryor, 1998). Schendel and O'Neill (1999), for example, focused on the use of student self-assessment for summative purposes, and they argued that "it may ultimately be a means by which traditional exercises of power of teachers over their students are reaffirmed as students implicate themselves" (p.206). This is similar to Brookfield's (2001) concerns about power and "liberatory practices" (p.22) in adult classrooms, and it challenges the assumption that participatory forms of assessment are necessarily fairer than conventional methods. Many of the ideas put forth in classroom assessment literature over the past few decades indicate that theoretical discussion amongst researchers about fair assessment now goes beyond technical concerns to tackle some of the fuzzier issues that arise in interactions between students and teachers in classrooms.

The expansion of assessment theory from pure measurement is also evidenced in the number of texts that mention fairness in relation to a recognition of social and cultural values in the content, criteria, processes or expectations of assessment (Airasian, 2005; Camilli, 2006; DeLain, 1995/2005; Gipps, 1999; Jimenez, 2004/2005; Lam, 1995; McMillan, 2007; Schendel & O'Neill, 1999; Suskie, 2002; Tierney, 1998/2005; Volante, 2006; Weeden et al., 2002; Whittington, 1999). A significant departure from the psychometric definition of fairness is seen here. In standardized testing, construct-irrelevant differences are controlled to permit comparisons, and test items are carefully examined for differential functioning (JCSEPT, 1999). In contrast, student diversity in the classroom is ideally nurtured and respected. While some theorists stress the need to avoid cultural bias or stereotypes (Airasian, 2005; Camilli, 2006; McMillan, 2007; Volante, 2006), Robert Tierney (1998/2005) argues that "culture-free tests" (p.29) are an impossibility because learning (literacy) is "inextricably connected to cultural background" (p.30). He proposes that educators aim for "cultural sensitivity" (p.30) in designing assessments so that they acknowledge and build on students' backgrounds. This is similar to the association of fairness and cultural awareness, which is expressed by others in the literature on classroom assessment (e.g., DeLain, 1995/2005; Gipps, 1999; Jimenez, 2004/2005; Schendel & O'Neill, 1999).

Acknowledging diversity in the classroom raises questions about the opportunities students have to learn. Fairness is often discussed in relation to opportunity to learn in classroom assessment theory, but some differences in perspective are seen (Airasian, 2005; Camilli, 2006; DeLain, 1995/2005; Drummond, 2003; Gipps, 1999; Jimenez, 2004/2005; Lam, 1995; McMillan, 2007; Shepard, 2000; Speck, 1998; Suskie, 2002; Weeden et al., 2002;

Whittington, 1999; Wormeli, 2006). For example, McMillan and Drummond both discuss opportunity to learn as an alignment of teaching, learning and assessment, but Drummond focuses on children's rights from a social justice perspective, and McMillan emphasizes the need for adequate time and materials from a more technical perspective. In comparison to the prevailing psychometric definition of fairness, where opportunity is defined narrowly in terms of test preparation (JCSEPT, 1999), opportunity to learn in classroom assessment theory is more broadly associated with the quality of student-teacher interactions, access to resources, and the teachers' understanding of students' diverse backgrounds (DeLain; Drummond; Jimenez, Shepard; Weeden et al.).

A similarity in the discussions about fairness in standardized testing and classroom assessment contexts relates to equality of outcome. Although the idea that assessment results for different groups must be equal has been "almost entirely repudiated in the professional testing literature" (JCSEPT, 1999, p.74), group differences are still thought to signal a need for further test analysis. Weeden and colleagues (2002) also express this view of group differences as a red flag in their discussion about the lack of equal opportunity in the English educational system. They argue that inequality of outcome should result in educational intervention, especially if it is persistently associated with gender, social economic status, and racial or ethnic groups. Camilli's (2006) position is similar in a chapter on test fairness that includes some discussion about fairness in classroom assessment. He argues that it is "unreasonable to define fairness as the elimination of individual differences" (p.248), and he recommends the use of assessment for learning to enable all students.

Some authors associate fairness with equal opportunity to demonstrate learning

(McMillan, 2007), while others recommend multiple or varied measures (Airasian, 2005; Camilli, 2006; Shepard, 2000; Suskie, 2002; Volante, 2006; Weeden et al., 2002). Table 2 shows the elements included in five definitions that aim to operationalize fairness for the classroom. Airasian emphasizes that assessment is a “human activity” (p.20) that should be guided by general ethical standards (e.g., respect, honesty), and he describes six additional aspects of fairness, which relate essentially to transparency, opportunity to learn, stereotypes and bias. Camilli begins by noting that the assessment purpose should be clear, and he discusses six additional practices for fair assessment. McMillan also describes six components for fairness in his textbook on classroom assessment, and Suskie suggests seven steps for fair assessment in higher education. Volante offers an interpretation that is similar to the psychometric focus on group differences, and he suggests nine strategies for fair classroom assessment. Similarities are seen in these interpretations of fairness. For example, four authors begin with an element related to transparency, and they all include suggestions for providing students with opportunity to learn.

All of the operational definitions provided by Airasian (2005), Camilli (2006), McMillan (2007), Suskie (2002) and Volante (2006) also recommend that teachers think about some aspect of their assessment practices. On this point, these definitions mirror classroom assessment theory as a whole, where recommendations for teachers to think about personal biases, expectations, roles, and power dynamics are abundant (e.g., DeLain, 1995/2005; Drummond, 2003; Schendel & O’Neill, 1999; Shepard, 2000; Speck, 1998; Weeden et al., 2002; Whittington, 1999, Wormeli, 2006). However, they also diverge

Table 2

*Aspects, Components, Strategies, Steps and Practices for Fair Classroom Assessment*

		Texts			
Steps	Airasian (2005)	Camilli (2006)	McMillan (2007)	Suskie (2002)	Volante (2006)
1	Informing students about teacher expectations and assessments before beginning teaching and assessment	Clear and reasonable assessment criteria	Student knowledge of learning targets and assessments	Have clearly stated learning outcomes	Examine how your classroom complements or conflicts with the school experiences and assessment techniques which are familiar to recent immigrants and indigenous people
2	Teaching pupils what they are to be tested on before assessment	Equity in assessment and instruction [focused on learning]	Opportunity to learn	Match your assessment to what you teach and vice versa	Develop test questions and other assessment items that reflect the multicultural and multiethnic composition of your school, district, region, province, country
3	Not making snap judgments and identifying pupils with emotional labels (e.g., disinterested, at-risk, slow learner) before you have spent time with them	Opportunity to learn	Prerequisite knowledge and skills	Use many different measures and many different kinds of measures	Utilize gender-neutral terms within tests, quizzes, and other forms of assessment
4	Avoiding stereotyping pupils (e.g., "He's just a dumb jock," "Kids from that part of town are troublemakers," and "Pupils who dress that way have no interest in school")	Sensitivity and construction of assessments	Avoiding student stereotyping	Help students learn how to do the assessment task	Provide modifications to tests and other assessment measures for students with exceptionalities and those with limited proficiency in English

Texts

Steps	Airasian (2005)	Camilli (2006)	McMillan (2007)	Suskie (2002)	Volante (2006)
5	Avoiding terms and examples that may be offensive to students of different gender, race, religion, culture, or nationality	Multiple measures	Avoiding bias in assessment tasks or procedures	Engage and encourage your students.	Take steps to reduce the effects of test-wiseness on academic achievement— design carefully constructed test items, utilize multiple-formats, provide reviews prior to testing
6	Avoiding bias toward pupils with limited English or with different cultural experiences when providing instruction and constructing assessments	Modelling fairness	Accommodating special needs	Interpret assessment results appropriately	Involve students in the development of evaluation criteria— develop rubrics with the assistance of students
7				Evaluate the outcomes of your assessments	Adopt a range of formative and summative assessment strategies that encompass different ways of demonstrating task mastery— speaking, writing, performing
8					Balance the weight given to different types of traditional and authentic performance-based assessment data when arriving at final course grades
9					Most importantly, reflect on preconceived notions that may affect the mark/grades you assign to particular groups of student

considerably their emphasis and wording on this point. While some focus on what to avoid, others recommend more active reflection. Camilli takes this idea the furthest in suggesting that teachers take an “overt approach to grading” (p.248) to model fairness for students. In addition, the variety in how these authors preface and present fairness (i.e., aspects, components, strategies, steps and practices), and unique inclusions by some authors results in considerable difference. These texts are helpful in providing concise tips for improving the fairness of classroom assessment. However, the brevity of justification for the various elements included in these definitions, taken into account with their differences, suggests that the reconceptualization of fairness for the classroom environment could bear much further investigation. None of these authors support the elements included in their definitions with empirical research. While avoiding offensive test content, involving students in creating rubrics, using multiple assessments, and modeling fairness for students may impact the fairness of classroom assessment, evidence to this effect is not offered.

Overall, the interpretations and aspects of fairness discussed in classroom assessment theory are varied, but in comparison to the committee documents, fairness is not viewed as comprehensively. Many of the aspects of fairness that appeared in the principles and standards are not mentioned at all, such as feasibility in terms of the scope of the assessment or balance in interpreting strengths and weakness. While there is discussion around the need for appropriate, unbiased assessments that are sensitive to students’ differences, there is scant mention of the importance of written assessment policy. Recommendations for the communication of results, the protection of privacy, or the use of multiple evaluators are rarely given. Instructions for determining fairness through statistical procedures are also

notably absent. In sum, the emphasis in classroom assessment theory is mainly on sharing expectations and criteria with students (i.e., transparency), and providing students with opportunities to learn and to demonstrate learning. With frequent acknowledgment of the influence of social values and cultural expectations, recommendations by classroom assessment theorists tend to stress the importance of teacher reflection for fair assessment.

### Empirical Studies Relating to Fair Classroom Assessment

The conceptual literature represents thinking about how fair classroom assessment should be practiced. However, principles, standards and theory do not provide insight into how fair assessment is actually practiced in classroom environments. In this section, I consider three groups of empirical research. The first involves inquiry into the moral and ethical world of teaching. The second contains research on classroom assessment where fairness issues emerge in the results or discussion, and the third group consists of research specifically focused on the fairness, equity or ethics of classroom assessment.

#### *Fair Classroom Assessment in Studies on the Ethics of Teaching*

I discuss three monographs in this section that draw on classroom-based research as a source for scholarly thought and discussion. These texts go well beyond reporting empirical results, and each is a major work on morality and ethics in teaching. Although this type of inquiry is broadly concerned with all classroom interaction, and does not aim to study fair classroom assessment in particular, these three texts illuminate fairness issues that arise as teachers and students engage in assessment interactions in classroom environments.

In *The Moral Dimensions of Teaching*, Buzzelli and Johnston (2002) look at classroom interaction through the lenses of language, power, and culture. They draw on their

own experiences as teachers and scholars, excerpts of classroom discourse from several of their studies, and the writings of other educators, philosophers and researchers for their analysis. Although their work is situated in the daily life of classrooms, the two activities that Buzzelli and Johnston identify as assessment are summative, and both are located in the chapter on morality and power in the classroom. They describe grading as a morally complex act and a “particularly stark instance of the use of power to regulate, categorize, and legitimate students” (p.63). Examinations are seen in the same light, leading to issues of trust and control, and the imposition of assessment practices on students. Buzzelli and Johnston note that ranking students for “administrative convenience” (p.66) is at odds with, and can take precedence over supporting student learning. In this way, they highlight the moral tension between the purposes of assessment, but without recognizing the existence of assessment for learning. Although Buzzelli and Johnston’s narrow definition of assessment limits the conclusions they draw on the subject, it does not lessen their contribution as far as understanding the moral nature of classroom interactions. As much of their discussion about teaching as a “socially negotiated activity” (p.122) involves “cultivating moral sensibilities” (p.129), it is valuable for thinking about fairness in assessment for learning, where teaching, learning and assessment are tightly integrated.

. *The Moral Life of Schools*, by Jackson, Boostrom and Hansen (1993), also offers insight into fairness in classroom assessment while being more generally concerned with the morality of “what goes on in classrooms” (p.xi). The authors discuss moral instruction (manifest) and morality in practice (latent) based on an ethnographic inquiry involving the observation of 18 elementary and secondary teachers and their students in a variety of

classroom activities in the United States over a period of two and a half years. They note the pervasiveness of social justice issues in a “cluster of assumptions” (p.28) that structure classroom interactions. Although Jackson and colleagues do not identify these as assessment interactions, many of the assumptions they identify involve practices in assessment for learning (e.g., teachers will ask fair questions) or assessment of learning (e.g., all students’ work will be graded on a similar basis). In this discussion fairness is interpreted as a quality of the interactions within the classroom environment, but Jackson and colleagues also describe fairness as a human quality held by individuals.

Of all the moral qualities a teacher might possess, a habit of being fair is surely one of the most highly praised. The rules of fairness call for treating all students alike, at least insofar as granting favors and privileges is concerned, although they usually allow special awards to be given to acknowledge outstanding performance of some kind. (p.216).

While equal treatment may seem fair, it can create a larger moral issue, which Jackson and colleagues illustrate with several examples of classroom (assessment) interaction. In one, a teacher responds with the same comment after each student reads aloud, and the “worth of his praise” (p.88) is reduced because he does not discriminate between the students’ performances. The absence of critical feedback from another teacher on a writing assignment leaves her students “academically shortchanged” (p.153). In their efforts to be fair, these two teachers give their students equally positive feedback, which voids it of any meaning that would otherwise support learning. This problem is akin to a moral issue raised by Buzzelli and Johnston (2002). They argue that a teacher’s “failure to assess” (p.66) can ultimately lead to “devaluing or disregarding” (p.66) student learning. Thus positive feedback lavished on students indiscriminately not only robs those who need constructive feedback to direct

further learning, it also communicates to students that the quality of their performance is of little import. With these observations in mind, it becomes apparent that treating students equally during assessment for learning interactions does not guarantee fairness.

Campbell (2003) also discusses several issues relating to the equal treatment of students in her work on the nature of ethical knowledge in teaching. She refers to her book, *The Ethical Teacher*, as a “culmination” (p.xvi) of conceptual work and empirical inquiry. More specifically, it draws on field studies that she and three other researchers conducted with elementary and secondary teachers (n = 12) in five schools in Ontario. Like Buzzelli and Johnson (2002) and Jackson and colleagues (1993), Campbell notes that treating students equally can become morally problematic. The example she gives is somewhat different, though, in that it involves uniformly low expectations rather than indiscriminately positive feedback. When one of the secondary school teachers in her study is informed by a school administrator that the current student failure rate is too high, he responds that applying lower expectations “would really do an injustice to the kids who are very capable who are put out in a competitive situation” (p.73). This teacher explains that “when you sell out your top students, then you’re not doing much good at the lower end either. That to me is a moral situation” (p.74). Unlike the teachers who seemed naively unaware of the moral complexities at play a decade earlier in Jackson and colleagues’ (1993) work, many of the teachers in Campbell’s study were quite conscious of, and articulate about the ethical dilemmas they face in their classroom practices.

Campbell emphasizes the fundamental nature of fairness in professional ethics. She also notes, though, that “the interpretation of what fairness actually means in practice is often

the source of confusion and dispute” (p.30), which suggests that consensus about the characteristics of fair classroom practice may be limited. Of the three texts discussed in this section, Campbell treats the concept of fairness most extensively, and several aspects of fairness (e.g. transparent, relevant, unbiased) that were noted in classroom assessment theory are mentioned by the teachers in her study. However, much of the concern about fairness centers on the tension between treating students equally and differentiating according to their educational needs.

Most teachers accept that fairness is best achieved when they are equally attentive to each student’s capacity and needs; as needs differ, the level of attention differs as well. Widely differing treatment, however, can become very unfair regardless of the good intentions or motive behind it, and the struggle for the ethical teacher is to be ever conscious of balancing the need to be fair to individuals and the need to be fair to the group. An added complexity is signalled in situations where other moral principles, those relating to the will to be kind, caring, and compassionate and those compelling us to be honest and trustworthy complicate the moral pursuit of what is just and fair. While this is often experienced by teachers most poignantly in their professional obligation to evaluate students, such ethical concerns pervade all aspects of school life. (pp.32-33)

Although it is not clear exactly which or how many teachers Campbell includes in “most,” her discussion highlights the complexity of fairness in the classroom environment. In all three of the texts discussed in this section, teachers are “moral agents” (Campbell, p.37) who negotiate, whether consciously or not, between conflicting demands, such as the different purposes of assessment, or the needs of individuals and interests of groups. Overall, fairness is an essential quality of both individuals and interactions in the classroom, and it is closely associated with other moral qualities such as honesty and respect. The focus is less on fairness as a characteristic of assessment instruments or processes, and more on the influence of human morality and ethical knowledge in classroom dynamics. As such, these texts raise

many questions for educators and educational researchers. What is most important for my work, though, is that fairness is seen to be in the hands of teachers who strongly influence the nature of the environment in which learning and assessment take place, regardless of their degree of ethical awareness. That issues relating to fair assessment emerge repeatedly in these studies of classroom life provides further argument for research that consciously undertakes inquiry into the ethical dimension of classroom assessment.

### *Fairness as an Emerging Issue in Classroom Assessment Research*

Classroom assessment research involves a wide range of topics, from interests in specific tools and methods to broader concerns about the influence of factors such as teacher education, policy reform, or external testing. In this section, I discuss seven studies that I collected over the course of several years as I became aware that fairness issues often emerge in the discussion sections of empirical research on classroom assessment. Fairness is not the main research topic in any of these studies, but they all contribute in some way to our knowledge about fair classroom assessment.

Most of the studies in this collection involve the concerns or beliefs of practicing teachers (Brighton, 2001; Eggen, 2004; Gummer, 2000; Ryan, 2001; Yip & Cheung, 2005), but the perspectives of different stakeholders, including pre-service teachers (Graham, 2005) students (Dann, 2002) and are also heard. Students' perspectives are particularly helpful for understanding the influence of power dynamics on fairness in classroom assessment. Dann (2002) describes a case study that she conducted in a small elementary school in England where the students in one class (Year 5) were engaged in a writing project. The students' grades for the project were supposed to be negotiated through a process that involved self,

peer, and teacher assessment. Although most of the students (11 of 15) said they felt the resulting grades were fair, Dann observed that they were reluctant to challenge the teachers' assessment. During their interviews with Dann, most of the students indicated that the final grade was "more of an imposed" (p.100) grade, than a negotiated grade. This harks back to the concerns expressed by Schendel and O'Neill (1999) about the fairness of participatory forms of classroom assessment.

Teachers' reflection about classroom assessment, which is also emphasized as an aspect of fairness in the conceptual literature, is seen in this collection as well. The difference, though, is that while all of these studies contain evidence of teachers reflecting on their assessment practices, many of the authors conclude that this reflection is somehow lacking. Three doctoral dissertations on classroom assessment provide examples. Brighton (2001) used a grounded theory approach to study the internal factors (i.e., beliefs and conceptions) that influenced 12 teachers (4 schools) as they engaged in professional development about differentiated instruction and assessment. She found that the teachers' "misunderstandings and confusion about equity and fairness . . . served as roadblocks" (p.129) in their professional development. Gummer (2000) analyzed the National Science Education Standards (USA) and worked with an experienced secondary school teacher to write a case study about the (in)consonance of assessment policy and practice. Although Gummer concluded that the teacher's concerns about fairness and equity in assessment should "not be dismissed" (p.266), she pointed out that the teacher failed to apply many assessment standards. Ryan (2001) emphasized the benefits of participatory research in his study with secondary teachers in Ontario, but he also commented on the limitations of his

participants' assessment knowledge. While the teachers in all three of these studies do seem to reflect on the fairness of their assessment practices, they are generally portrayed by the researchers as being hampered by gaps in their knowledge about classroom assessment.

Despite their limited assessment literacy, many of the teachers portrayed in these studies clearly express strong convictions about classroom assessment. The notion that consistency improves the fairness of classroom assessment is frequently heard (Brighton, 2001; Eggen, 2004; Graham, 2005; Gummer, 2000; Ryan, 2001; Yip & Cheung, 2005). For example, Yip and Cheung surveyed biology teachers ( $n = 351$ ) about a new assessment system in secondary schools in Hong Kong. They explain that although only one survey item (of 25) asked about the fairness of the new system, the teachers often expressed concerns about fairness in their responses to open-ended items, particularly in relation to the inconsistency of the assessment tasks and procedures used by teachers across schools. In most of the other studies, though, teachers concerns about consistency center on their own assessment practices. Two secondary science teachers in different studies, one in Eggen's (2004) ethnographic study in Norway and the other in Ryan's (2001) participatory action research in Ontario, both refer to fairness in describing similar grading processes:

When correcting tests I always view the first five students again to adjust the grades to what I have given to the last students. Then it is fair because I have a tendency to be stricter in the beginning and then little by little adjust the grading when evaluating what the other students have answered. (Eggen, p.345)

I sometimes go over the paper 3 or 4 times because I don't want to be unfair to the students. When you evaluate something it's only human nature. I mean, I just try to sit down and do it all at once so I'm in the same mood. It's just human nature, you can't be perfect and your mood affects the student's mark. (Ryan, p.173)

Both of these teachers equate fairness with the consistent application of evaluative criteria,

and they attempt to control the degree of variance in their judgments. In measurement terms, these teachers' interpretations of fairness relate to personal bias (i.e., generosity, severity, central tendency, halo effect) in rating (Linn & Gronlund, 2000).

Some teachers' concerns about consistency relate to the tasks or tools that they use to assess student learning. Jan, the teacher in Gummer's (2000) case study, taught several chemistry classes at the same level, so she developed multiple versions of her tests to militate against the "grapevine" (p.190) between classes. The equivalency of the test items was important to Jan because she believed that for the different versions to be fair, they needed to be equally difficult for all students. The middle school teachers that Brighton (2001) worked with expressed similar ideas about maintaining the equivalency of assessment tasks for all students. They strongly believed that "equity and fairness for students means all students do the same thing" (p.129), and they resisted the differentiated assessment practices that were promoted through Brighton's project. One teacher was particularly concerned about the fairness of differentiation because she saw it as "spoon feeding" (p.70) or "watering down activities" (p.130). She felt this lowered expectations, which she believed was a disservice to her students. This is similar to the moral issue identified earlier by Buzzelli and Johnston (2002), Campbell (2003), and Jackson and colleagues (1993) where lowering expectations to the degree that all students can succeed ultimately shortchanges students and robs them of the opportunity to learn.

In this collection of studies on classroom assessment, differentiation is usually discussed in tension with the need to maintain consistency, either in the tools or the processes of assessment. However, a few studies suggest that there may also be a relationship between

flexibility and fairness in classroom assessment. One example is from a two-year research project in a teacher education program in the United States where a team of researchers studied the changes in pre-service English teachers' (n = 38 in 2 cohorts) assessment theories and practices (Graham, 2005). The pre-service teachers found writing assessment to be challenging in general, and concerns about fairness emerged in the reflective responses that they wrote at different points in their program. Although they had questions about various aspects of using rubrics to assess student writing (e.g., how to involve students in selecting criteria), concerns about fairness centered on identifying appropriate criteria and providing constructive feedback for weaker writers. For these pre-service teachers, learning to assess meant learning to adjust assessment so that it supported learning for particular students.

In Graham's (2005) study fair assessment is linked to the provision of equitable opportunity for students to learn. More often in this collection of studies, though, opportunity to learn is discussed in terms of equality. For example, one of the teachers in Eggen's (2004) study believed that to be fair, his assessments must be based on the textbook, rather than on classroom activities, because all students have equal opportunity to study the textbook. For the teacher in Gummer's (2000) study, being "equally unfair to everyone" meant that the "playing field" (pp.189-190) was level. The importance that many of these teachers place on equal opportunity could stem from their focus on assessment of learning. Although teaching, learning and assessment are linked by the researchers, only a few teachers express interest in using assessment for learning (e.g., individuals in Dann, 2002; Eggen, 2004; Graham, 2005; Gummer, 2000). Overall, a narrower understanding of fairness is seen in this collection than in either the conceptual literature, or in the ethical inquiries. Assessment

qualities that are emphasized in classroom assessment theory (e.g., transparency) are not mentioned here, and neither are any of the administrative issues (e.g., protection of privacy in the assessment process) that appeared in the principles and standards. The teachers in these studies voice concern about their assessment practices, but they do not seem to recognize the values embedded in the assessment criteria and processes that they use. Their interpretations of fairness relate more specifically to the consistency of assessment criteria and the equal treatment of students in the assessment process. While this collection of studies on classroom assessment offers glimpses of fair and unfair practices, it does not fully open the window on fairness as a quality of classroom assessment. The fairness issues that emerge here do, however, indicate the need for more focused research on fairness in classroom assessment.

#### *Research Targeting Fairness, Equity or Ethics in Classroom Assessment*

Despite searching through multiple databases (i.e., ERIC, Scopus, Proquest Dissertations, PsycInfo and Scholar's Portal) on several occasions, I uncovered relatively little empirical research specifically on fairness in classroom assessment. Eight studies that focus on fairness, equity or ethics in classroom assessment are included in this section. As they are closest to my work in terms of the subject of study, I describe each one below in methodological terms, and I have included an overview of the purpose, participants, data collection and educational context for each in Table 3. Following this information, I discuss the aspects and interpretations of fairness seen in these eight studies in relation to the other literature reviewed in this chapter.

Bursuck, Munk and Olson's (1999) study is unique in this group because they

Table 3

*Empirical Research on Fairness, Equity and Ethics in Classroom Assessment*

Authors (Date)	Purpose	Participants	Data Collection	Educational Context
Bursuck, Munk & Olson, 1999	To determine students' perceptions about fairness of grading adaptations and identify student factors that influence these perceptions.	275 students (stratified by grade and GPA) in 1 school; 15 of 275 were formally identified with learning disabilities	Survey (pilot-tested and reviewed by special education teachers) and follow-up interviews (11)	No subject specified Secondary Education Midwestern USA
Green, Johnson, Kim & Pope, 2007	To determine teachers' ethical positions on classroom assessment practices.	169 teachers (114 pre-service and 55 inservice) taking classroom assessment courses in 2 programs	Survey (based on principles & standards, pilot-tested, administered online)	Teacher Education Southern USA
Morgan & Watson, 2002	To discuss the fairness and equity of teachers' assessments based on two studies by the authors.	Study A: 1 teacher Study B: 11 teachers from 5 schools	Study A: Observation and interviews, field notes Study B: Think aloud with samples of students' work	Mathematics Intermediate & Secondary Education UK
Phillips, 2002	To question the fairness of equal treatment of students in the teacher-student interactions of classroom assessment	1 teacher	Reflective case (based on researcher's teaching experience)	Mathematics-Algebra Intermediate Education Western USA
Szpyrka, 2001	To explore the relationship between classroom assessment practices and equitable assessment practices.	6 teachers in 6 schools	Questionnaire, observation, classroom profile (all adapted from existing instruments), interviews and assessment artifacts	Science Intermediate Education Southeastern USA

Authors (Date)	Purpose	Participants	Data Collection	Educational Context
Watson, 1999	To examine differences in the epistemological and pedagogical assumptions underlying teachers' informal assessment practices.	30 teachers in 3 local school authorities	Observation and interviews over 3 years	Mathematics Elementary & Secondary Education UK
Yung, 2001	To examine teachers' views about fairness in a school-based assessment scheme.	3 teachers	Observation (4 lessons per teacher) and interviews	Science - Biology Secondary Education Hong Kong
Zoeckler, 2001	To understand the moral dimensions of grading.	13 teachers in 1 school department	Interviews (multiple repeated over 1 year), grading artifacts (e.g., school guidelines, teachers' gradebooks), field notes	English Secondary Education Northeastern USA

surveyed and interviewed secondary school students (n = 275, 11 interviews) in the United States about their perceptions of grading adaptations for final report cards. The seven other studies all involve teachers as participants. Green, Johnson, Kim and Pope (2007) surveyed teachers and pre-service teachers (n = 169) who were taking classroom assessment courses in two university programs in the United States. Their survey contained 36 short scenarios based on existing principles and standards (JAC, 1993; JCSEE, 2003), anecdotes related by teachers, and the authors' own teaching experiences. The response options were dichotomous (ethical or unethical), and the survey aimed to gather information about the ethics of specific classroom assessment practices. This work was pioneering in the sense that it was the first to investigate teachers' ethical positioning specifically in relation to classroom assessment, and it provides considerable information about the preparation and administration of standardized tests (in the classroom), multiple assessment opportunities, communication about grading, grading practices, teacher bias, and confidentiality. However, the survey design was criticized when the results were presented at an educational research conference. Brookhart (April 11, 2005), a specialist in classroom assessment research who was the session discussant, noted that the brevity of the scenarios left so much open to interpretation that the reliability of the forced responses was suspect. Essentially, Brookhart argued that ethics of classroom assessment are sufficiently complex and contextual in nature to belie a simple survey with dichotomous response options. Morgan and Watson (2002) avoided this problem with the use of different types of information (e.g., classroom observations, think aloud sessions) gathered from two of their previous studies in England. They consider the fairness and equity of teachers' interpretive judgments of student learning in mathematics. Phillips (2002) turned

the lens on her own assessment practices, and she questioned the need for equal treatment in the teacher-student interactions of classroom assessment. Although this 'case' is but a brief account, it provides insight in the critical reflection of a middle school teacher who journeyed on to become an educational researcher and university professor in the United States.

Szpyrka (2001), Watson (1999), Yung (2001b) and Zoeckler (2001) all used data collection techniques that are associated with qualitative methods, such as classroom observation, interviewing, and extended field visits. Szpyrka (2001) also adapted several existing instruments, including a rubric designed to profile an equitable classroom (Rose, Kolb & Barra Zuman, 1991), in a doctoral dissertation that explored the relationship between assessment practices and equitable practices in six teachers' classrooms in the United States.

Watson (1999) was sufficiently concerned about the impact of a decade of reform in England to meet with 30 teachers over three years, and she looked at differences in the epistemological and pedagogical assumptions underlying their informal assessment practices.

In an article that focuses on the issue of fairness in classroom assessment, Yung (2001b) discusses the views of three secondary science teachers who participated in a larger study on the implementation an assessment program in Hong Kong. The final work in this section is a doctoral dissertation by Zoeckler (2001). Using a framework based on the moral dimensions of schooling by Jackson and colleagues (1993), Zoeckler worked with 12 English teachers in a secondary school in the United States to better understand the moral dimensions of grading.

These eight studies share several similarities with the conceptual and empirical research discussed in the previous sections. The importance of opportunity to learn for fairness comes to the forefront, and once again it is understood in different ways. Sometimes

the need for students to have opportunity to learn is in the present. One example of this is the opportunity that students are given to prepare for a large-scale summative assessments (e.g., Yung, 2001b; Zoeckler, 2001). Others are more concerned with the impact of teachers' assessment decisions on students' opportunity to learn in the future, which relates to the consequences of assessment. Morgan and Watson (2002), for example, emphasize the "interpretive nature" (p.103) of all assessment, including what they define as formal (i.e., large-scale performance assessment of learning) and informal (i.e., interaction in assessment for learning) assessments, and they conclude that both can have long-lasting consequences for students' future opportunities to learn.

Often just beneath the issue of opportunity to learn lies the more contentious conflict between consistency and differentiation. This is evidenced by the students in Bursuck and colleagues' (1999) study who indicated a tolerance for differentiation in assessment response modes (e.g. oral versus written) when they are necessary for students with special learning needs. However, they perceived changes to the degree of an assessments' difficulty, either in the task design or application of criteria, as unfair. Bursuck and colleagues explained that the students often felt that grading adaptations would "undermine the motivation of higher achievers" who were attempting more difficult courses instead of taking the "easy way out" (p.91). A similar idea is expressed by one of the teachers in Zoeckler's (2005) study in a comment about grading the work of lower achieving students in her class.

It's a challenge, because you want to give the same fairness; they're in the same classroom, they need to take the same test as everyone else. You can't just give them points for doing nothing because that detracts from the top student who is really achieving (p.155)

The comments by this teacher and the students in Bursuck and colleagues' study highlight a

difference noted earlier in the studies on the ethics of teaching. While Jackson and colleagues (1993) describe scenarios that show how giving all students equally positive feedback shortchanges learning and devalues achievement, Campbell illustrates how lowering expectations can have the same effect. As such, it appears that neither equal treatment, nor differentiated treatment are universally fair or unfair in classroom assessment.

The tension between consistency and differentiation is further evidenced in some of the differences in teachers' beliefs, views or positions in these studies (Green et al., 2007; Morgan & Watson, 2002; Phillips, 2002; Szyrka, 2001; Yung, 2001b; Zoeckler, 2001). Szyrka (2001) determined that the teachers whose classrooms were more equitable (based on Rose et al., 1991) were more concerned with students' learning than grading, and they focused more on students' achievement than on effort. In the more equitable classrooms, students were offered a variety of opportunities to learn, and the assessment criteria and processes were made explicit. The teachers used assessment to inform instruction and give students feedback, and they considered multiple aspects of learning to determine grades. Despite these commonalities, Szyrka also noted significant differences in these teachers' beliefs about fairness. For one teacher, assessment was fair if it was "modified in some way so that students can be successful" (p.82), whereas another teacher thought fair assessment meant that "every one is measured on the same standard" (p.87). These different beliefs exemplify what Lam (1995) referred to as the "two antithetical views of fairness: equality and equity" (p.2). Some teachers aim for equality in the classroom, and they seem to resent the need to "bend the rules" (Szyrka, 2001, p.93) for exceptional students. Others are confident that they have the "legitimate authority" (Zoeckler, 2005, p.89) to make

adjustments that are in the best interest of students, particularly well-behaved students with borderline grades. Throughout all of the empirical research, a confusion between equality and equity was seen in how teachers and students understood fairness. Whereas equality is associated with sameness, equity refers to appropriateness, and as Messick (2000) suggests, “not all inequalities are inequities” (p.12). An example of inequality being equitable is the differentiation of assessment for specific groups, such as atypical or dominant language learners. Concerns about equality are overwhelmingly present and frequently manifested in references to the consistency of assessment. This tension is a fundamental area of debate in the ethics of teaching (Strike & Soltis,1998), so it is not surprising that many teachers struggle morally with the tension between treating all students equally and ensuring the equitable treatment needed by some students for learning in the classroom environment.

Teachers’ beliefs about fairness may differ diametrically even when they work in the same educational context. The three teachers that Yung (2001b) described all taught biology in secondary schools in the same educational system in Hong Kong, and the 13 teachers in Zoeckler’s (2005) dissertation all taught English in the same school department in New York. In both of these studies, the teachers expressed considerable concern about assessing fairly, but their interpretations of fairness were radically different. The variation in teachers’ thinking is particularly evident in Green and colleagues’ (2007) study where they found high levels of disagreement among teachers about the ethics of routine classroom assessment activities (e.g., grading essays, writing report card comments). Moreover, some of the grading practices that show higher degrees of agreement (e.g., conflation of effort and achievement, use of late penalties, etc.) conflict with the recommendations of assessment

specialists. As a result, Green and colleagues conclude that the ethics of classroom assessment is a “realm without professional consensus” (p.1009) that would benefit greatly from further research and dialogue.

Inconsistency between teachers leads some of these researchers to express concern about inequity across an educational system (Morgan & Watson, 2002; Yung, 2001b). This is the same concern voiced by the teachers in Yip & Cheung’s (2005) survey, and also the underlying issue that drove Watson’s (1999) inquiry. Defining inequity as “the making of different decisions and offering of different opportunities, in similar circumstances, by different teachers, in ways which might affect pupils’ futures” (p.109), Watson concentrated on teachers’ assessment practices in one small national system (UK). She uncovered six paradigmatic conflicts about teaching, learning and assessment in mathematics, and she argued convincingly that these are sources of social inequity. Overall, in the literature on fairness in classroom assessment, consistency is viewed as a desirable aspect of fairness, both within individual teachers’ classroom practices and across an educational system.

What is not emphasized in the empirical work are aspects of fairness that fall outside the interactions between teachers, students and students’ work. This means that there is little discussion about some of the aspects of fairness that might interest school leaders, which were included in the principles and standards. Several examples of this void are that fairness issues relating to the storage of assessment information (i.e., right to access, protection of privacy) are not raised, and that the relationship between fairness and collaborative assessment within a school is not explored. As a whole, empirical research that sheds light on fairness in classroom assessment is concerned primarily with the issues or dilemmas that

teachers face as individuals as they engage in the process of assessing student learning.

While the teachers in these studies do aim to support student learning, the focus tends to be on more assessment of learning rather than assessment for learning, and the literature does not yet provide a clear picture of fairness as a quality of the ongoing interactions involved in assessment for learning. The teachers do repeatedly raise concerns about fairness, and their voices provide evidence about the moral complexity of fair classroom assessment, but they are not necessarily using reflection as a conscious strategy to improve the fairness of their practices. As Watson (1999) notes, even in a “highly regulated system” (p.114), teachers have considerable autonomy in their classroom assessment practices, and their interpretations of fairness influence how they approach and resolve related issues and dilemmas. As such, gaining insight into teachers’ interpretations of fairness in classroom assessment has much potential in both practical and theoretical terms.

#### Fairness in Writing Assessment in Ontario

In Ontario’s publicly funded educational system, student writing is assessed both in the classroom and as part of a large-scale assessment program. Although my study is concerned with fairness specifically in the classroom environment, I provide a brief overview of research on fairness in relation to one large-scale assessment, the OSSLT, because it sheds light on how the concept of fairness has been interpreted in research literature situated in the same educational context as my study.

The OSSLT is the only large-scale writing assessment administered to students after the elementary years in Ontario. Students are required to pass the OSSLT to graduate from secondary school, and the percentage of students who pass it in one attempt is very high

(84% in 2008). The OSSLT is developed and scored annually by the Education Quality and Accountability Office (EQAO), which claims that it is “fair for all students” (2008b, p.37). However, the pass rates for specific groups, including English-language learners (38%), students with special needs (45%), students in applied English classes (56%), and students who previously failed the test (31%) are all much lower. As these statistics suggest that the OSSLT is significantly more challenging for certain groups of students, considerable research interest has been generated in the fairness of this test for these groups. One team of researchers in Ontario collaborated on a multi-phase project that concentrated on the perceptions, experiences, and performance differences of English as a Second Language (ESL) and English Language Development (ELD) students taking the OSSLT (Cheng, Fox & Zheng, 2007; Cheng, Klinger & Zheng, 2007; Fox & Cheng, 2007; Luce-Kapler & Klinger, 2005). The results of this research project raise serious validity and fairness concerns about the OSSLT for students who are learning English in English-language school boards in Ontario’s educational system. In addition to the this teams’ work, three doctoral research projects highlight fairness issues with the OSSLT for specific populations. Brackenreed (2006) questioned the fairness and effectiveness of test accommodations for special needs students, and Kearns (2008) explore the impact of the OSSLT on students who had failed the test at least once. Spencer (2006) engaged in an ethnographic study at one school, and she argues that the OSSLT creates a “problem population” of students. In sum, a body of empirical work has been developed by educational researchers in Ontario that expresses strong concern about the fairness of the OSSLT. In most of this work, fairness is considered in terms of group differences, which is in keeping with the interpretation of fairness that is

usually seen the literature on standardized testing.

Despite the increase in the amount of classroom assessment research in recent years, work relating specifically to writing assessment in Ontario's elementary and secondary classrooms is very limited. A few studies have been done in Ontario on the use of rubrics in writing assessment. Engemann and Gallagher (2006) discuss rubrics as tools for valid, reliable and fair writing assessment, and they briefly describe a study where the consistency (reliability) of scoring improved when teachers discussed their judgments about writing samples. As their coupling of reliability and fairness is cursory, Engemann and Gallagher seem most interested in emphasizing the need for collaborative professional development to improve the quality of teachers' assessments. Ross, Rolheiser and Hogaboam-Gray (1999) and Laveault and Miles (2002) also investigated rubric-based writing assessment in Ontario, but both of these teams focused on the accuracy of student self-evaluation. Although these authors did not explicitly raise the question of fairness, rubrics are used to share assessment criteria with students, which is one of the aspects of fairness emphasized in classroom assessment theory. Ross and colleagues used a quasi-experimental design with 15 teachers in the treatment group, whose students in Grades 4 to 6 were taught self-evaluation in writing, and 15 teachers in a control group who did not attend in-service sessions or implement the self-evaluation strategies. This research team arrived at two conclusions. First, that students in the treatment group were more accurate in their self-evaluations, and second, that the self-evaluation process helped weaker students improve their writing. Laveault and Miles were interested in whether students' characteristics were associated with their accuracy in using rubrics. They had students in Grades 5 to 8 ( $n = 770$ ) score writing exemplars with a rubric,

and they found differences in accuracy related to students' writing ability, grade, gender, and their membership in an ability group (i.e., learning disabled, gifted, talented, control). These two studies indicate that rubric-based self-assessment is less accurate for certain students, specifically those who are not explicitly taught to self-assess, those who are weaker writers, those who identified as learning disabled, and those who are male. As the accuracy of an assessment affects its effectiveness in supporting learning, these results also raise questions about equity in terms of opportunity to learn, and hence about the fairness of rubric-based self-assessment for students in these groups.

The issue of gender equity in writing assessment was the concern of another team of researchers in Ontario. Peterson, Childs and Kennedy (2004) sent a set of student papers to Grade 6 teachers (n = 108) in 17 schools in Ontario for grading and comments. Each set contained four papers (one persuasive and one narrative paper identified as a girl's and one persuasive and one narrative paper identified as a boy's) that were combined differently to make four study "conditions" (p.164). Peterson and colleagues (2004) did not find any significant differences, based either on teacher gender or student gender, in how the papers were scored. They conclude that "the efforts of teachers in this study to mark and respond to students' writing in an unbiased way appear to be successful, for the most part" (p.176). However, they did note some patterns that led to a further analysis of the comments. Peterson and Kennedy (2006) subsequently found that while there were no differences in the number of comments, based again on the gender of either the teacher or student, there were some "interesting gender trends" (p.54). Specifically, comments for boys tended to contain more commands, and praise for both boys and girls showed evidence of same-sex depreciation.

Also, male teachers commented more on artistic style, while female teachers commented on conventions and organization. As Peterson and Kennedy noted, these results should be of “interest to educators and researchers who strive for gender equity in writing instruction and assessment and for enhancing all students’ development as writers” (p.58). Although the body of research on writing assessment in Ontario’s educational system is relatively small, there are findings that suggest a need for more conscious attention to fairness. With such a small collection, the results of these studies also beg the question about what transpires beyond the realm of research, in the ongoing interactions between teachers and students who are developing as English writers in Ontario.

#### Chapter Summary

In this chapter I looked at how fairness is understood in different types of literature. In the documents containing principles and standards for classroom assessment, fairness is associated with all phases of assessment, and as a result, its interpretation far exceeds the breadth of its psychometric definition. References to standardized procedures are replaced with a strong emphasis on the need for transparent criteria and procedures, written policy, and clear communication of assessment results (see Appendix B). Many of the aspects of fairness found in these committee documents are reflected in classroom assessment theory. As a body of work, however, classroom assessment theory presents a narrower view of fairness, with the emphasis primarily on transparency, opportunity to learn and demonstrate learning, and the recognition that social values and cultural expectations influence classroom assessment. While the fairness of standardized test results may be determined through statistical procedures, such differential item analysis, classroom assessment theorists ask

teachers to reflect critically about their personal biases, roles, and power dynamics in assessing student learning. Classroom-based empirical research emphasizes an even different view of fairness. The teachers whose voices are heard through empirical research give evidence to the complexity of fair classroom assessment. They engage in reflection about their assessment practices, but not necessarily with the intention of improving fairness, as theorists recommend. In the ethical inquiries, fairness is seen as an essential moral quality for a teacher, as well as a quality of classroom interactions, which the teacher influences as a moral agent. In contrast, fairness is discussed in classroom assessment research as a quality of assessment tasks or process, and the opportunity to learn that is provided. The tension between differentiating for individual students and maintaining the same criteria or tasks for all students is most prominent. A confusion between equality and equity seems to underlie how fairness is interpreted throughout the literature. A small body of work relating to fair writing assessment has been produced in the Ontario educational context. To date, this work has tended to concentrate on group differences, whether it be differences in how certain groups of students (i.e., English language learners, students with special learning needs) experience or perform on the OSSLT, differences in the effect of rubric-based classroom assessment for specific groups (i.e., based on student abilities), or gender-based differences in teachers' comments. Thus far, fairness has not been considered as the subject of empirical research in Ontario in terms of the assessment interactions that take place regularly between teachers and students. In the next chapter, I discuss my approach to understanding fairness in the ongoing interactions of classroom assessment through the practical wisdom of teachers.

## CHAPTER FOUR: TURNING TO TEACHERS' PRACTICAL WISDOMS

*Consider theorizing in this light: imagine developing and exchanging theories the way you create a recipe and share it with a friend. What would such theorizing look like?*

Lisa Heldke (1988, p.23)

In the opening chapter of this dissertation, I gave a brief account of critical pragmatism, and I identified two implications of this philosophical approach for my work. Because I view knowledge as being located in time and place, I discussed the classroom environment and the key concepts of assessment for learning and fairness in both general and specific terms. Because I view knowledge as changeable, I showed in the previous chapters how these concepts have evolved within the educational community. Here, I elaborate further on my understanding of critical pragmatism in order to discuss a third implication for my work, which relates to the value I place on teachers' practical wisdom as a source for understanding fairness in classroom assessment. Following this, I elaborate on the concept of practical wisdom, and I consider factors that can influence teachers' practical wisdoms in Ontario's educational context. In the final section of this chapter, I summarize the key concepts in this study in relation to a visual diagram, and I situate each of the three research questions accordingly.

### Critical Pragmatism

Pragmatism began in the late 19<sup>th</sup> century as an American philosophical movement based on the ideas Charles Peirce, William James, and John Dewey, and it resurfaced during the later years of the 20<sup>th</sup> century in the works of Richard Rorty, Hilary Putman, Richard Bernstein, Cornell West, and Cleo Cherryholmes (Maxcy, 2003; Ozmon & Craver, 2008; Warms & Schroeder, 1999/2009). Margolis (2006), a contemporary American philosopher,

argues that a third wave of pragmatist thought could have considerable meaning for philosophical discussion in the 21<sup>st</sup> century. My interest in critical pragmatism is at once more humble and more self-interested; it helps me understand and articulate my beliefs about the nature and generation of knowledge in the context of educational research and practice.

With critical pragmatism, I adopt a stance toward educational research and practice that is both critical and hopeful. In questioning accepted practices and assumptions in education, critical pragmatism aims to move beyond the nihilistic and elitist orientation of some postmodernist theory (McLaren, 2003b; Ozmon & Craver, 2008). Although it avoids defining itself in the negative terms often used in critical pedagogy (e.g., anti-x in McLaren, 2003b), critical pragmatism does share an important idea about the nature of knowledge with both critical theory and sociocultural theory. Knowledge is thought to be socially constructed in all three of these perspectives, but the idea is discussed somewhat differently in each.

Sociocultural theory focuses on explaining how learners construct knowledge both individually and socially (Shepard, 2005), whereas critical theory emphasizes the power of the dominant culture in conceptualizing knowledge as a social construction (McLaren, 2003a). Critical pragmatism is also concerned with power dynamics, but not necessarily those of the dominant culture. Most importantly, it is regarded as less of a theory, and more of a way of theorizing (DeWaal, 2005). Warms and Schroeder (1999/2009) describe pragmatism as a “way of doing philosophy that weaves together theory and action, each continuously modifying the other and maintaining their mutual relevance” (p.271).

Cherryholmes (1988) explains that critical pragmatism is a “continual movement” (p.97) that not only involves the construction and deconstruction of educational practice, but also the

construction and deconstruction of the criticism of practice within an educational community. In this way, knowledge is socially constructed because it is endlessly adjusted through the interaction of all our theories and practices, and our critical thinking about those interactions.

Heldke (1988), a philosopher whose writing is located at the intersection of feminism and pragmatism (Sullivan, 2006), expresses a similar view of theorizing with her recipe analogy (in the epigraph to this chapter). She compares philosophical inquiry and cooking, and she argues that theories are like recipes because they are “tools we may choose to use, outlooks we may elect to assume. Some are more useful than others; none are universally reliable” (p.17). Theory is not venerated from this perspective, but it is valued in terms of its usefulness for inquiry. Critical pragmatism assumes that different forms of knowledge are valuable for inquiry (Bernstein, 1985; Biesta & Burbules, 2003; Maxcy, 2003). The implication for educational research is that the forms upon which it has conventionally relied, namely theoretical and empirical knowledge (Fenstermacher, 1994; Kessels & Korthagen, 1996), are neither discarded, nor assumed to be superior to other forms of knowledge, particularly ethical, practical and narrative knowledge. From this perspective, teachers’ practical wisdoms are valued as a source for deepening our understanding of both theories and practices.

### Teachers’ Thinking

This study draws on the practical wisdoms of experienced teachers to understand fairness in classroom assessment. The term *practical wisdom* is used here as the translation of the Ancient Greek term *phronesis*. The term *phronimos* refers to a person who has developed practical wisdom. Different translations of these terms are found in the literature

(Connolly, 2005; Kern, 2007). Phronesis is sometimes translated as practical reasoning (e.g., Fenstermacher, 1994), practical-moral judgment (e.g., Connolly, 2005), or practical knowledge (e.g., Waluchow, 2003). Many authors preface their writing about teachers' thinking with comments about the "terminological diversity" in this area (Ariogul, 2006, p.15; also Meijer et al., 1999). For the sake of brevity in describing groups of work, I use the short umbrella term *teachers' thinking* to include teachers' knowledge, beliefs, reasoning, understanding, and wisdom. I use more specific terms when needed to indicate different forms of knowledge, or distinct conceptualizations of teachers' thinking. Borg (2003) argues in his review of research on teachers' thinking that the "superficial diversity created by the terms . . . should not mask the considerable overlap which exists among them" (p.83). As such, I begin with a brief review of three main interests in teachers' thinking in educational research, followed by a summary of the major conceptualizations of teachers' thinking before concentrating further on the concept of practical wisdom. My reason for doing this is twofold. First, I wish to clarify how my interest in teachers' practical wisdom compares to other studies that involve teachers' thinking, and second, I want to show how different conceptualizations have emerged and relate to the concept of practical wisdom.

### *Interests in Teachers' Practical Thinking*

The thinking that teachers use in practice has been of considerable interest to researchers in recent years. A perusal of educational databases indicates that researchers' interests basically fall into three main groups in this growing body of work. First, there are a multitude of studies interested in how teachers develop practical thinking, and often they focus on the effect of a certain educational approach or professional development program.

An example is Choi's (2007) study on the influence of an inquiry-based science course on the beliefs, practical knowledge, and practices of elementary science teachers. Second, many recent studies are interested in the nature or content of teachers' practical thinking in specific contexts. An example is Meijer's (1999) examination of the content of practical knowledge used to teach reading comprehension in the Netherlands. The third research interest is in using teachers' practical thinking as a source for understanding an issue in practice. An example is Yen's (2006) work with three exemplary teachers to gain insight about improving education for Indigenous minority students in Taiwan. In summary, recent empirical research involving teachers' practical thinking clusters into three main groups, with one group interested in how it is affected, a second group interested its content and nature, and a third group interested in using it as source of information. My interest in teachers' practical wisdom aligns most closely with the third group. However, none of these studies used any form of teachers' thinking for the same purpose as my study, which is to understand fairness in classroom assessment.

### *Conceptualizations of Teachers' Thinking*

In addition to differences in researchers' interests in teachers' practical thinking, there are also significant differences in how it is conceptualized. Several researchers whose studies are about teachers' thinking have provided comprehensive historical reviews of the literature in this area (see Ariogul, 2006; Borg, 2003; Meijer, 1999). Given that my study is about fairness, my purpose in discussing some of these studies here is not to review research on teachers' practical thinking, but to explain why I chose the concept of practical wisdom for my study. Most of the recent empirical work in this area builds on three schools, which

essentially involve conceptualizing knowledge as an inventory, as narrative, or as a network.

*Inventory of Knowledge.* This conceptualization examines teachers' knowledge in order to describe its nature and content. It was developed in the early 1980's by researchers such as Elbaz (1980) and Shulman (1987) who wanted to understand the knowledge that teachers use in practice at a time when there was little pre-existing research on the subject. Elbaz conducted an in-depth, single case study that explored the practical knowledge of a secondary English teacher in Canada. She described her participants' thinking as consisting of five categories: knowledge of curriculum, subject matter knowledge, instructional knowledge, knowledge of self, and knowledge of the milieu. Based on the assumption that practical knowledge is "held in active relation to the world of practice" (p.101), Elbaz also looked at five aspects of its orientation: theoretical, situational, social, personal and experiential, and at three levels through which practical knowledge is structured: rules of practice, practical principles and images. Elbaz rejected the view, which was common at the time, of the teacher as a "passive transmitter" (p.294) of curricula, and she emphasized the complexity of teachers' practical knowledge.

Shulman's (1987) inventory of teachers' knowledge was similar in that he also viewed teachers' abilities in a positive light, and he drew on some of the same theoretical sources as Elbaz (e.g., Dewey, Schwab). However, the underlying goals of these two works were quite different (Fenstermacher, 1994). Elbaz' (1980) work was completed as a doctoral dissertation, and it is descriptive in nature. When Shulman's (1987) inventory was published in the *Harvard Educational Review*, he was already a well-established researcher, a past president of the American Educational Research Association, and involved in several major

projects relating to teachers' knowledge. As his research was undertaken at time when the quality of teacher education was under political and public scrutiny in the United States (Bullough, 2008), it was more "normatively oriented" (Fenstermacher, 1994). Shulman argued for the professionalization of teaching, and he identified four sources for developing teacher knowledge: content scholarship, educational materials and structures, educational scholarship, and the wisdom of practice. Shulman's inventory, or knowledge base for teaching, contained seven categories: content knowledge, general pedagogical knowledge, curriculum knowledge, pedagogical content knowledge, knowledge of learners, knowledge of educational contexts, and knowledge of educational purposes and values. Although this inventory generated criticism at the time, Shulman's work continues to inform current research. The most notable example of his influence is seen in the educational community's adoption of the term *pedagogical content knowledge*. Shulman defined it as the "blending of content and pedagogy into an understanding of how particular topics, problems, or issues are organized, represented, and adapted to the diverse interests and abilities of learners, and presented for instruction" (p.228). He noted that pedagogical content knowledge is of "special interest" (p.228) among the other forms of knowledge because it is distinct to teaching, which supports the view of teachers as professionals. An indication of the widespread acceptance of the concept of pedagogical content knowledge is seen in its addition to the ERIC database thesaurus in 1998. A search using pedagogical content knowledge as a descriptor (1986 to 2008) now results in over a thousand records.

*Knowledge as Narrative.* Elbaz' work (1980) also continues to influence current research, and it is directly connected to a second way of thinking about teachers' knowledge,

which is as narrative. Elbaz noted the influence of her thesis supervisor, Michael Connelly, in her discussion about teacher agency. From the early 1980's on, Connelly collaborated with Jean Clandinin on a program of research relating to teachers' knowledge (for a review see Wright, 2002). In various publications spanning over a decade, they described a method for studying teachers' knowledge, which is identified as a form of narrative inquiry (e.g., Connelly & Clandinin, 1986; Connelly, Clandinin & He, 1997). As Connelly and Clandinin (1986) explain, they coined the term *personal practical knowledge* to define their "interest in understanding teaching acts in terms of personalized accounts of people knowing" (p.297). In their later writing, they stressed the personal aspect of teachers' knowledge because they believe that "knowledge is not something objective and independent of the teacher to be learned and transmitted but, rather, is the sum total of the teachers' experiences" (Connelly et al., 1997, p.666). They also emphasized that teachers develop personal practical knowledge in a professional environment, called the *professional knowledge landscape* (p.673). Connelly and Clandinin's conceptualization of teachers' knowledge has been taken up by many researchers, especially when the goal is to understand and share teachers' experiences in a particular educational context. Evidence of this is seen in the increasing number of dissertations (i.e., ProQuest Dissertations database from 1983 to 2009) involving teachers' personal practical knowledge and narrative inquiry. In their recent work, Connelly and Clandinin have continued to clarify their ideas about the role of narrative in educational inquiry.

People shape their daily lives by stories of who they and others are and as they interpret their past in terms of these stories. Story, in the current idiom, is a portal through which a person enters the world and by which their experience of the world is interpreted and made personally meaningful . . . To use

narrative inquiry methodology is to adopt a particular view of experience as a phenomenon under study. (2006, p.375)

The distinction that Connelly and Clandinin make between narrative as a way of knowing and narrative as inquiry is important for my work. Although I use narrative in several ways in my study, its design differs from narrative inquiry. I discuss this further in the next chapter.

*Knowledge as a Network.* In a review of research relating to teachers' knowledge, Fenstermacher (1994) suggested that although the inventory and narrative conceptualizations are both based on a "high degree of regard for teachers" (p.32 ), and they present practical knowledge as complex and contextual, they are not entirely satisfactory for describing the practical wisdom that is operative in classrooms. Along with several contemporary educational philosophers (Bernstein, 1985, Biesta, 2007, Dunne, 1993; Orton, 1997), Fenstermacher argues that the Aristotlean concept of phronesis, which he translated as practical reasoning, is more complete for understanding teachers' knowledge because it includes a moral dimension.

Practical reasoning may also address the moral aspects of action, indicating that it was fair, right, or the best of a number of poor alternatives. As such, it permits us to consider the moral dimensions of teaching, dimensions that others . . . argue are at the very heart of the enterprise of teaching. That it permits consideration of both epistemological and moral dimensions of teaching is a compelling reason for entertaining practical reasoning as an approach to the study of the practical knowledge of teachers (1994, p.45).

This is not to say that the inventory and narrative conceptualizations completely ignore the moral dimension of teacher knowledge. Clandinin and Connelly (1996) noted that the professional knowledge landscape is both intellectual and moral because it is "composed of relationships among people, places, and things" (p.25), and both Shulman and Elbaz addressed the moral dimension of teaching in the decades following their initial inventories.

Shulman (1992/2004) responded to criticism of his work by pointing out that pedagogy aiming to be effective and responsible is “inherently moral” (p.378), and Elbaz (1992) later studied the characteristics of moral knowledge in teachers’ stories. Although she concluded that it was important for researchers is to enable teachers’ “moral voices” (p.431), Elbaz did not consider how these voices interact with their other forms of teachers’ knowledge. Thus, these authors do consider the moral dimension of teachers’ knowledge, but they do not emphasize it to the same degree as the other forms of knowledge they discuss (i.e., personal, experiential, theoretical, technical, and contextual). It is only when the Aristotlean concept of practical wisdom is used to understand teachers’ thinking that moral beliefs and ethical knowledge are highlighted alongside the other forms of knowledge. Connolly (2005) explains that practical wisdom draws on a mental network that includes multiple elements, namely theoretical knowledge, technical knowledge, understanding of particulars, experience, personal characteristics, and moral beliefs and ethical knowledge. I prefer this conceptualization of teachers’ thinking because it acknowledges the ongoing interaction of all these different elements.

In choosing to conceptualize teachers’ thinking as a network, I am not excluding what has been learned from thinking about teachers’ knowledge as an inventory of types, or as narrative. Although the discussion above dealt only with the key figures associated with the inventory and narrative conceptualizations, they represent large bodies of research that are respected within the educational community. They are also helpful for understanding Aristotle’s concept of practical wisdom in contemporary terms within an educational context, especially in terms of their commonalities. For example, Shulman (1987) and Elbaz (1980)

both included content knowledge in their inventories, which falls less distinctly in the category of theoretical knowledge in practical wisdom. All of these authors also situate knowledge in some way. Shulman (1987) identifies knowledge of educational contexts, Elbaz (1980) discusses knowledge of milieu, Connelly and colleagues (1997) refer to the professional landscape, and an understanding of the particulars of practice is one of the elements of practical wisdom.

### *The Concept of Practical Wisdom*

Practical wisdom is one of the five types of knowledge discussed by Aristotle in the *Nichomachean Ethics* (trans.1925/1975). Three of these types of knowledge, being philosophic wisdom (*sophia*), reason (*nous*), and knowledge (*episteme*), are in the theoretical realm. Technical knowledge (*techne*) and practical wisdom are both forms of practical knowledge, but they differ significantly (Aristotle; Dunne, 1993; Warne, 2006). Technical knowledge is associated with production (*poiesis*); it is a means to an end, involving general guidelines and the “capacity to make” (Aristotle, p.141; Dunne). In contrast, practical wisdom is associated with the particulars of ethical practice (*praxis*); it is an end in itself, involving moral decisions and the “capacity to act” (Aristotle, p.142; Dunne). Of the five types of knowledge, practical wisdom is the most complex, flexible, and responsive. Although it is identified primarily as an Aristotelean concept, practical wisdom has been discussed from classical times by a multitude of philosophers, from Plato to the present (Dunne). Contemporary philosophers have suggested that practical wisdom is an ideal source for delving into the personal, moral, and often tacit aspects of practice, which might not be visible with technical or theoretical approaches (Bernstein,1985; Biesta, 2007; Dunne;

Fenstermacher, 1994). I explain each of the elements of practical wisdom in turn below:

*Theoretical Knowledge.* Theoretical knowledge (episteme) is distinct from practical wisdom because it involves abstract universals rather than contextual particulars, and it is justified by conceptual rather than perceptual reason (Fenstermacher, 1994; Kessels & Korthagen, 1996). However, a “resourcefulness of mind” (Dunne, 1993, p.272) in applying theoretical knowledge in difficult circumstances is a key characteristic of practical wisdom. In classroom practice, a teacher’s theoretical knowledge would involve content knowledge (e.g., English Literature) and theories about teaching, learning and assessing (e.g., Vygotsky’s zone of proximal development). In discussing phronesis as a model for reflective teaching and teacher education, Birmingham (2003) explains that “reflection begins with the experience, seeks out information that can be found in educational theory as it is needed and relevant, and uses this information to help answer questions about the experience” (p.192). The example she provides to illustrate her point describes a situation where pre-service teachers misinterpreted events in practice because of their limited theoretical knowledge. As Connolly (2005) observes, strong theoretical knowledge provides teachers (professors) with a broader range of “action options” and “more complex ways of seeing” (p.162). This highlights the notion that practical wisdom is more than a linear implementation of theory into practice, and that there is movement between theoretical knowledge and understanding an event, which takes its root in practice.

*Technical Knowledge.* Technical knowledge (techne) is closely related to theoretical knowledge, despite its co-location in the practical realm, because it involves general rather than particular knowledge (Dunne, 1993). Technical knowledge is also like theoretical

knowledge in that it informs practical wisdom, but in a non-reciprocal relationship such that practical wisdom “carries the others” (Connolly, 2005, p.67). Technical knowledge in teaching is the pedagogical knowledge that involves knowing how to teach and assess effectively (e.g., steps for a clear and engaging assessment activity). While technical knowledge is critical for classroom teaching, it is in itself not sufficient. Field and Latta (2001) explain that this is because teachers are “simultaneously confronted with the demand, not simply to know how, but to know when, and how much and with whom” (p.891). Simultaneous demands are repeated over and over in the ongoing interactions of classroom education, and choosing the best course of action in any particular situation requires technical knowledge to work in tandem with the other elements of practical wisdom.

*Understanding of Particulars.* Aristotle emphasizes that practical wisdom results from experiencing the particulars of circumstance over time (trans. 1925/1975; Dunne, 1993; Kessels & Korthagen, 1996). It operates as a “type of mediation” (Bernstein, 1986, p.110) between universals and particulars. In practice, teachers might draw on their understanding of a particular context (e.g., school community), a particular situation (e.g., illness of a parent), or a particular student (e.g., learning habits) to respond to a student’s work and make assessment decisions. In their discussion about experience in the context of teacher education, Field and Latta (2001) argue that because practical wisdom is informed by the particulars of circumstance and place, it should be understood, not as a codifiable knowledge, but as a dynamic experience. Dunne also emphasizes this point, explaining that practical wisdom “is realized always in concrete applications and never resolves itself into formulated knowledge that can be possessed apart from these applications” (p.127). Thus, practical

wisdom requires knowledge, but cannot be held with any permanence as knowledge.

*Experience.* Experience plays a critical role in the development of practical wisdom. Aristotle emphasizes the importance of experience throughout the *Nicomachean Ethics* (trans.1925/1975). For example, in Book Three, he states that professional soldiers are the “most capable in attack and in defence” (p.68) because of their practical experience, and in Book Ten he reasons that “those who aim at knowing about the art of politics need experience” (p.274). In Book Six, he associates practical wisdom with a person’s “time in life” (p.153), and he argues that a young person cannot be a phronimos because “such wisdom is concerned not only with universals, but particulars, which become familiar from experience, but a young man has no experience, for it is length of time that gives experience” (p.148). More currently and in relation to education, Berliner (2004) notes that teachers spend “hundreds and thousands of hours” (p.201) gaining practical experience, and he argues that “wisdom is derived from teaching experiences” (p.206). Kessels and Korthagen (1996) believe that experience is a “prerequisite” for teachers’ practical wisdom because the particulars of practice “only become familiar with experience, with a long process of perceiving, assessing situations, judging, choosing courses of action, and being confronted with their consequences” (p.20). The importance of classroom experience for teachers is acknowledged in the requirement of a practice session (practicum) in most teacher education programs.

However, it cannot be assumed that a positive linear relationship always exists between practical wisdom and teaching experience. In Book Ten, Aristotle notes that although doctors are not “made by a study of text-books” (1925/1975, p. 275) alone, and that

they require a “practiced faculty” (p.275) for their decision-making, they must also study “the universal, and come to know it as well as possible” (p.273). Drawing on one of the examples that Aristotle uses to distinguish between theoretical knowledge, experience, and practical wisdom, Dunne (1993) explains, in essence, that experience is necessary, but not sufficient for practical wisdom. In educational literature, the caveat that experienced teachers are not necessarily wise or expert is often stated (e.g., Berliner, 2004; Connolly, 2005; Shulman, 1987/2004), and experienced teachers have been found to vary significantly in their knowledge and practices (e.g., Morgan & Watson, 2002; Yung, 2001b). Nevertheless, it is generally accepted that the wisdom of practice is gained gradually through experience, and thus, that experienced teachers are more likely to have developed practical wisdoms than are novice teachers.

The practical wisdoms of experienced teachers may be particularly helpful for understanding fairness in classroom assessment for two reasons. First, although classroom assessment is an area of practice that can be challenging for both experienced and novice teachers (e.g., Borko et al., 2000; Butterfield et al., 1999; Graham, 2005), experience advantages teachers in dealing with the complexities of assessment. For example, of the teachers who responded to Bol, Stephenson, O’Connell and Nunnery’s (1998) survey on classroom assessment practices (n = 893), those with the most experience (20 + years) reported the highest levels of confidence in using varied assessment methods. This is supported by Cowie and Bell’s (1999) observations over a two-year period where the novice teachers in their study found formative assessment difficult because it involves recognizing and responding to student learning spontaneously in the classroom (p.109, 111-112). A

second reason that experienced teachers' practical wisdoms may be a particularly good source of information is that experienced teachers' concerns about classroom assessment tend to go beyond the technical or managerial concerns of novice teachers (Cheung, 2002), and they may also have more confidence in critiquing the "predominant professional ideology" of their educational context (Zembylas, 2003, p.313).

*Personal Characteristics.* As an experienced teacher is not necessarily a wise teacher, it seems that personal characteristics must also be at play in the development of practical wisdom. Dunne states that practical wisdom "expresses the kind of person that one is" (p.244). A variety of personal characteristics might influence a teacher's practical wisdom about classroom assessment, such as gender, learning preferences, or family background. This is particularly evident in a case study on the change process in a professional development program, where Borko and colleagues (2000) attributed the variation in how two elementary teachers changed their assessment practices to the differences in their personal beliefs and life circumstances. Another example of the influence of personal characteristics is seen in Connolly's (2005) research with three university professors. He observed that one professor maintained a particularly technical stance to teaching, and he concluded that with this approach, efficiency rather than wisdom results from experience.

*Moral Beliefs & Ethical Knowledge.* Moral beliefs and ethical knowledge play similar roles, and they are central to fair decision-making in practice (Bernstein, 1985; Fenstermacher, 1994). Moral beliefs are those "which are evaluative in nature...which distinguish, whether consciously or unconsciously, between what is right and wrong, good

and bad” (Buzzelli & Johnston, 2002, p.3). Ethical knowledge is similarly defined in that ethics “concerns what kind of actions are right or wrong, what kind of life is a good life, or what kind of person is a good person” (Strike & Soltis, 1998, p.4). Campbell (2003) remarks on the conceptual similarity of ethics and morals in that they “both address virtue and basic principles of right and wrong as they influence belief, intention, and behaviour” (p.17). In this way, a moral decision may also be an ethical decision, but they are not necessarily one and the same. The difference lies in their expression. Ethical knowledge is a “more collective and public” (Campbell, 2003, p.17) expression of morals, which are often stated as “codes of practice” (Buzzelli & Johnston, p.5). Ethical codes are usually created within professions to provide practitioners with the principles of practice, which are in essence the moral beliefs that are accepted and shared within a community at a particular time. For instance, *The Ethical Standards for the Teaching Profession in Ontario* (OCT, 2008a) “describe the professional beliefs and values that guide the decision-making and professional actions of College members in their professional roles and relationships” (p.5). Beliefs and values are also closely linked, but Campbell (2003) and Strike and Soltis (1998) identify values more closely with preferences. Values may be “non-moral preferences that individuals hold in relative ways” (Campbell, p.17), such as matters of popular taste or personal liking. Although some value judgments fall in the moral dimension, those of a non-moral nature are, as Strike and Soltis explain, “quite different from ethical judgements” (p.6). To avoid confusion, I do not use the term *values* in this dissertation, and in using the terms *moral beliefs* and *ethical knowledge*, I adhere to the definitions given above.

Although I have described the elements of practical wisdom separately to facilitate

my discussion, it is important to note that they are actively intertwined in their use. Field and Latta (2001) explain that practical wisdom does not draw on any one element alone; it involves teachers in being “mindfully embodied” (p.892) while availing themselves to their knowledge, beliefs, understandings, characteristics and experience. A key idea in thinking about how practical wisdom works is that it is not a static collection of knowledge types, but a “fluid” and responsive process (Dunne, 1993, p.127). In approaching my research, I understand practical wisdom about fairness as an ever-changing configuration, where some elements come to the forefront and others recede, depending on the individual teacher responding and the particular circumstances involved. For my study, I believe it should also be emphasized that although practical wisdom is individual, it is “acquired and deployed in one’s actions with one’s fellows” (Dunne, p.244; also Kessels & Korthagen). A teacher’s practical wisdom does not operate in isolation because it is “influenced by, and influences social practices” (Connolly, 2005, p.67). For that reason, it is important to consider teachers’ practical wisdoms within a particular educational context.

### *Teachers’ Practical Wisdom in Ontario*

There are well over 100, 000 elementary and secondary teachers working in the publicly-funded educational system in Ontario (Government of Ontario, 2008). As practical wisdom is individual, it cannot be understood in collective terms for this entire group. However, the requirements for teaching in an educational system provide evidence regarding the knowledge, experience, and characteristics that are encouraged, and ultimately expected of its teachers. As the requirements for teaching change over time and place, it cannot be assumed that the same elements of teachers’ practical wisdom are expected or encouraged in

every context. O'Donoghue and Whitehead (2008) found, for example, that the emphasis on content knowledge, pedagogical technique, and ideological beliefs has varied considerably as teacher education programs have developed over the years in different countries. In the following section I discuss the current requirements for teaching in Ontario in relation to several elements of teachers' practical wisdom. In doing this, I aim to situate the practical wisdom of my participants relative to teachers in Ontario in general because it brings the individuality of their practical wisdoms into relief against the commonalities of the context.

*Teachers' Theoretical Knowledge in Ontario.* Teachers must be certified by the OCT to teach in Ontario's publicly-funded educational system. For certification, the OCT requires a Bachelors degree and the successful completion of a teacher education program. Teacher education programs in Ontario must include studies about learning and the development of students, teaching methods, curriculum, and educational legislation (OCT, 1996). Teachers in intermediate and senior divisions must also have post-secondary education that is relevant for their teaching subjects. Teachers may also specialize by taking additional courses, such as English as a Second Language or Reading, but these are not required for certification. This means that to be certified to teach English Language Arts in Grades 7 to 12 in Ontario, a teacher would need to provide evidence of having gained theoretical knowledge, specifically content knowledge (i.e., English Literature) and general knowledge about learning and learners, at a recognized post-secondary institution. However, almost a third (29.0%) of the English teachers in Ontario who completed the 2002 SAIP questionnaire indicated that how they teach Language Arts is restricted to some degree by "limits" in their content knowledge (CMEC, 2002a, p.30). In their responses to the 2005 OCT survey, some teachers (18%) also

felt that their assignments were not “well suited given their content knowledge” (p.24). This suggests that although the educational requirements for certification aim to ensure that teachers in Ontario have sufficient knowledge for their teaching assignments, a notable minority of practicing teachers indicate that they are lacking content knowledge for their particular assignment.

*Teachers' Technical Knowledge in Ontario.* Technical knowledge involves the skills or techniques used in practice. The OCT (2008a) lists five general standards for teachers' practices in Ontario. Classroom assessment appears in one standard, which states that professional practice involves using “appropriate pedagogy, assessment and evaluation, resources and technology in planning for and responding to the needs of individual students and learning communities” (p.13). Because the standards apply to all teachers, they do not specify assessment strategies or techniques for any given level or subject area. Additionally, there are no regulations relating to the inclusion of classroom assessment as a topic in teacher education programs in Ontario (OCT, 1996). Nonetheless, teachers in Ontario appear to be very confident in their assessment skills. Almost all (92%) of the teachers who responded to the 2004 OCT survey felt that teachers' assessments (e.g., observations) are an accurate measure of students' abilities, and many (82%) also rated teacher-made assessment tasks highly (4-5 on 5-point scale). This is particularly notable in comparison to the small percentage (11%) who rated provincial standardized tests as an equally accurate measure.

To understand teachers' technical knowledge in Ontario specifically for the purpose of this study, I looked at the items on the 2002 SAIP questionnaire that related to classroom assessment. Almost all of the respondents (99.1%) agreed that “assessment must be an

integral and ongoing part of the learning process, not limited to final products” (CMEC, 2002a, p.14). This indicates that English teachers in Ontario overwhelmingly support the idea of assessment for learning. As this questionnaire focused on measuring the frequency of specific practices, it provides evidence that many teachers were using some of the core practices associated with assessment for learning in 2002. For example, many were involving students in peer and self assessment activities, sharing assessment criteria through writing exemplars or rubrics, and giving students feedback on homework and assignments. These results suggest that teachers in Ontario have some degree of technical expertise in classroom assessment, but they say little about the quality of the assessment interactions between teachers, students, and students’ work in English Language Arts classrooms.

*Teachers’ Experience in Ontario.* Teacher education programs in Ontario are required to provide beginning teachers with practical experience in schools (OCT, 1996), and teachers in Ontario tend to view practical experience as the most effective means of learning about teaching and assessing (OCT, 2003). There is evidence to support the notion that teachers’ concerns in Ontario evolve with experience, as found in other educational contexts. For example, the 2005 OCT survey showed that teachers’ concerns about class size decrease with experience. There also seems to be a relationship between teachers’ experience and their classroom practices. Analysis of the 2002 SAIP data revealed that some of the classroom assessment practices of a group of experienced specialist teachers (n = 105) differed significantly from a comparison group of teachers with less experience and specialization in teaching English Language Arts in Ontario (n = 40). The experienced specialist teachers tracked homework completion and gave feedback to the whole class less frequently, but they

used exemplars from established writers and held writing conferences (i.e. individual or small group feedback) more frequently than the teachers in the comparison group. They also spent far more time marking student work outside of regular classroom hours. Another important difference between these two groups of teachers relates to their perception of difficulty in assessing writing. Almost all of the experienced specialist teachers (96.2%) disagreed that it is “really difficult to judge good writing” (CMEC, 2002c, p.13), whereas fewer (65%) in the comparison group disagreed. This indicates that experienced specialist teachers find assessing student writing less challenging, which suggests that confidence in assessing student writing is gained through experience in teaching English in Ontario.

*Teachers' Ethical Knowledge and Personal Characteristics in Ontario.* Four ethical standards for teaching have been established by the OCT (2008a). They relate to care, trust, respect, and integrity, which are moral virtues in philosophical terms. Although these virtues are not discussed specifically in relation to classroom assessment, the ethical standards are intended to “guide the decision-making and professional actions” (p.5) in classroom practice, which necessarily includes assessment. As such, care, trust, respect, and integrity can be seen as the characteristics that are expected for classroom assessment interactions. Some strong patterns are evident among teachers in Ontario that relate to these characteristics. For example, when the 2003 OCT survey asked teachers to identify the features that are the “most important for good teaching” from a list of options, the two most frequently selected were “inspiring love of learning” and “showing that they care about their students” (p.10). Teachers in Ontario have repeatedly expressed that inspiring student learning is the best part of teaching (OCT, 2003; 2004). This collective belief is important in relation to practical

wisdom about fairness because it suggests that teachers in Ontario orient their practices primarily to helping students learn.

In summary, most English Language Arts teachers in Ontario appear to be fairly confident in their theoretical and technical knowledge. Teaching experience is highly valued, and teachers' confidence in assessing student writing seems to increase with experience specifically in teaching English Language Arts. Many English teachers do report using some of the core practices of assessment for learning, but the existing data reveal little about the quality of assessment interactions in Ontario's classrooms. Most importantly, teachers in Ontario consider their interactions with students, which are at the heart of this study, to be the most important aspect of teaching.

#### Concepts and Questions Situated in the Framework

In the final section of this chapter, I summarize my discussion surrounding the key concepts for this study in relation to a visual illustration of the conceptual framework. My purpose in this section is to indicate how all the concepts discussed to this point are positioned in relation to each other and to my research questions. As such, the balance of this chapter is organized in the order of my three research questions.

#### *Fairness in an Assessment for Learning Interaction*

Figure 2 shows the concept of fairness situated in an assessment for learning interaction in a classroom environment. Fairness is a complex concept, and consensus is limited in how it is interpreted in the existing literature. Although it is broadly associated with all phases of classroom assessment in the principles and standards proposed by educational committees, classroom assessment specialists emphasize a smaller number of

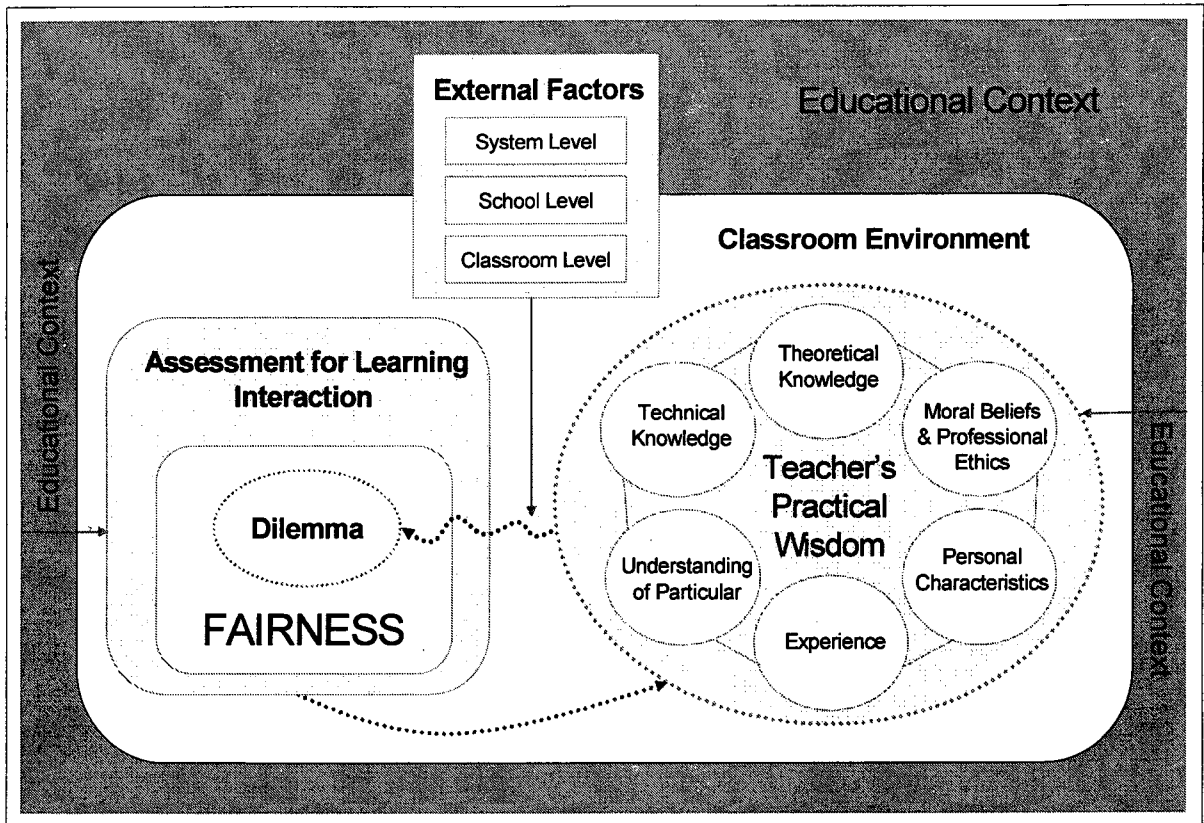


Figure 2. Visual illustration of the conceptual framework for this study.

key aspects, particularly transparency (e.g., clear learning expectations and assessment criteria), opportunity to learn (e.g., alignment of teaching, learning and assessment), opportunity to demonstrate learning (e.g., multiple or varied assessment tasks), and the need for teacher reflection (e.g., avoiding bias, recognizing dynamics). As evidenced in the empirical research, teachers do voice varied concerns about fairness, but the tension between maintaining consistency in their practices (e.g., equivalency of tasks, equal treatment of students) and differentiating according to students' needs (e.g., equitable assessment through modifications or accommodations) comes to the forefront.

In much of the literature reviewed earlier, fairness is discussed in relation to

summative tasks or grading (assessment of learning), and less attention has been devoted to understanding fairness in assessment for learning. In the classroom environment, assessment for learning involves interaction between teachers, students, and students' learning (Allal & Pelgrims Ducrey, 2000) that can be both planned and interactive (Cowie & Bell, 1999).

Assessment for learning is a composite practice that is ideally situated in a constructive environment where teaching, learning and assessment are integrated. Five core practices are involved: a) clear learning expectations and assessment criteria, b) varied approaches used to elicit learning, c) balanced and descriptive feedback for students, d) the adjustment of teaching as a result of the assessment, and e) the active involvement of students in a learning-assessment process (Tierney & Charland, 2007). Teachers' practical wisdoms about fairness in the interactions of assessment for learning may involve aspects of fairness that have already been discussed in

the literature, but others aspects may also emerge. I do not specify in my framework which aspects might be heard in a teacher's practical wisdom because this is my first research question: What are teachers' practical wisdoms about fairness in classroom assessment?

#### *A Teacher's Practical Wisdom about a Fairness Dilemma*

A teacher's practical wisdom is a flexible mental network that draws on theoretical knowledge, technical knowledge, experience, personal characteristics, moral beliefs and ethical knowledge, as well as an understanding of particular circumstances, events or students (Aristotle, trans. 1925/1975, Connolly, 2005, Dunne, 1993). Practical wisdom is situated in the particulars of practice, and it develops through a teacher's learning and experience. The moral-ethical element of practical wisdom is key in understanding practical

wisdom as a teacher's "capacity to act with regard to human goods" (Aristotle, trans. 1925/1975). Figure 2 shows a teacher's practical wisdom as a network of thought about a dilemma relating to fairness in assessment for learning. The notion that a dilemma is a matter of perception is important because differences in teachers' practical wisdoms may mean that a situation is less of a dilemma (issue) for some teachers than others. Fairness dilemmas (or issues) relate to aspects or interpretations of fairness, such as transparency of learning expectations and assessment criteria, opportunity to learn or to demonstrate learning, and teacher reflection about bias. In Figure 2, an arrow returns from an assessment for learning interaction to a teacher's practical wisdom to show that encountering and thinking about a dilemma (or issue) becomes part of a teacher's experience. The illustration is thus one of a cycle, where classroom assessment interaction feeds back into practical wisdom, interplaying in an complex internal network. I do not identify which elements of practical wisdom a teacher might draw on in relation to a fairness dilemma (issue) because this is the second research question: How do these practical wisdoms about fairness inform teachers' decisions in classroom assessment dilemmas?

#### *External Factors Within an Educational Context*

Numerous factors that are external to teachers' thinking have been found to facilitate and hinder classroom assessment practices. Figure 2 shows external factors within the classroom environment and the broader educational context in which the classroom is situated. The straight arrow indicates that although they are external, these factors might influence the teacher's practical wisdom about a fairness dilemma (or issue). For purpose of discussion, I organized these factors into three levels (i.e., classroom, school and system) but

they may interact in practice. I have not identified which factors might facilitate or constrain a teacher's practical wisdom about fairness in this framework because this is the third research question for my study: Which external factors facilitate or constrain teachers' practical wisdoms about fairness in classroom assessment dilemmas? In the next chapter, I explain the design of my study, and the methods that I used to gather and analyze information relating to these three research questions.

## CHAPTER FIVE: METHODOLOGY

*Aristotle believed that if one's subject matter is the practical and communal life of persons then one must renounce the methodological purism and the possibilities of generalization and precision that are legitimate aspirations in properly theoretical endeavours."*

Joseph Dunne (1993, p.18)

This chapter focuses on the design, information sources, tools, process, and strategies used in this study. I begin by giving my rationale for choosing a case study design, and then I describe the sources from which I obtained information. I also describe the tools I used to gather information, and how I went about developing them. I explain the process of my research in the general sequence that it unfolded. In the final sections of this chapter, I discuss the role of narrative in my project, and I describe the strategies I used to ensure the quality of my work.

### Research Design

This research project was designed as an instrumental multi-case study using complementary methods to describe six individual cases embedded in a specific educational context. The individual cases represent experienced English teachers, and the educational context is Writing in Grades 7 to 12 English Language Arts classrooms in Ontario. I chose to use a case study design because I was seeking to understand a complex concept in practice. Case study research is well-suited to this purpose because of its "particularistic, descriptive, and heuristic" nature (Merriam, 1998, p.29). Essentially, this means that case study research involves rich description that aims to shed light on a specific concept, event, person or program. Merriam explains that the "specificity of focus" makes case study research "an especially good design for practical problems – for questions, situations, or puzzling

occurrences arising from everyday practice” (p.29), and she notes that case studies “can bring about the discovery of new meaning, extend the reader’s experience, or confirm what is known” (p.30). Given that I aimed to use the individual cases in my study to provide insight into fairness in classroom assessment, the design is also instrumental. According to Stake (2005), a case study is instrumental when its value is not purely intrinsic, and it is “examined mainly to provide insight into an issue” (p.445). With the diversity of interpretation currently found in the literature on fairness in classroom assessment, I felt that a small number of cases would reveal more of its complexity than a single case study, without overly restricting the depth of understanding needed for my research questions. I designed my study to use both qualitative and quantitative information in a complementary manner to produce a “fuller picture” than would be possible with a single method (Erzberger & Kelle, 2003, p.470). While neither qualitative nor quantitative information is privileged from my philosophical perspective, I relied more heavily on qualitative information in this study due to the nature of my questions.

#### Information Sources

I gathered information from multiple sources to inform the development of research tools and answer my research questions. Specifically, I worked with existing large-scale survey data, a variety of educational documents, and teacher participants.

##### *Large-Scale Survey Data*

Before working directly with teachers, I used large-scale survey data to establish the recruitment and selection criteria, support the development of research tools, and help describe the educational context of this study. The large-scale data were drawn from the

*Teacher Questionnaire* of the 2002 Writing School Achievement Indicators Program (SAIP), which is a pan-Canadian assessment program undertaken by the Council of Ministers of Education, Canada (CMEC). The SAIP, which is now called the Pan-Canadian Assessment Program (PCAP), is designed to sample the academic achievement of 13 and 16 year olds across Canada. Four subjects, being science, mathematics, reading and writing, are assessed in rotating years (CMEC, 2009). In 2002, the SAIP focussed on writing, and the teachers of the students who were selected for the assessment completed a questionnaire containing 33 items that asked about their years of teaching experience, type of education, beliefs about writing pedagogy, and classroom practices (CMEC, 2002b). The CMEC gave me written permission to use the 2002 SAIP data on October 1<sup>st</sup> 2007. I explain more about how I used this dataset in the section on my research process in this chapter.

### *Existing Documents*

In keeping with my philosophical orientation, and with the emphasis on context in case study methodology (Stake, 2005; also more generally in Patton, 2002), I sought to situate the individual cases in my study in relation to the broader educational context. There are many kinds of documents can be used to gather information unobtrusively, which Merriam (1998) notes are “particularly good sources for qualitative case studies because they can ground an investigation in the context of the problem being investigated” (p.126; also Marshall and Rossman, 1999). In addition to the existing research literature, I used a variety of documents from the following educational organizations in Ontario to investigate and describe the educational context:

- Council of Ministers of Education, Canada
- Education Quality and Accountability Office

- Government of Ontario
- International Association for the Evaluation of Educational Achievement
- Ontario College of Teachers
- Ontario Ministry of Education
- People for Education
- Toronto District School Board

The documents I used included reports based on the surveys conducted annually for the OCT, as well as curriculum and policy documents, a legal document, information pages from Internet sites, a school board census, and summarized results from large-scale assessments. These documents are publicly available on the Internet and included in my reference list.

### *Teacher Participants*

The teachers who participated in this study were all volunteers on the *Teachers' Network*, which was an interactive feature of the OCT Internet site that encouraged teachers to share their expertise with other teachers. I purposively selected a pool of teachers from the *Teachers' Network* volunteers as potential participants for this study. Patton (2002) explains the rationale for purposefully selecting participants in his text on qualitative methods:

The logic and power of purposeful sampling derive from the emphasis on in-depth understanding. This leads to selecting information-rich cases for study in depth. Information-rich cases are those from which one can learn a great deal about issues of central importance to the purpose of the research, thus the term purposeful sampling. (p.46)

Similar ideas have been expressed specifically in relation to case study research. Stake (2005) argues that cases should be selected, first and foremost, based on the learning opportunity they provide for researchers, and Flyvbjerg (2004) notes that selection on this basis amplifies the utility of a small number of cases, which economizes research resources (i.e., time and money). To determine the characteristics of teachers who would be most likely to provide information-rich cases for my study, I analyzed the pan-Canadian data from the 2002 SAIP

teachers' questionnaire. Specifically, I looked at the relationship between years of teaching experience, type of degree, and level of specialization in teaching English Language Arts. I concluded that teachers who had at least 10 years of teaching experience, held a relevant undergraduate degree, and identified themselves as specialists in teaching English Language Arts would be the most likely to provide information-rich cases. As a result of having been part of a purposefully selected pool, the six teachers who agreed to participate in this study share these characteristics. They engaged in this study voluntarily, and they were not paid as participants. As a token of my appreciation for their generosity in sharing their time and practical wisdom, I sent each participant a bookstore gift certificate following the final phase of information gathering for this study.

#### Tools for Gathering Information

Three research tools were developed specifically to collect information for this project: a background questionnaire, a vignette questionnaire, and a template for individualized interview protocols. In planning this study, I responded to Brookhart's critique of the short dichotomous items used by Green and colleagues (2007) in their survey on the ethics of classroom assessment (discussed in Chapter Two). I included features in the design of my questionnaires and the protocol template to allow for the flexibility I needed to capture my subject's complexity, while providing sufficient structure to aid participants in elucidating their practical wisdom such that it related to fairness as much as possible.

#### *Background Questionnaire*

The main purpose of the Background Questionnaire was to collect information for the description of the individual cases, focusing in particular on areas that were pertinent to the

participants' practical wisdoms. Although background information could have been gathered orally with the small number of participants in this study, I chose to develop a written electronic questionnaire to reduce the time and travel needed for interviews, and to permit participants to answer quickly and easily. I used the encrypted version of Survey Monkey, which is an Internet service that provides software to create surveys and collect responses securely online. I determined the formatting for each item based on guidance from Colton and Covert (2007), and by looking at existing teacher questionnaires (e.g., SAIP, OCT). Knowing that I would be analyzing responses from a small number of purposefully selected participants, rather than the large numbers usually associated with survey design, I was able to include more open-ended options and minimize the length of the questionnaire. A copy of the Background Questionnaire is appended (Appendix C).

The blueprint that I used for the development of the Background Questionnaire is provided in Appendix D, along with the rationale for the items included in each section. The items relating to experience and education (sections 3 to 6) provided information that was useful for my second research question. The items on the ethics of classroom assessment (section 7) were included primarily to provide a sense of each participant's general approach to assessment as part of their case portrait, but some of this information was also useful for my first research question. The nine items in this section were drawn from Green and colleagues' (2007) survey and adapted to suit the context of my study. I chose these particular items from the 36 in Green and colleagues' survey because they showed the highest levels of disagreement among respondents, and I felt that this would help highlight the differences and similarities between the participants in my study. While some of these

items relate to assessment of learning rather than assessment for learning, they are all associated with aspects of fairness that have been identified in the existing literature (e.g., opportunity to learn, teacher bias, etc.). Before using these items, I discussed my intention with one of the authors of the original survey, and I received feedback on the wording of my adaptations. The instructions and items of this questionnaire were also revised based on feedback from two colleagues in the doctoral program who both had experience as educators in Ontario's educational system.

### *Vignette Questionnaire*

The second tool that I developed for this study was the Vignette Questionnaire. Vignettes are brief stories or simulations that are used in social research to elicit responses from participants, and they have been found to be especially useful in research on ethical dilemmas and sensitive topics (Barter & Renold, 1999; Hughes & Huby, 2004; Wilks, 2004). A copy of the Vignette Questionnaire is appended (Appendix E).

This questionnaire contains eight written vignettes, each of which is followed by the same three open-ended questions that correspond directly to my research questions. Specifically, these questions were designed to provide information about each participant's recommendations for fair assessment (RQ #1 on practical wisdom *about* fairness), how they came to this recommendation (RQ #2 on how their practical wisdom *informed* their decision), and which factors they thought might facilitate or hinder a teacher in following their recommendation (RQ #3 on external factors). Appendix F shows the blueprint I used to develop the Vignette Questionnaire. The content of the vignettes was inspired by my own experience and by the experiences of the teachers who are described in existing empirical

research on classroom assessment (discussed in Chapter Two). I designed them to reflect the aspects of fairness that are prevalent in the conceptual literature, and then tailored them specifically for the study's educational context. The aspect of fairness illustrated in each vignette is identified in Appendix G specifically in reference to McMillan's (2007) components of fairness and *The Student Evaluation Standards* (JCSEE, 2003). The fairness issue in each vignette involves at least one of the core practices of assessment for learning (Tierney & Charland, 2007), where the teacher is poised at a moment of decision in either a planned or interactive cycle (Cowie & Bell, 1999). Each vignette is also situated in a writing activity that reflects the mandated curriculum for English Language Arts in Grades 7 to 12 in Ontario (based on analysis of OME documents), and that is commonly used by English Language Arts teachers in Ontario (based on secondary analysis of the SAIP 2002 data). A short description of each vignette follows:

*Vignette A.* The first vignette describes a point in interactive assessment for learning where a teacher notices three students who are having difficulty summarizing an article. One of the three missed the lesson on summarizing while he was on a school band trip, and the other two were present, but inattentive during the lesson. The teacher adjusts her plans and sets out to work with the band student, but she is unsure of how to respond to the inattentive students. While she realizes that the band student needs the same opportunity to learn as the rest of the class, the inattentive students have already had this opportunity. The inattentive students might benefit from the additional lesson, but this differentiation would give them more opportunity to learn than the rest of the class.

*Vignette B.* This vignette describes a teacher who is in the process of interpreting

information about a student's writing during a planned assessment for learning event, based on set ideas about the nature of good writing. Before the teacher is able to act by giving the student written feedback, his interpretation is unintentionally challenged by a colleague's ideas about the student's writing. While the teacher wonders how to proceed, and seems to be teetering on the edge of change, it is not clear whether he is aware that his preference for a certain writing style might bias his assessment.

*Vignette C.* The teachers in this vignette are planning a cross-curricular unit on writing historical fiction. They develop a rubric to give to their students at the beginning of the unit, but one of the teachers is unsure about sharing the learning expectations and assessment criteria with the students beforehand. The issue here is transparency, which is not only an aspect of fairness emphasized in classroom assessment theory, but also underlies one of the core practices of assessment for learning. This vignette illustrates a controversy in the use of rubrics for teaching writing as some teachers believe they provide parameters to scaffold learning while others feel they constrain creativity with narrowly prescribed rules.

*Vignette D.* This vignette illustrates a point in the interactive assessment process where a teacher notices a problem, and has some idea of its significance, but doesn't know how to respond to it. The problem comes from involving students in an assessment activity without first establishing a constructive classroom environment with clear assessment criteria and procedures. As a result, a student who values mechanics over content feels entitled to ridicule the writing of another student during a peer assessment session. Peer assessment of this type is not only biased, it is counter-productive to learning in that it may have a negative impact on the emotional and academic development of the students involved.

*Vignette E.* This vignette describes a planned assessment event where a teacher is in the process of giving students written feedback on their research project plans. When she realizes that she has written more on some students' work than others, she is concerned about the consistency of her feedback. The teacher's uncertainty suggests that she wants to treat her students equally so that they have a common learning experience, but she has other demands on her time. Although equality is not an aspect of fairness that is emphasized in classroom assessment theory, teachers in classroom-based research do frequently express concern about the equal treatment of their students, and in principle, consistency in assessment procedures is associated with fair assessment.

*Vignette F.* The teacher in this vignette is engaging in the first planned assessment event of the school year when her assumptions about the content of a particular student's writing folder are challenged. Knowing that the student is formally identified as gifted, the teacher is surprised by the average writing in the folder. She does not seem to question the stereotype informing her expectation that this student's writing skills would be advanced. The need for teachers to reflect critically about their biases, expectations, power and roles comes up in both the conceptual and empirical research relating to fairness in classroom assessment. Although this teacher is considering how to act on the information that has been elicited, it is not evident that her reflection will lead to fairer assessment.

*Vignette G.* This vignette describes a moment in the interactive assessment process where a teacher notices a problem, recognizes its significance in relation to her expectations, and is considering how to respond. Circulating around the classroom as the students write in their journals, the teacher notices many mechanical errors, and she expects a higher quality of

writing at this level. Although she wonders about giving the class this feedback immediately, she does not seem to have considered making her expectations clear beforehand. The teacher makes two assumptions in this activity. First, that the students have the knowledge and skill to produce the quality of writing that she expects, and second, that all forms of writing must adhere to the conventions of standard English.

*Vignette H.* In this vignette, a teacher assesses her students' writing portfolios at a planned point partway through a course. She is pleased with the excellent quality of the writing in one student's portfolio, but she finds that it is missing several assignments. She considers how to act on this information, and she is not sure if she should talk to the student about the missing assignments, or have the student either complete all the tasks already assigned or work on a more challenging task. The issue illustrated here is whether the students in a class all need to complete same number of tasks for the assessment to be fair. Essentially, the teacher in this vignette has noticed that the student's writing exceeds the learning expectations, and she wonders whether she should insist on consistency across the class, or differentiate the number and nature of assessment tasks for this student.

In my decision to use vignettes for this project, I considered both their strengths and limitations. One of the benefits of using vignettes in discussing a topic such as fairness is that they can reduce the social desirability or defensiveness of responses. Barter and Renold (1999) explain that:

As commenting on a story is less personal than talking about direct experience, it is often viewed by participants as being less threatening. Vignettes also provide the opportunity for participants to have greater control over the interaction by enabling them to determine at what stage, if at all, they introduce their own experiences to illuminate their abstract responses. (p.3)

A strategy that is suggested to support the disclosure of information with vignettes is to direct participants to “adopt the role of consultants” (Hughes & Huby, 2004). As such, in the instructions included on the Vignette Questionnaire, I asked participants to recommend the fairest course of action for the fictional teacher in each vignette. A challenge in using vignettes is that the relevance of information elicited is dependent on the quality of their construction. Not only does the content need to be closely aligned with the research topic, the format must be appropriate for the age and type of participants, and the characters and events portrayed in the vignettes need to engage and maintain the participants’ interest (Barter & Renold, 1999; Hughes & Huby, 2004; Wilks, 2004). A strategy to ensure the quality of vignettes is to base their construction on a combination of practical experience and research literature, as shown in my blueprint.

Vignettes can be beneficial for research on complex topics where a “uniform situation” helps focus responses and provides a basis for comparison (Barter & Renold, 1999, p.3). On the informal written feedback that some of my participants provided, they commented that the vignettes were enjoyable, relevant to their practice, and helpful in guiding their responses. However, the structure provided by vignettes is problematic if participants’ responses are directed to the degree that the research is circular (Merriam, 1998; Wilks, 2004). Several authors recommend strategies, such as using open-ended options, probing responses (e.g., follow-up interview) or complementary methods, to ensure that participants are not limited to a set of predetermined options based on what has already been established in the literature (Barter & Renold, 1999; Wilks, 2004). I used all three of these strategies in this research project.

The development and review process for the Vignette Questionnaire was lengthy. Initially I drafted 15 vignettes, and I developed a supporting structure for each one in which I identified the aspect of fairness involved (McMillan, 2007; JCSEE, 2003), as well as the core practice and process in assessment for learning (Cowie & Bell, 1999; Tierney & Charland, 2007), and the writing activity (OME, 2006, 2007a, 2007b; CMEC, 2002b). In total, I asked six reviewers for feedback on different aspects of the Vignette Questionnaire. The content of each vignette was reviewed by two established scholars. One has produced numerous texts on grading, classroom assessment theory, and educational research, and the other has worked on a several research projects with a focus on classroom interaction and assessment for learning. Feedback on the relevance of the vignettes for the particular context of the study was provided by two former teachers with experience in Ontario's educational system. One was the head of an English Department at a large secondary school (now retired), and the other is a school board curriculum consultant. The final two reviewers were the same colleagues who reviewed the Background Questionnaire. They engaged in a systematic review of the Vignette Questionnaire, and they provided feedback on the clarity of the instructions, the content and relevance of the vignettes, and the format of the questionnaire overall. They also provided timed responses to a selection of vignettes to help me determine the total number in the questionnaire. Based on the information provided by all six reviewers, I selected and revised the eight vignettes that were the most relevant for the purpose and context of this study. The Vignette Questionnaire was administered through the encrypted version of Survey Monkey, and the participants were able to respond in writing at their leisure and convenience.

### *Interview Protocols*

I created an individualized protocol for each interview based on the participants' responses to my two questionnaires. The template for these protocols is appended (Appendix G). In determining which items or issues to discuss further during the interviews, I considered patterns, gaps, and unanticipated responses in relation to the research questions. The interview questions served three main purposes. Some were designed to confirm or clarify my interpretation of the participants' written responses, and others aimed to elicit more detail in specific areas, especially about any responses that were ambiguous or confusing. Questions were also designed to encourage participants to discuss *any* aspect of fairness that they felt was important for classroom assessment, whether or not it was presented in the vignettes.

### Research Process

The design of my research project required an iterative process with the gathering and analysis of information occurring at multiple points. I have organized my narration of this process to follow the sequence of main events, beginning with gaining access and ethical approval, and conducting the initial analyses, and then moving to selecting and working with participants, and finally to analyzing and interpreting the information they provided.

#### *Gaining Access and Ethical Approval*

I submitted the initial application for ethics approval for this research project to the Social Sciences and Humanities Research Ethics Board (SSHREB) at the University of Ottawa on August 31<sup>st</sup>, 2007. However, I could not commence the research until I had written approval from a school board to gain access to teachers. I approached three school

boards with this goal. One school board was concerned that my study would conflict with a professional development program that they were implementing, and declined to participate. A second school board expressed considerable interest in the vignettes, but they wanted to use them in focus groups with both novice and experienced teachers. They envisioned the process as a professional development activity, in which I would be permitted to use data with the approval of a school board representative. Despite several lengthy conversations and a revised proposal, I was unable to come to an agreement with this school board that served both my research purpose and their interests. The third school board I contacted did not have a formal process for applications from external researchers, and several months passed as my proposal moved along an administrative chain. During this time I was occupied with the initial analysis and the development of my research tools, but with the end of the school year fast approaching, and no guarantee of a positive response from the third school board, an alternate approach was needed. Following a committee member's suggestion that I might try recruiting participants through a professional organization, I contacted the OCT about using the *Teachers' Network*. As a certified teacher in good standing with the OCT, and a volunteer on the *Teachers' Network*, I already had access to the Members' Area of the site, so I was able to email (blind) other teachers with an invitation to contact me about participation. I revised my application to the SSHREB at the University of Ottawa accordingly, and received full approval to proceed with this research project on April 23<sup>rd</sup> 2008.

#### *Analyzing Secondary Data and Relevant Documents*

I analysed large-scale data for three purposes: to establish the criteria for selecting participants, to support the development of research tools, and to help describe the

educational context of the study. For the first purpose, I conducted an initial analysis using the data from the 2002 SAIP questionnaire for teachers across Canada ( $n = 4,494$ ). I used SPSS, which is a software program for quantitative analysis, to look at the variables relating to teachers' education, their years of experience, and teaching preferences. I ran frequency distributions, cross-tabulations, and chi-square tests (with Cramers'  $V$ ) to determine the strength of association between these variables. I also calculated the index of qualitative variation (IQV) to determine the characteristics of teachers who identified themselves as specialists. The IQV is "based on the ratio of total number of differences in the distribution to the maximum number of possible differences within the same distribution" (Frankfort-Nachmias & Leon-Guerrero, 2006). With this analysis I identified the characteristics of teachers who would be most likely produce information-rich cases for my study.

For my second purpose, which was to support the development of my research tools, I analysed the responses of a sub-sample of teachers who responded the 2002 SAIP questionnaire. All of the teachers in this sub-sample ( $n = 105$ ) shared four characteristics. They all identified themselves as specialists, had at least 10 years of experience teaching English Language Arts, had a relevant bachelors degree, and taught English writing to 13 or 16 year-old students in Ontario. I examined the responses of these experienced specialist teachers to 16 categorical items relating to their teaching and assessment practices. The resulting descriptive statistics provided the detail I needed to tailor the vignettes meaningfully for the educational context of this study.

For my third purpose, which was to help describe the educational context, I looked at the 2002 SAIP data for English teachers in Ontario ( $n = 415$ ), as well as several reports on

the surveys conducted for the OCT from 2003-2006 (n = 1000) and a variety of local documents (described previously). In this process I was looking for two specific types of information. First, I was interested in any factors that would shed light on teachers' practical wisdoms in Ontario, such as the educational requirements for teaching. I particularly wanted to know about teachers' assessment knowledge, and the degree to which English Language Arts teachers in Ontario might be using assessment for learning. The second type of information I was looking for related to any external factors that might influence classroom assessment in Ontario's educational system. The understanding I gained through this early work ultimately helped me interpret the participants' practical wisdoms about fairness relative to the educational context. Gaining this understanding was especially important for my research design, where multiple cases with varying classroom environments were embedded in one educational context.

#### *Recruiting and Selecting Participants*

In discussing research designs, Patton (2002) notes that there are “no rules for sample size in qualitative inquiry” (p.244). However, the rough number of participants to include is suggested by the purpose of a study, especially in terms of the tension between depth and breadth. In the initial plan for this study, I followed Patton's (2002) advice in setting a minimum number for “reasonable coverage of the phenomena given the purpose” (p.246). I proposed to include at least three participants, with a maximum of six to ensure that it would be feasible for me to gather in-depth information for each case. As such, I set out to purposefully recruit and select a small number of participants based on the likelihood of their becoming information-rich cases for my study.

I used the search feature of the *Teachers' Network* (OCT) to identify teachers who had volunteered to help other teachers specifically in the areas of classroom assessment, English curriculum, or literacy. I then used the *Find a Teacher* feature of the OCT site to check teachers' Certificates of Qualification and determine if they met the selection criteria (i.e., a minimum of 10 years experience, relevant undergraduate degree, and qualifications to teach English Language Arts in Ontario) for my study. Teachers' Certificates of Qualification are publicly available on the OCT Internet site. Through this process I identified 35 teachers who seemed to meet the selection criteria. The *Teachers' Network* (OCT) had an email feature that allowed one member to contact other members without revealing destination addresses. As such, I emailed these teachers with an invitation to email me if they were interested in participating in my research project. The email text was approved by the SSHREC at the University of Ottawa (Appendix H). Of the 35 emails sent, 21 did not reply, and 3 were not deliverable (i.e., non-functioning email addresses). Of the 11 teachers who did express interest in the study, four were currently teaching elementary students, so they were necessarily excluded. I sent the remaining seven an information sheet (Appendix I) and a consent form (Appendix J), which were required and approved by the SSHREC. One teacher subsequently decided that the time commitment might be problematic for her schedule, and the remaining six consented to participate in the study. I stopped recruiting participants at this point because I felt that a larger number would compromise the depth of understanding that I wished to attain from each individual case.

### *Working with Participants*

As I received participants' signed consent forms, I opened a file for each participant,

and I used Survey Monkey to send the Background Questionnaire to an email address they provided. The six participants returned their consent forms between April 29 and May 19, 2008, which staggered the gathering of information in time. Once they had completed the Background Questionnaire, I sent each participant the Vignette Questionnaire using the same system. Participants were encourage to take their time in responding to the vignettes, and most took over a week (range from 2 to 26 days). When participants returned the Vignette Questionnaire, we scheduled a time and place at their convenience for the follow-up interview.

The interviews were digitally recorded and lasted from 50 to 75 minutes. I followed many of Seidman's (1998) interview techniques to ensure that the interviews were effective and fluid. For example, I concentrated on listening actively at different levels in order to understand to the participants' voices (i.e., public, personal) as well as their words. I kept my questions open, and I went back to points of interest afterwards to allow them to complete their thoughts or stories. I was flexible in following the protocol during the interviews in order to give the participants the opportunity to discuss any aspect of classroom assessment that they felt posed a fairness dilemma or issue, and to give them the lead in when or how they discussed their own practices. I also adopted the same approach that I took in designing the Vignette Questionnaire, where participants were encouraged to respond as a consultant (i.e., recommendations for the fictional vignette teachers). As the participants were all experienced teachers who had previously volunteered to assist other teachers, this way of working together seemed natural and comfortable. The participants were given a copy of their questionnaire responses during the interview, and a verbatim transcript of the recording

afterwards, and the informal feedback from several suggests that they were also comfortable with the working process.

### *Analyzing and Interpreting the Information Gathered*

Once interview recordings were fully transcribed, they were coded along with the participants' responses to the vignettes and the open-ended items from the Background Questionnaire. I used both deductive and inductive processes for the analyses. Deductive analysis involves matching the participants' responses to a set of categories that are established beforehand (Patton, 2002; Yin, 2006), whereas inductive analysis involves what is often referred to as a constant comparative method where categories emerge (and merge) as the analysis proceeds and patterns become evident (Glaser & Strauss, 1967; see also Creswell, 2008; Merriam, 1998). Patton notes that despite the association of deduction with experimental research and induction with naturalistic inquiry, they are often combined in research practice. To understand the various interpretations or aspects of fairness that my participants discussed, I initially coded their responses with a list derived from the principles and standards, but these categories were adjusted as my analysis progressed. Some were more clearly defined, some were eliminated or amalgamated, others were expanded, and some categories emerged as new themes. This required that I return to the responses that had been previously coded, moving between cases during the analysis. For all three research questions, the analysis resulted in adjustment to the initial categories that I began with, and these results will be presented and discussed in the following chapters.

### Role of Narrative

In simple literary terms, a narrative is a "story, actual or fictional, expressed orally or

in writing” (Harris & Hodges, 1995). In social research, the term *narrative* carries epistemological and political meaning that goes beyond its literary definition. For example, in *The Dictionary of Critical Social Sciences*, Young and Arrigo (1999) define narrative as a “term that connotes doubt about the value of grand theories, universal truths, global explanations, totalizing practices, or master codes with which to understand everything” (p.222). Methodological texts note that the growing interest in narrative in the later half of the 20<sup>th</sup> century has resulted in the development of narrative designs and analysis in social research (e.g., Creswell, 2008; Riessman, 2004). In education, narrative inquiry has emerged as a distinct methodology that studies personal experience as story (Connelly & Clandinin, 2006). My work is similar to narrative inquiry in that I value the experiences of individuals, and in some cases I was able to get beneath the “surface manifestations” of my participants’ beliefs to hear their “experiential narrative origins” (Connelly & Clandinin, 1986). Narrative is clearly seen in my work in several ways. I used narratives, in the form of vignettes, to elicit information from my participants. I also included their narratives within the cases reported in the next chapter, and the description I give of my research has a narrative style.

However, my work differs from narrative inquiry in three important ways. First, it uses hypothetical dilemmas involving fictional teachers and it is strongly rooted in research literature, whereas narrative inquiry focuses on the personal experiences of living teachers without emphasizing problematic situations or pre-existing knowledge (Creswell, 2008; Connelly & Clandinin, 1986). Second, the proponents of narrative inquiry are interested in individual experience over a long period. For example, Clandinin and Rosiek (2007) see narrative as a “form of representation that describes human experience as it unfolds over

time” (p.40), and Polkinghorne explains (in a transcribed conversation, Clandinin & Murphy, 2007) that he encourages his students to “do individual case study narrative life stories” (p.633) that describe the “life movement of a particular person” (p.634). In contrast, I used short stories related by my participants about specific moments or events in their lives, and I look at themes that emerged across their cases to understand a problem that I had identified beforehand. The third way that my work differs from narrative inquiry is in the process of text construction. Whereas the researcher and participant work closely together to develop the story that is to be told in narrative inquiry, I retained control of the content of my dissertation. Stake (2005) explains this difference in discussing the role of storytelling in case study research:

Even when emphatic and respectful of each person’s realities, the researcher decides what the case’s ‘own story’ is, or at least what will be included in the report. More will be pursued than was volunteered, and less will be reported than was learned . . . . It may be the case’s own story, but the report will be the researcher’s dressing of the case’s own story. This is not to dismiss the aim of finding the story that best represents the case, but to remind the reader that, usually, criteria of representation ultimately are decided by the researcher (p.456)

My case reports are based on my participants’ responses and stories, but the writing was not collaborative. Given the strong role that narrative plays in my research, it is in a “borderland space” (Clandinin & Rosiek, p.58) with narrative inquiry. However, with my use of hypothetical dilemmas rather than lived stories, my interest in a particular concept rather than a person, and my analysis across as well as within individual cases, I do not follow the tenets associated with narrative inquiry. My work is a multi-case study that draws on different ways of knowing, including theoretical, empirical, practical, moral, and narrative knowing, for an instrumental purpose.

## Research Quality

Essentially, the quality of all research depends on the thoughtfulness of the design, the honesty of the engagement, and due care in the process (Borman, Clarke, Cotner & Lee, 2006). Research quality has traditionally been judged in terms of reliability and validity. However, the acceptance of “nonpositivist orientation” (Lincoln & Guba, 2000, p.163) and the use of multiple methods open rhetorical options for discussing the quality of a research project (Sandelowski, 2003). As recommended by Creswell and Miller (2000) and Patton (2002), I considered the criteria that were most consistent with my research design and that reflected my philosophical orientation, and I concluded that I wanted my work to be credible, thorough, and meaningful. As a quality for research, credibility refers to the relationship between the researcher’s interpretations and the participants’ views (Tobin & Begley, 2004). For research that aims for depth rather than breadth, such as a case study, completeness is considered a more appropriate quality than either objectivity or confirmability (Tobin & Begley). Although completeness is meant to refer to the thoroughness of the results, it suggests that the research topic is closed. Because I see research as part of an ongoing dialogue, I wish to leave the possibility of further work open. As a result, I aim simply for thoroughness as a desirable quality for my work. I also saw meaningful, rather than generalizable or transferable, as a more appropriate and useful goal for the results of my study. Generalizability is often considered a desirable quality for studies that aim to apply results from a sample to a larger population. Transferability has been suggested as an alternative for interpretive inquiry (Tobin & Begley, 2004), but the meaningfulness of transferred results has been questioned (Patton, 2002). My aim with this research was to

provide insight from a set of cases in a particular educational context, which would be meaningful for similar cases in the same or comparable contexts. To enhance the credibility, thoroughness and meaningfulness of the results of my work, I incorporated seven strategies in its design and execution, which were to reflect, document, anchor, respect, triangulate, return, and listen. Each of these seven strategies is described below:

*1. Reflect:* The need for researchers to reflect on their roles and assumptions in the research process is often emphasized in methodological texts (e.g., Maxwell, 2005; Seidman, 1998; Yanchar & Williams, 2006). Recording events as they transpired in a journal helped me see the progress of my project, which was useful in directing my next steps during the early phases of the project. During the lengthy process of transcribing, analyzing and writing, I used my journal frequently to keep track of emerging or changing categories, and the thoughts and realizations I had in the process. Although keeping a journal was helpful in this way, I found that the most productive form of reflection during my research was in collaboration with colleagues. I was able to discuss any uncertainties I felt in the process, and work through many questions and challenges with a few close colleagues who were similarly engaged in doctoral research. This collegial support in the reflective process was invaluable.

*2. Document:* To obtain the degree of detail that is desirable for description and interpretation in case study research (Borman et al., 2006; Merriam, 1998), I documented all particulars of the educational context and the individual cases that appeared to be relevant for my research questions. Interview recordings were also transcribed verbatim, as Seidman (1998) and Maxwell (2005) recommend, and I logged questions and ideas that came up during the analysis. Detailed files were organized for each case and for the cross case

analysis in both hard copy and electronic format. Following the advice of Anfara, Brown and Mangione (2002) and Marshall and Rossman (1999), I aimed to make my research process as transparent as possible to enhance the credibility of my interpretations and provide evidence of the thoroughness of the work.

3. *Anchor*: For research to serve the future, I believe that it must be anchored in the past and present. I hope I have done this to some extent by drawing on the existing literature for my conceptual framework and to develop my research tools. A review of relevant research is a good first step for building questionnaires (Colton & Covert, 2007), and I also communicated directly with the authors upon whose work I drew to ensure the accuracy of my interpretation of their work. The analysis of secondary data and documents also helped me understand the characteristics of the educational context in the present.

4. *Respect*: As the sole researcher, I collected, analyzed, and interpreted information provided by the participants. In this regard, I maintained fairly conventional researcher-researched relationships during this project. What was important to me, as an educator, was to approach each of the participants with professional courtesy and appreciation for their time and the information they shared. I believe that I maintained a respectful stance during the interviews, and I gave each participant the opportunity to correct or clarify my interpretation of their responses. From my perspective, this was beneficial because the collegial tone of the interviews permitted frank discussion about fairness issues in classroom assessment.

5. *Triangulate*: The use of multiple information sources and analytical methods for triangulation is a strategy that is frequently discussed in literature on case studies (e.g., Merriam, 1998; Stake, 2005; Yin, 2006). However, Erzberger and Kelle (2003) note that

triangulation is a metaphor with different meanings, and its use should be clarified. I used multiple methods (e.g., secondary analysis, written questioning, and oral interviewing) to gather and analyze information from multiple sources (e.g., large-scale data, research literature, government documents, internet sites, and individual responses). Given that I used these methods and sources more for complementary rather than confirmatory purposes, the triangulation in this study is essentially a “complementarity model” (Erzberger & Kelle, p.469). In this type of triangulation “qualitative and quantitative results may relate to different objects or phenomena but may be complementary to each other and thus can be used to supplement each other” (Erzberger & Kelle, p.466). However, I did find during the analysis that multiple sources also served to confirm some information, and eliminate small errors. For example, one participant’s response for the number of years she had been teaching in Ontario did not correspond with other items on the Background Questionnaire or with the information on her Certificate of Qualification (OCT). During the interview she confirmed that the incongruent response was an error (typo). More importantly, I found that the use of multiple and varied sources resulted in a wealth of information for each case.

6. *Return*: Merriam (1998) describes analysis as “having a conversation with the data, asking questions of it, making comments to it, and so on” (p.181). Thinking about analysis this way means that the first pass through each transcript or set of responses is just a beginning. Having a follow-up interview scheduled after the participants’ responded to the vignettes allowed me to return to them with specific and probing questions, and we did, literally, have conversations about the data. I also returned to the participants’ interview and questionnaire responses repeatedly as I constructed and rearranged working tables with

evolving categories. As Merriam suggests, this is a necessary process for meaningful sense.

7. *Listen*: The final strategy that I used was adapted from the notion of member checking in qualitative research (Creswell, 2008; Merriam, 1998; Maxwell, 2005). The term *member checking* suggests the possibility of a single, correct truth (Tobin & Begley, 2004), and it concentrates on the content of an interview (Clarke & Robertson, 2001). I prefer the notion of listening because it suggests that differing perspectives may be heard while the control of the final text remains with the researcher. Listening to participants' responses was obviously central to my study, but I found listening to participants respond to my interpretation of their responses equally important. In keeping with Clarke and Robertson's suggestion, I sought informal feedback by email from participants about their experiences in the research process. In addition, I listened to a small group of researchers and educators who reviewed the design and content of my research tools. With feedback from these colleagues and the study's participants, I have benefitted not only for the purpose of this research, but in my learning about the process of conducting research in education.

In summary, I designed this research project to better understand fairness specifically in assessing student writing for learning in the classroom environment. I chose a multi-case study approach to draw on the practical wisdom of six experienced specialist teachers embedded in the context of English Language Arts in Grades 7 to 12 in Ontario. The results of this research are presented in the following two chapters, first as individual cases, and then thematically across the cases.

## CHAPTER SIX: CASE BY CASE

*Their story, yours and mine – it's what we all carry with us on this trip we take, and we owe it to each other to respect our stories and learn from them.*

William Carlos Williams to Robert Coles (Coles, 1989)

In this chapter I describe six individual cases. Each case involves one participant as a teacher working in Ontario's educational system, and each is presented using the same organizational structure. The first section of each case is a portrait that includes the participant's current teaching assignment, teaching experience and education, views about teaching, and position on ethical issues in assessment. The second section contains the participant's recommendations for fair assessment, and the third describes the elements of practical wisdom that informed the participant's recommendations. The fourth section gives the external factors identified by the participant, and it is followed by a brief summary of the case. The six participants, using the pseudonyms Kevin, Lorena, York, Tanar, Lucy and Amada, are all experienced teachers who have specialized in teaching English Language Arts in Grades 7 to 12 in Ontario's publicly-funded educational system.

### Kevin: The Same Hoops

#### *Kevin's Portrait*

The quotations given in Kevin's case portrait are from his responses to the background questionnaire (05 May 2008), the vignette questionnaire (03 June 2008), and his interview transcript (09 June 2008).

*Teaching Assignment.* Kevin taught Grade 9 and 10 students in applied and academic English courses and Grade 12 students in an English university preparation course during the 2007-2008 school year. He taught in a small public secondary school (under 400 students) in

a rural area that drew students from a nearby town and a military base. Of the students in this school who were eligible to take the Ontario Secondary School Literacy Test (OSSLT) for the first time in 2008 (n = 94), 77% passed, which was below the provincial pass rate of 84%. While statistics were not available for the entire school population, the OSSLT data indicated that almost all of the students spoke English at home (91%). None were identified as English language learners, but 17% were identified as having special learning needs. Many of the students had parents who were engaged in active military duty, and the school community was often impacted by absences, injuries and fatalities among family members and friends. As such, the school staff were as concerned about the students' psychological well being as their academic achievement. Despite this demand, Kevin enjoyed his teaching assignment because the school was located in a "nice little community." He described himself as "lucky" because his classes were relatively small and student movement (due to military postings) occurred mainly during the summer months. Kevin was comfortable knowing all of the staff and most of the students in the school, and he was "happy" teaching in an environment where "everybody gets along."

*Teaching Experience and Education.* Kevin had been teaching for 13 years. He first worked as an occasional teacher and then taught in an program for young adults in Alberta for two years before moving to Ontario in 1997. Most of his teaching experience was in English at the secondary level, but he also had some experience teaching Drama, Math and Parenting courses. However, Kevin preferred teaching English, and he considered himself a specialist based on his teaching experience and education.

Kevin graduated from a university in Alberta with a Bachelor of Fine Arts (Drama)

and a Bachelor of Education in 1997. He continued his professional development at a university in Ontario to become certified as an Honour Specialist in Dramatic Arts (2003), and he acquired Additional Basic Qualifications in Senior English (2002). Kevin recently attended several professional development workshops, including one organized by the Council of Ontario Drama and Dance Teachers on using performing arts to improve students' moral reasoning and the school climate, and another by the Ontario Secondary School Teachers Federation that dealt more generally with secondary school pedagogy. Kevin had also engaged in several professional development opportunities related to assessment. He had recently attended a workshop on classroom assessment where he found that the ideas and strategies presented were "nice" but "not always practical." Kevin also participated in several events relating to the development of the OSSLT, and he particularly enjoyed sharing teaching strategies with other secondary English teachers one summer as part of the test development team.

*Views about Teaching.* In discussing his views about teaching, Kevin emphasized the importance of structure and routine. He justified many of the strategies that he uses on the basis that "it just makes it easier" and that "it's not going to come back and get you in the end." Kevin explained that this was because he was likely, as a teacher in a small school, to have students in his Grade 12 class that he had taught in earlier grades. It seems, though, that Kevin intended to serve his students interests as well as his own with this approach. He noted that structure and routine "help them as they go along too." Kevin found that many of his students "don't have anybody to . . . really look after them . . . and teach them right from wrong." As a result, he felt that it was important to be a role model for his students, to "give

them a little bit of a moral backbone,” and show them how to regulate their lives with good work habits. This was an aspect of teaching that had resulted in some frustration for Kevin, and he described feeling like his hands were “tied” in situations where ministry policies clashed with his beliefs, or when parents were not as supportive as he felt they could have been. Kevin related that when he became a parent he realized that his students were “somebody’s kids,” and he aimed to treat them the way he wanted his child to be treated at school. Overall, Kevin seemed comfortable with his teaching and assessment practices. He related that “one of the best things about teaching” is that students “always remember you . . . whether they liked you or hated you . . . enjoyed your class or despised it.”

*Ethics of Classroom Assessment.* Nine items were included on the Background Questionnaire for the participants to indicate their position on some of the ethics of classroom assessment. Of these items, Kevin initially identified one as ethical, three as dependent on circumstances, and five as unethical (see Table 4).

The only item that Kevin felt was clearly ethical was about a teacher creating learning activities similar to the tasks in the provincial assessment (SIMILAR). Kevin offered comments on all three of the items that he thought depended on circumstances. He felt that adding surprise items to a test (SURPRISE) would be ethical “as long as it’s material that has been covered in class.” Regarding the calculation of homework in final grades (HOMEWORK), Kevin pointed out that homework “could actually be a better indicator than a test” for some students, assuming it was collected and evaluated. On the subject of report card comments (POSITIVE), Kevin suggested that there were other means of communicating with parents, but he stressed that “if a teacher can offer comments or suggestions to help a student improve,

Table 4

*Participants' Initial Responses to Items on Classroom Assessment Ethics*

Items ▾	Participants					
	Kevin	Lorena	York	Tanar	Lucy	Amada
Surprise	Depends <sup>a</sup>	Depends <sup>b</sup>	Unethical	Unethical	Depends <sup>b</sup>	Unethical <sup>b</sup>
Change	Unethical <sup>b</sup>	Ethical	Ethical	Depends	Depends <sup>b</sup>	Unethical <sup>b</sup>
Homework	Depends <sup>a</sup>	Depends <sup>b</sup>	Unethical	Depends <sup>b</sup>	Ethical	Unethical
Zero	Unethical	Unethical	Unethical	Unethical	Unethical	Unethical
Guessing	Unethical	Depends <sup>b</sup>	Depends <sup>b</sup>	Unethical	Depends	Depends <sup>b</sup>
Similar	Ethical	Ethical	Ethical	Depends <sup>b</sup>	Ethical	Ethical
Answer	Unethical <sup>a b</sup>	Ethical	Ethical	Depends	Unethical	Unethical
Identity	Unethical <sup>a b</sup>	Ethical	Ethical	Ethical <sup>a b</sup>	Ethical	Depends <sup>a</sup>
Positive	Depends <sup>a</sup>	Depends <sup>b</sup>	Depends <sup>a</sup>	Ethical <sup>b</sup>	Ethical	Unethical <sup>b</sup>

<sup>a</sup> Additional comment provided on questionnaire.

<sup>b</sup> Clarification given during follow-up interview.

then he or she should do so.”

Kevin initially identified the remaining five items as unethical. During the follow-up interview, he described circumstances that might change the ethics of one these items. Kevin felt that because of his school board’s use of “peg marks,” a certain degree of flexibility was acceptable for senior students’ grades (CHANGE), especially if they were applying for post-secondary education. Kevin explained that peg marks are a set selection of percentages (e.g., 78%, 81%, 86%) that teachers must choose from when writing report cards, and he noted that a small increase in a final grade could make “all the difference in the world” for a student. Kevin concluded that bumping a student’s grade up one peg would not only be justified in terms of the consequences for the student, but also if it was a better representation of a student’s “most recent work.”

Kevin discussed two other items that he had initially identified as unethical, but without essentially changing his position. Although he noted that showing a student how to use a test response form (ANSWER) was “not necessarily a bad thing to do,” he felt that it was better to follow the instructions closely for administering large-scale tests, especially as he had heard of teachers “getting into trouble for deviating from it.” Kevin also felt that knowing students’ identities when grading essays was not in itself unethical (IDENTITY), noting that “you have a pretty good idea who it is anyway.” He suggested, though, that by masking students’ names, he could keep his assessment process “a little fairer.” He described several strategies, such as using the computer lab for essay tests and putting names on the last page of an assignment, to reduce the likelihood and perception of teacher bias. These nine items suggest that Kevin’s experience developing and administering a

standardized test informed his position on the ethics of classroom assessment, but he also felt that the fairness of an assessment could change depending on the circumstances.

### *Recommendations for Fair Classroom Assessment*

Kevin shared his practical wisdom through the recommendations he gave for fair assessment. Of the multiple aspects of fairness that he discussed, Kevin stated that he felt most strongly about four, which essentially involved sharing assessment criteria, offering appropriate tasks, ensuring that interpretations of student learning are well-informed, and maintaining consistent expectations and procedures.

Kevin emphasized that for an assessment to be fair, students must have a clear understanding of its purpose and the learning expectations. He believed that teachers should be “up front” about what they expect “right from the very beginning” so that the students “know you’re going to be fair, you’re going to treat them well.” Kevin recommended that students and parents be given learning expectations and assessment dates on the first day of a course, and that the criteria for an assessment task (rubric) be given at the outset of an assignment because it tells students “where we’re going . . . and it helps them get there.” Kevin’s comments indicated that his understanding of fairness included transparency, not only in the learning expectations, assessment criteria and procedures for a course overall, but also for each of the course assignments.

A second aspect of fairness that Kevin discussed related to the appropriateness of the assessment for the students. He believed that to be fair, assessment tasks had to be engaging as well as clear.

I think also appropriate as well, as far as something they can see, kind of like

why they're doing it. It's not just some assignment that I've sat up late saying, "Hey, what am I going to do with these guys . . . hey let's throw them, let's have them do this" . . . . so I think if it's something that they can see it's relevant, or appropriate . . . . there's some purpose behind it, and it's not just, you know, an assignment for the sake of having an assignment.

He suggested that this could be accomplished by varying the types of assessment tasks (e.g., writing a short story rather than unit test) and by offering a choice of tasks for students to demonstrate their learning. Kevin recommended that teachers:

Try to offer different activities . . . with each unit. So for example, when we do short stories, you're going to have to write one. Ah, maybe we'll do a novel, and you'll have to do . . . some type of other assignment with it, maybe like a test or something . . . . we'll get into plays, and maybe the big thing there is you'll have to reenact a scene, or maybe take a scene and modernize it. So, I usually try and stay away from just . . . one type of assessment . . . . I usually try and offer them, just vary stuff as we go . . . sometimes what I'll do is I'll even make a list of projects for a particular unit, and you choose which one you want to do.

He believed that variety and choice made assessment fairer because it gave students "more than one chance or opportunity to be successful" and he emphasized that when assessments "hit different interests . . . different talents, it is fair because. . .it goes back to Bloom's Taxonomy, not everybody is or does the same thing." In this light, Kevin saw variety not only as a means of improving the appropriateness of an assessment for students, but also as a way of enhancing his understanding of student learning. In essence, Kevin recommended using varied approaches to elicit learning, which is one of the core practices of assessment for learning, to ensure that the teacher's interpretation is well-informed. Kevin added, though, that for this type of assessment to be fair with a large class, a teacher would need to be organized and "stay on top" of students so they don't get "lost in the shuffle."

Kevin believed that treating students consistently was one of the most important

aspects of fairness. The issue of consistency came up repeatedly in his responses, especially in relation to students completing assignments. For example, in his response to the vignette about missing work, Kevin wrote:

I think that it is necessary for a student to hand in all assignments, no matter how small or unpopular the assignment may be. I will call home or give detentions until such time as the outstanding assignment is in. . . I want to try my best to be fair with all of the students in my classes. By letting one student off of the hook for an assignment or letting him or her hand it in late without any consequences I'm not being fair.

Asked how he maintains consistency when he offers students choices in classroom assessment, Kevin explained that the same criteria can still be used to assess student learning, regardless of the response format (e.g., poem, short story, personal reflection on a reading). Different response formats give students “a chance to build on something [they’re] . . . actually good at and show it off” while still “demonstrating understanding” of the learning expectations. Kevin stressed that this type of differentiation in assessment is fair because “everybody’s kind of jumping through the same hoops, even though it’s a little different . . . everybody in the end is still doing the same thing.”

Kevin gave two additional examples that illuminated his understanding of fairness further in terms of consistency and differentiation. In one he explained that he prefers not to “force” students to do oral presentations because he believes that it is necessary to be “flexible” for students who are “really uncomfortable” in front of the class. He noted, however, that “by the same token it’s not fair” if a student does not try at all. For this kind of circumstance, Kevin recommended that the assessment be adjusted, but he cautioned that “to keep it fair, you really do have to keep to the structure” of the original task. A modification he

recommended was to have the student present to a smaller group of peers. Kevin's second example involved a situation where he felt that the outcome was completely unfair. A student was absent for over half of the morning classes one semester, but was permitted to complete the course work at home because she had a note from a psychiatrist. Kevin stressed that while there were "legitimate" reasons for extended absences, he felt that this student was "boycotting mornings." Kevin disagreed with the principal's decision to give the student "the pass" because it wasn't "really fair" when other students had "put in" their class hours. The fairness issue in this circumstance was not related to the students' academic achievement, but rather to the lack of value placed on developing steady work habits, which Kevin believed were absolutely necessary for students to become "good people."

#### *Elements of Practical Wisdom*

In making recommendations for fair classroom assessment, Kevin seemed to draw to some degree on all the elements of practical wisdom. He frequently referred to his own teaching experience in expressing general knowledge about learners (i.e. theoretical knowledge), and he used examples of particular students and situations to explain his responses. His recommendations were imbued with pedagogical knowledge (i.e. technical knowledge) that often involved tools or methods to help structure learning for students (e.g., checklist for peer-editing, steps for essay writing). Kevin expressed an openness to other teachers' strategies, and he related that when he heard a "neat idea" he'd try it to see if it "actually works" in his classroom.

The relationship between generally sound pedagogical practices and fair assessment was sometimes blurry in Kevin's responses. For example, when I asked him to elaborate on

why he felt structure made assessment fairer, he was not initially able to articulate his thoughts, replying instead “mmm . . . I know what I want to say, I’m just having a problem saying it.” However, he was able to talk through this, and eventually he described the tension between consistency, which centers on fairness for all, and flexibility, which addresses the need to be fair to individual students. When his recommendations focused explicitly on fairness, Kevin drew on his moral beliefs and professional ethics more than any of the other elements of practical wisdom. For example, he responded to one vignette by writing:

I am a very firm believer in the peer-editing process and its value. I would speak to the class as a whole about the difference between constructive and destructive criticism. . . I have a very strict policy about what is okay and not okay to say in my classes. Students know this. By voicing an opinion or criticizing someone's work for the amusement of others is not an appropriate or acceptable behaviour.

During the interview, Kevin provided a pedagogical rationale for this, stating that some students “close up” with criticism from peers, and then the teacher has to “try like anything to get them out of their shell to try again.” He also felt this was a “a moral belief” because it “falls under criticizing kids as well. . . as far as bullying kids, or making fun of kids . . . so you nip it in the bud.” Kevin emphasized the need for educators to consider “what we are really teaching these kids” and this imperative seemed to be central in his practical wisdom.

### *External Factors*

Kevin identified a wide range of factors that could influence the fairness of classroom assessment. At the classroom level, Kevin pointed out the benefit of having small class sizes for both planned and interactive assessments. He noted, for example, that it is easier with smaller numbers to “keep track” of student learning, and to write comments or suggestions

on assignments. He also felt that smaller classes afforded more “freedom” during classroom interaction because there’s more “one-on-one time” for feedback, and the teacher can “go around and actually help kids” as opposed to keeping a large group “inline.”

The majority of Kevin’s comments on factors that might influence fairness related to student behaviour and attitude. When asked directly if assessment was fair when it was differentiated for students, Kevin replied, “I think before I differentiate something, or just change something, or even tweak something for somebody, I kind of want to see some type of effort there.” Taken in conjunction with Kevin’s earlier recommendation regarding choice and variety in assessment, and his emphasis on the importance of consistency, this comment suggests that while he was not opposed to flexibility, he believed that student behaviour dictated the degree to which it was warranted. In his recommendations for several vignettes, Kevin also noted that the behaviour of some students could have a negative impact on how much feedback or opportunity to learn the teacher might be able to provide. He discussed this issue further during the interview, indicating that he sees a reciprocal relationship between fair assessment and classroom management. He explained that “if you’re fair in they way you mark it also makes it fairer . . . a little bit easier as the classroom discipline stuff goes because everybody just . . . falls more in line.”

Beyond the classroom, Kevin identified system policies and school leadership as influences that could work in tandem on the fairness of an assessment. He explained that even though teachers are no longer “allowed to deduct marks for lateness,” they still need “some way to encourage students to get stuff in on time.” He believed that it helped to “penalize kids with detentions or longer assignments,” but he noted that he could not do this

without his administrations' support. "If they back me up, my job is far easier." Although the OME policy on late and missed assignments aimed to improve the fairness of classroom assessment by having teachers report on students' academic achievement and learning skills separately (see *Growing Success*, 2008, pp.6-i-6-iv), Kevin felt that it hindered fairness because the lack of consequences for late or incomplete work meant that the assessment process had become inconsistent from student to student in his school.

### *Summary of Kevin's Case*

Kevin's approach to teaching seemed to fit well with his school environment. He emphasized the need for structure, routine, and discipline, but at the same time, many of his comments suggest that he tended to be flexible, caring and constructive in his interactions with individual students. Although he discussed multiple aspects of fairness, he placed most emphasis on four: the use of transparent learning expectations and assessment criteria, appropriate assessment tasks to meaningfully engage students, multiple and varied assessments tasks to inform his interpretations of student learning, and maintaining consistency in the treatment of students. Kevin's recommendations for fair assessment tended to focus on the pedagogical knowledge he had developed through his teaching, but they were also strongly guided by moral beliefs and professional ethics. Kevin wanted to be a role model for his students, and he aimed to be fair to all students in his decisions and interactions with them. His beliefs about the importance of good work habits surfaced in how he considered student effort before modifying classroom assessments, and also in his disagreement with a policy that forced him to pass a student who met the course expectations without attending class regularly. Although Kevin identified many external factors that could

impact the fairness of an assessment, he repeatedly referred to students, and their behaviour seemed to play a strong role in the degree to which he used assessment to support learning.

### Lorena: Who the Learner Is

#### *Lorena's Portrait*

The quotations given in Lorena's case portrait are from her responses to the background questionnaire (19 May 2008), the vignette questionnaire (25 May 2008), and her interview transcript (27 May 2008).

*Teaching Assignment.* During the 2007-2008 school year, Lorena was a special education resource teacher, the teacher-librarian, and the literacy lead teacher in a relatively small public middle school (under 400 students) in the center of a large urban area. As a resource teacher, Lorena taught Grade 7 and 8 special education students in pull-out groups that concentrated on improving skills in English and Mathematics. As the teacher-librarian, she worked with classroom teachers to support student learning. For example, she introduced novels for literature circles and taught mini-lessons on questioning skills. As the literacy leader, Lorena worked as part of a team that was responsible for communicating strategies to improve literacy within a family of schools (i.e., elementary, middle and secondary schools in geographic proximity within a school board).

In addition to the regular Grade 7 and 8 program, Lorena's school offered special education classes for students with a variety of disabilities, including learning, behavioural, developmental, and physical disabilities. Although the school was located in a culturally diverse area, many of the students spoke English as their primary language (approximately 60%). Because the EQAO does not administer tests in Grades 7 and 8, there are no large-

scale results to help describe the students in Lorena's school. The two secondary schools that serve the area are quite different in that one is technically oriented with a low pass rate for the OSSLT (58%), and the is other more academically oriented with a higher pass rate (81%) that is close to the provincial rate (84%). This information, taken with a comment by Lorena about student achievement in one class spanning four grade levels, suggests that the range of academic ability was quite wide in her school's student population.

*Experience and Education.* Lorena had been working in urban schools for 14 years. She was an educational assistant for two years before she became an elementary teacher, and she had been teaching at the intermediate level for seven years. In her previous assignment as a Grade 7 home room teacher, Lorena taught Mathematics, English and Social Studies, and she remembered in particular that she enjoyed integrating the English and Social Studies curriculum. Lorena indicated that she preferred teaching English Language Arts more than other subjects, and she considered herself a specialist based on her experience and education.

Lorena graduated from a university in Ontario in 1995 with a Bachelor of Arts (Drama) and a Bachelor of Education. Beyond her basic qualifications for teaching in the primary, junior, and intermediate divisions, Lorena had engaged extensively in professional development through Additional Qualifications courses. She was certified as a Special Education Specialist (1997-1999) and a Dramatic Arts Specialist (1995-2000). She had also taken Additional Qualifications in Reading (Part 1, 2001), English as a Second Language (Part 1, 2002), Librarianship (Part 1, 2005; Part 2, 2006), and the first two parts of the Principal's Qualifications Program (2006). Looking over her qualifications, Lorena noted that in addition to these university-based courses she had also been involved in a "good deal

of professional development” in her current role as the literacy lead teacher, and previously as a literacy coordinator. However, she also stressed that she hadn’t sought these leadership roles. She explained that she had been selected by school administrators, and that her interest in leadership had consequently “come about” and “grown” over the years.

*Views about Teaching.* Lorena expressed her views about teaching openly and clearly, and she asked many questions during our interactions. For example, one of her initial written responses contained a statement followed by a question: “I believe that the teacher’s role is to facilitate and coach student achievement. Help is NOT a 4 letter word. I wonder if this is because I’m an elementary teacher?” Lorena believed that it was important to encourage Grade 7 and 8 special education students to advocate for themselves in preparation for their transition to the secondary school environment. However, she also indicated that as a resource teacher she accepted considerable responsibility for her students’ learning in a way that she related to parenting:

Well if a child is not getting concepts, if they’re not being successful . . . .I need to find out why. . . . and being the Italian Catholic mother that I am, I always look at myself first. What did I do that he doesn’t know how to do this? . . . . so I mean, I look at my practice.

Lorena was critical, not just in reflecting on her own teaching practice, but in thinking about issues in Ontario’s educational system. For example, she described a “transition meeting” with secondary school educators, and after noting differences in the delivery of special education at that level, she concluded that “we’ve really got to do something about creating this continuum . . . otherwise . . . the children aren’t going to have any consistency.” Lorena expressed considerable interest in educational research. She posed several questions about

this research project, and she described an action research project in which she was currently engaged. She and her colleagues were sponsored by the Ontario Elementary Teachers' Federation to study the process of teacher moderation in assessment as they implemented it in her school. In keeping with her extensive and ongoing professional development, Lorena felt that was important for teachers to "be aware of trends and changes" in education. Lorena seemed to feel that learning was a key aspect of teaching. For example, she joked that teachers should avoid having their "day plans laminated," and she emphasized that teachers "can't be stagnate" in their practices.

*Ethics of Classroom Assessment.* Of the nine items on the ethics of classroom assessment included in the Background Questionnaire, Lorena identified four as ethical, another four as dependent on circumstances, and one as unethical (see Table 4).

The items that Lorena indicated were ethical were about teachers changing a student's final grade based on the work completed (CHANGE), creating learning activities similar to tasks on the provincial assessment (SIMILAR), helping a student use a test response form accurately (ANSWER), and grading essays tests where the student is identified (IDENTITY). The only item that Lorena felt was clearly unethical was about a student receiving a zero for not returning a form with a parent's signature (ZERO).

During the follow-up interview, Lorena elaborated on all four of the items that she had identified as dependent on circumstances. She thought that it would be ethical to include a surprise item on a test (SURPRISE) if it paralleled information discussed in class. The purpose of this would be to evaluate students' application of knowledge, which is one of the four categories in the OME achievement chart. Lorena was less certain about weighing

homework heavily in final grades (HOMEWORK), stating that she was “torn about the whole homework issue.” She felt that often with homework students made “further mistakes and ground themselves in those mistakes.” As a result, Lorena didn’t see much merit in calculating homework itself into students’ grades, but she did believe that students’ homework habits (i.e., finishing work or preparing for class) should be considered in the assessment process. For the item about penalizing guessing (GUESSING), Lorena remembered having a secondary school teacher who scored tests this way, and she thought it “kind of made sense at the time.” Although she felt that this practice might be okay for “structured multiple-choice where you’ve got it all set up as such,” she clearly did not design assessments this way in her own practice. For the final item about writing only positive comments on reports cards (POSITIVE), Lorena’s response indicated that she wasn’t entirely comfortable with this practice either. She began with “I know that we do that, but I don’t know that it’s always in the best interest of the child.” She went on to explain that while it would be unethical to provide inaccurate information about a student’s learning, teachers did sometimes avoid negatives and use “teacherese” (e.g., beginning to, with assistance) to signal lower achievement. Lorena suggested that with the limited space for comments on report cards, “talking to the parents” was the best way to communicate about a student’s difficulties. Lorena’s responses to these nine items suggest that she considered the purpose of the assessment in determining whether a practice was ethical or unethical. While she seemed to identify some items as dependent on circumstances because they weren’t within her experience (e.g., test development), several other items generated discussion about common practices in elementary education, such as marking homework and writing report card

comments, that she had already questioned herself.

### *Recommendations for Fair Classroom Assessment*

Lorena shared her practical wisdom about fairness through her recommendations for writing assessment. Of the multiple aspects of fairness that she discussed, Lorena stated that she felt most strongly about four, which involved using assessments that are appropriate for individual learners, ensuring that interpretations of student learning are well-informed, reflecting on classroom assessment with colleagues, and increasing consistency across teachers' assessment decisions.

Ensuring that instruction and assessment are appropriate for individual learners was an important aspect of fair assessment for Lorena. She associated fairness with differentiation because, as she explained, some students need to be challenged while others need to be given tools to support their learning. She gave an example of a "child who has really great ideas, but physically he has difficulty writing" and said, "well heck, let's give him a computer . . . to me that's being fair." To differentiate appropriately, though, Lorena argued that it would be absolutely necessary to understand the individual learner's needs:

Knowing who the student is, is being fair. It's like being a parent, and you've got three kids, and one of them has worn out their shoes. Well, in some circles you get everybody a new pair of shoes, but no, it's only child number two who needs the new pair of shoes, and that's being fair.

The notion of understanding the student's needs leads to another aspect of fairness that Lorena emphasized. In order to differentiate assessment for learning, Lorena pointed out that a teacher would need information about the whole learner, the individual beyond the classroom walls.

If a child doesn't have paragraph structure in their writing . . . . I still have to see that, but I also have to understand how am I going to approach this lack of knowledge. Like, how am I going to help him with his paragraph writing. Is it because he broke his leg when he was in grade three and was away from school for two months. Is it because there's a second language at home? Is it because he's only got one parent who works three jobs? You know, there's all those things that have an impact on their learning . . . . and if we don't take that into consideration, then how can we be fair?

For Lorena, fair assessment requires information. She added a caveat in stating that she would not include this information as "part of the actual mark," but she believed that "all those things" should be taken "into consideration when assessing a student." Continuing with the example of paragraph writing, Lorena stressed that assessment results should be investigated in order to inform teaching and learning.

If a child is not getting concepts, if they're not being successful . . . . I need to find out why. . . . I can't just say, 'okay, they don't know how to write paragraphs.' I need to find out what in their background . . . . so I do an OSR search, I need to find out from previous teachers . . . other people that have worked with them. . . I look at my practice . . . and to be fair in my assessment, I need to do that. . . . rather than just giving them an R for the sake of marking the paper.

Although Lorena's focus on the whole child may have been a product of her experience as a special education resource teacher, seeking information from multiple sources, whether it be from the Ontario Student Record (OSR), other teachers' observations, or current assessment results, is consistent with the rationale underlying the use of varied approaches to elicit learning, which is one of the core practices of assessment for learning.

Lorena also identified consistency as being an important aspect of fairness. In response to one vignette, she felt that giving students the same amount of feedback was less of an issue than consistency in the content of the feedback. She explained that if a lesson had

focused on semi-colons, for example, then “everybody would get feedback on semi-colons.” Drawing on her action research on collaborative assessment, Lorena also discussed the benefits of consistency across classrooms. She related that in her school “those of us all teaching Grade 7 English are looking at the papers and saying, ‘okay this is the rubric, this is the criteria, what mark will we give this paper?’” And as the teachers progressed, they developed “common goals and common understandings,” which Lorena thought would “bring out more fairness and a more level playing field.”

Given her involvement in collaborative assessment, it was not surprising that Lorena identified teacher reflection as a key aspect of fairness in classroom assessment. She mentioned the benefits “speaking with colleagues” and holding “team meetings” several times, and when asked directly what teachers should do to assess student writing fairly, she replied: “I think a really important thing here . . . is involving your colleagues. Like, talking to your colleagues about how you’re coming up with these marks.” However, Lorena’s responses also suggest that she engaged in critical reflection independently. Ultimately she described her “approach to fairness” as being multifaceted, explaining that “part of it is who I am, part of it is just personality and core beliefs, and part of it is also research, and professional development, and speaking with colleagues, and understanding how a student would benefit . . . they sort of feed off each other.”

### *Elements of Practical Wisdom*

In describing her approach to fairness as an interconnection of elements, Lorena was essentially describing her practical wisdom. In many of her responses she linked different elements of practical wisdom. For example, she noted that her student-centered approach to

working with intermediate students was influenced by her earlier elementary teaching experience. She also explained that her planning was based on her understanding of a particular group of students, stating that it “depends on who’s in the class . . . I have a general idea – this is what the curriculum wants, but the way I’m going to do it, and what day I’m going to do what depends on the class.” In many of her recommendations, Lorena supported theoretical knowledge, which can be seen in statements such as “assessment is an ongoing process,” with considerable technical knowledge. For example, she described the “gap analysis” needed to use assessment information to adjust teaching, and she compared the diagnostic merits of different reading assessment tools. As such, Lorena’s practical wisdom about fair assessment was informed by different elements of practical wisdom, including the theoretical and technical knowledge she had gained through teaching experience and professional development.

Lorena’s sense of professional ethics was also strongly at play in her recommendations for fair assessment. She noted that assessment can be “really draining” emotionally because of the tension between building a student “up for success” and being “honest” to support long term learning. Many of Lorena’s responses to the vignettes, such as “I am not in a classroom to ambush or ‘catch’ students for a lack of knowledge” and “are there gaps for which I am responsible,” suggest that she sees a constructive orientation in teaching as a professional obligation. Her comments about needing to “find out why” a particular student was struggling, and her emphasis on the importance of ongoing professional development held the same tone. As she discussed her practical wisdom, Lorena recognized that the elements worked together to inform fair assessment, and she commented

that “a lot of it has to do with the moral and ethical knowledge, and the personal characteristics, and then that’s supported by professional development . . . it’s a blend.”

### *External Factors*

In discussing the influence of classroom factors, Lorena focused on students and their parents. She noted that parents are not always a positive influence at two different points during the interview. When she explained her concerns about including homework in the assessment process, she noted that parents’ help with homework varied because sometimes “they’re using different words than the teacher used, or they’re approaching it from a different point of view . . . and that’s not really helping.” Lorena also mentioned parental influence in relation to the formal identification of students as exceptional learners. She gave a particular example, saying “this one student I have in mind right now, I think . . . he’d probably rather not be identified. I think his parents like the idea of identifying him. There’s sort of an elitism that goes with that.” In her recommendations for fair classroom assessment, Lorena emphasized the importance of “communicating with the parent and having them understand, not only what you’re reporting on, but what you’re teaching the student . . . and having them at least be aware of . . . the spectrum that goes on in the grade.” Although she noted that it might not impact the fairness of an assessment directly, Lorena clearly felt that communication with parents was a key component of a constructive learning environment.

Lorena discussed student behaviour more than any other external factor. She believed that the “student teacher relationship is very important,” and she stressed that when students “become hostile or alienated” or “defensive,” not only is there an “unpleasant tension in the

classroom,” but “learning is interrupted.” Lorena’s comments on student behaviour, however, focused heavily on the teachers’ responsibility in creating a “classroom atmosphere” that permits fair assessment.

Understanding our students and their development helps us when dealing with certain situations. Sometimes it's personalities, other times maturation, or others still social circumstances that can influence behaviours . . . Fostering a supportive environment is paramount to a class community. Being proactive . . . will set up the classroom community to be respectful and everyone will be aware of the common expectations.

While this response was given specifically for the vignette about peer-assessment, Lorena referred repeatedly to the importance of a constructive and safe learning environment.

Lorena also identified the school environment as a potential influence on the fairness of classroom assessment, but most of her comments related to colleagues rather than school administrators. She referred to teacher friends, and underscored the importance of “dialogue amongst teachers.” As noted earlier, Lorena was engaged in a collaborative project that aimed to improve the quality of classroom assessment through discussion. In addition, she observed that reflecting on assessment, which she saw as a “big part” of fairness, could also be done less formally with a colleague.

Lorena mentioned multiple factors at the system level that could influence the fairness of classroom assessment, but the only one that was prominent in her responses was report cards. She had determined from her discussions with parents that they often placed more importance on grades than comments, and she noted that the standardized report card limits what can be communicated because “there’s only so many lines on the report card . . . the learning skills box is only so big, and the English box is only so big.” As a result she felt that the message conveyed to parents was not clear, and she emphasized this point by relating

that a friend of hers, who was a lawyer and adept at deciphering text, had difficulty understanding her child's report card. Lorena thought this was due to the standardized vocabulary that teachers in Ontario were supposed to use for their comments. "The way we've been instructed to report – I must comment on the expectation. But unless they have what we call an R, like under 50, we don't put 'does not have the expectation' . . . because they're working towards it." One of Lorena's colleagues had apparently explained that the words "always, sometimes, never" indicated the achievement chart levels, and "if the kid does it regularly you leave it blank." But as Lorena noted, "you can't write a whole report card with sometimes, sometimes, sometimes." Although report cards are associated with assessment of learning, Lorena's comments suggest that she saw them as a form of communication that could serve the fairness of assessment for learning. She concluded, though, that with the limited space for comments and the standardized vocabulary of report cards, teachers had to "find other ways of communicating" with students and parents in order to support student learning.

#### *Summary of Lorena's Case*

As a teacher, Lorena valued and engaged in ongoing learning and educational inquiry. She questioned the merit of practices that are common in her educational system, and she appreciated the opportunity to reflect critically with colleagues. Lorena discussed many aspects of fairness, but she placed the most emphasis on differentiating appropriately for individual learners, gathering information to understand student learning, reflecting with colleagues, and increasing the consistency of classroom assessment. In thinking about the fairness of classroom assessment, Lorena drew on all the elements of practical wisdom, and

she was cognizant of their inter-working. Her strong sense of professional ethics seemed to permeate all aspects of her practice, including her student-centered approach and her use of assessment to inform her teaching. While Lorena recognized the potential influence of a wide range of factors on the fairness of classroom assessment, she tended to focus on the role of teachers in creating a learning environment to support fair assessment.

### York: As Much of the Picture as Possible

#### *York's Portrait*

The quotations given in York's case portrait are from his responses to the background questionnaire (02 May 2008), the vignette questionnaire (09 May 2008), and his interview transcript (26 May 2008).

*Teaching Assignment.* York taught three Grade 11 and 12 English college and university preparation courses, and one Grade 12 Geography course during the 2007-2008 school year. He taught in a very large public school for young adults (over 2000 students) that offered multiple programs, including cooperative, apprenticeship, special education, English literacy, and general academic credit programs. The school was located in an area that had been subsumed by urban sprawl and was an inner-ring suburb crisscrossed by highways. York worked closely with a small group of like-minded colleagues in a program that focussed on supporting students between 18 and 21 years old who had experienced difficulty in regular secondary schools. In York's words, these were students who fell "between the cracks for all different reasons: personal, emotional, physical, academic." The school year in this program was divided into four intensive ten-week sessions. Students took one course at a time in small classes to earn credits toward an Ontario Secondary School

Diploma (OSSD). York described the academic achievement of the student population as a “mixed bag” overall, with a few students aiming for college or university, while others were fulfilling parole requirements. Most, he believed, were “aspiring for a 50 . . . to get their OSSD.” The 2008 OSSLT results for this school were not released “to protect against disclosure of personal information associated with individual students” (EQAO, 2008a). Given the small class sizes and the nature of the programs at this school, it is likely that very few students attempted the OSSLT.

*Experience and Education.* York had been teaching for 18 years. He taught Grades 7 and 8 for 15 years before moving to his current assignment three years ago. York explained that he was hired for this specialized program because of his experience in the intermediate division, which has generally been more student-centred than secondary academic programs. In addition to English, York had previously taught Math, Science, Computers, Media, Geography, Drama and Social Studies courses. Based on his experience and education, York considered himself a specialist in English Language Arts. However, he had come to prefer teaching subjects that were reflective of the “real world outside the school,” and he described a particular course on Media and Computers that he had recently enjoyed teaching. He explained that this was because when he and his colleagues were hired for their program they “basically had to figure out what courses [would] keep these kids in school,” and they had found that this course was more popular than the regular English curriculum.

York completed his elementary and secondary education in England before moving to Canada. He graduated from a university in Ontario with a Bachelor of Arts and a Bachelor of Education (1989). His initial qualifications were in junior and intermediate English, and he

had since taken additional qualification courses in Computers (Part 1, 1990), Special Education (Part 1, 1995; Part 2, 1999), and Librarianship (Part 1, 2003). He was certified in Senior English (2005) and as an Honours Specialist in English (2006). York frequently attended workshops sponsored by his school board, and he was very familiar with assessment documents produced by the OME and his school board. He had been involved in multiple scoring sessions with the EQAO, and he had scored students' responses for the OSSLT, as well as Grade 3 and 6 assessments in Math and English.

*Views about Teaching.* In discussing his teaching practice, York emphasized that his school was unlike most secondary schools. He described it as “kind of like a train station, you get what you get . . . they come when they come and they go when they go.” What was “very pointed” in this program, he explained, was that he and his colleagues were “not teaching curriculum, we’re teaching students.” While he felt that his elementary experience had helped with this “kind of angle,” he also noted that the issues he faced as a secondary teacher in this program were very different. For example, in response to one of the vignettes, he raised the issue of control.

I think for high school teachers control is a big issue . . . coming from elementary where we worried about things like wearing hats and chewing gum to a place where some of these students have come out of jail, or where they’re going home to an abusive partner . . . issues of control take on a whole new meaning.

York noted that “too much control” and rigidity in classroom practices could be problematic, and the flexibility he had in working with a small number of students seemed to be the aspect of his teaching assignment that he enjoyed the most. He remembered that in his secondary school in England, some of the teachers treated the students like “pond scum,” and he was

quite clear about not wanting to follow their example. Overall, though, York felt that his own educational experiences had been very positive, and he seemed to carry this over into his teaching practice. While he was aware of many challenges, and he was open to critical reflection about different aspects of classroom assessment, York's focus as a teacher was primarily on the best interests of his students.

*Ethics of Classroom Assessment.* Of the nine items on the ethics of classroom assessment that were included in the Background Questionnaire, York identified four as ethical, two as depending on circumstances, and three as unethical (see Table 4). The four items that York indicated were ethical were about teachers changing a student's final grade based on the work completed (CHANGE), creating learning activities similar to tasks on the provincial assessment (SIMILAR), helping a student use a test response form accurately (ANSWER), and grading essays tests where students' names were evident (IDENTITY). Except for the latter, which involves the issue of teacher bias, the items that York identified as ethical are practices that would improve the accuracy of an assessment because they avoid assessing irrelevant skills or behaviour (e.g., assessment of ability to use a form rather than relevant knowledge). The three items that York thought were clearly unethical were about including surprise items on an exam (SURPRISE), weighing homework heavily in final grades (HOMEWORK), and giving a zero to a student for not returning a form with a parent's signature (ZERO). All three of these items introduce elements into the assessment that could be irrelevant for the assessment.

York commented further on the two items that he had identified as dependent on circumstances. During the follow-up interview, York corrected his initial response to the

item about a teacher deducting points for wrong answers to discourage guessing on a test (GUESSING). York felt that there were no circumstances in which this practice would be ethical, and he stated emphatically that he “just wouldn’t do that.” For the final item about writing only positive comments on reports cards (POSITIVE), York pointed out that they should not always focus on strengths because “sometimes you need to point out areas for growth and improvement.” These nine items suggest that York’s orientation in classroom assessment was constructive in that he focussed on what would be best for student learning in identifying assessment practices as either ethical or unethical.

#### *Recommendations for Fair Classroom Assessment*

York shared his practical wisdom through the recommendations he gave for fair writing assessment. He discussed different aspects of fairness, and he offered many strategies in response to the vignettes. For example, he suggested involving students in assessment to make it fairer, and he described one of his own practices.

So often what I do, and I certainly don’t do this always, but often – you create the rubric with the class . . . you pick the criteria. If it’s an English essay, it’s fairly standard. But if it’s an assignment, like a project or scrapbook or something . . . we’ll pick the criteria together. I usually guide them for the things that I, the ‘must do’. But often the focus might be on . . . creative interpretation, or even in some cases, it’s just completion.

York believed that students understand the criteria better when they are involved in its determination, but he emphasized the need for transparency even with teacher-made rubrics, recommending that “you always share the rubric with them upfront, so that . . . they know what the rules of the game are.” In several of his written responses, York questioned the purpose of the assessment, asking “What are the marks for?” and “What were the assessment criteria for the assignment?” York also gave clear instructions for the teacher in each

vignette, with directions such as “talk with the offending student privately” or “accept the work as is” and “review the rubric.” During the follow-up interview, York indicated that felt most strongly about two aspects of fairness, which related essentially to the information used to interpret student learning, and to teachers’ critical reflection in the assessment process.

York emphasized several times that for classroom assessment to be fair, teachers should give students “a chance to show . . . what they know.” He indicated that, ideally, the process should involve more than one chance, explaining that in class “we cover . . . these three or four expectations, and give you an assignment, and you are assessed, but I have to give you another opportunity to demonstrate that you've achieved those expectations.” York also believed, though, that it was not necessary for a student to take up every opportunity offered to achieve the learning expectations.

If you show enough potential . . . how many short stories do you need to be able to analyse. Well, we're supposed to do one a week, so that's 10 short stories . . . but if you've analysed five really well. If I take a look at your mark, well then, you've got five zeros, but I know you know how to do it . . . . . you've met the criteria.

This comment suggests that York believed assessment decisions should be criterion-referenced, which is consistent with his earlier response (to one of the items on the ethics of classroom assessment) where he gave more weight to mastering learning expectations than completing work. So while offering multiple opportunities to demonstrate learning was part of York’s recommendations for fair writing assessment, he seemed to see the need for a student to use those opportunities only in relation to the degree to which they contributed to the teachers’ ability to develop an understanding of the student and the student’s learning.

I guess you'd have to really try and put yourself in the shoes of the student . . . and just try – you don't even have to agree, you have to understand. And

you've been in that situation as a learner yourself. And it's not about cutting someone a break, or cutting them some slack, it's about understanding as much of the picture as possible so that when you make a decision it's based on as many facts as you can.

Here he indicated that what is most important for decision making is the completeness of a teachers' understanding. However, many of York's comments also suggested that he considers, first and foremost, the consequences for an assessment for a student's future. Discussing grading and report cards, for example, York said that his reporting "really depends on what the student needs . . . I usually err in saying what I think is best for the student . . . both in the short term and in the long run."

For classroom assessment to go beyond record keeping, as York recommended with his emphasis on understanding, teachers' assessments must not only be well-informed, but also subject to reflection. York openly discussed the subjectivity of classroom assessment interactions, noting that in "dealing with human beings, you intuit a lot . . . that can be bias, there is always that risk . . . my intuition of you may be based a lot on my own personal experiences . . . so I might be completely wrong." York indicated that he did not view this as an insurmountable problem, however, and he offered several strategies to help teachers avoid biased assessments, including listening to colleagues, questioning assumptions, and documenting classroom interactions. Based on his experience, York felt that many teachers are quite competent in this area.

A lot of teachers are actually very good at getting to know where the student is coming from and who they are as a person. Of course you have to still . . . balance that with being professionally careful in your judgment – some students think just because they're nice to teachers, they should get a higher mark. Well what if you and I hated each other, should I give you a mark based on how I feel about you? Well no, that's not fair.

Overall, in the recommendations that York gave for fair assessment, fairness was a quality of classroom interactions that demanded understanding, empathy and reflection from teachers.

*Elements of Practical Wisdom*

The combination of professional and personal stances that York alluded to in discussing fair classroom assessment is reflected in how he expressed his practical wisdom. For example, York recommended that the teacher in one vignette get the “point across without embarrassing anyone,” and he explained that he responded this way because it was how he “would want to be treated.” Essentially, York espoused the golden rule as a guide for classroom assessment interaction. Although he gave voice to relatively few moral beliefs, it was evident that his own experiences as a student did inform his recommendations. York explained, for example, that his association of fair assessment with well-informed decision making came from his “personal experiences of what it’s like to be . . . judged on an incomplete picture, the teacher not really knowing what’s going on.” York also commented on the difficulty of teasing his experience as a learner apart from the other elements of his practical wisdom, saying that it was “hard to pinpoint particular experiences” because “it’s very . . . cumulative . . . I’ve basically been schooled for 40 years, so some of that kind of seeps out in other ways.”

The practical wisdom that York shared suggested that he had acquired pedagogical knowledge from multiple sources, including knowledge of students and knowledge from students. For example, he said that one of his responses was based on knowing “what works for students,” whereas he indicated that another was informed by “feedback from students.” York expressed considerable knowledge about classroom assessment, which seemed to

derive from both teaching experience and professional learning. In discussing the merits of written feedback, for example, he justified his preference for oral feedback based on his teaching experience.

I tend not to write a lot of feedback . . . it's much better to talk it, to say it, and explain it . . . in this program I actually have the luxury to give oral feedback. In elementary we did what was called conferencing where you sit down . . . but in high school most of the feedback is in writing . . . and it mostly – it does get read, but when you talk about it, it is much easier to integrate.

In contrast to this experiential basis for technical knowledge, when York discussed assessment issues specific to Ontario's educational context (e.g., 70%/30% split for final grades, the weight of achievement chart categories), he referred to his knowledge of his school board's assessment policy. He described this document in considerable detail.

It's a fleshed out version of the what the ministry policies are on assessment and evaluation and reporting. And it's actually, it's an excellent document, and it is very, very research-based. It's very student-oriented, even the high school version . . . works in the best interest of the students.

York's familiarity with school board and ministry assessment documents was evident in the knowledge he expressed about classroom assessment, and it played a strong role in the practical wisdom he shared about fairness in classroom assessment

### *External Factors*

York identified several factors that could influence the fairness of classroom assessment at the system level. Assessment policies clearly influenced his thinking in this area. In addition, he discussed time in relation to student numbers and curricular demands. He recognized that with the small classes in his program he had more time to give students individualized oral feedback on their writing, but he also noted that a lack of time sometimes impacted his choice of assessment methods. For example, he explained that involving

students in assessment (e.g., developing rubrics with students) requires time, and “sometimes it’s just faster” to tell students “okay, this is what we’re going to do.” York noted that even in his program, where the “focus is more on the student than the curriculum” the number of learning expectations that have to be addressed within the time frame of a course can impact the fairness of classroom assessment.

Sometimes there’s just not enough time to do another assignment based upon the same kind of learning . . . because I can’t just drop something totally new on you . . . and then have you demonstrate your proficiency in these expectations. So sometimes it’s a one-shot deal. We cover it . . . I think we used to do that a lot more with the curriculum change in ‘97, but sometimes there’s just not enough time to give you more opportunities to demonstrate.

Although he clearly valued student feedback, York related that with time pressures, his communication with students sometimes became “very . . . goal oriented . . . like, you need to hand this in by such and such a date, we have to get this done, I have to mark it, and report on it.” He also commented on a “disconnect” in current policies in relation to time, specifically because teachers are mandated to instruct a set number of hours, but there “really is no onus on the students to be there for the 110 hours to get that instruction.” In essence, York felt that the policy dictating his professional obligation in terms of time was not supported by the de-emphasis on student attendance in Ontario’s current assessment policies.

York identified student ability and student behaviour with equal frequency when he discussed the influence of classroom level factors on fair assessment. He identified the “level of help certain students need” as a potential hindrance for fair classroom assessment, and he stressed the importance of flexibility in grading, particularly for the lower-achieving students attending his program.

For students who are not ready right now to go on to college, or who just

need the OSSD, I will round a 45 or 48 up to 50. There's no point in making them repeat things over again . . . because if they don't get that 50 they may get discouraged and not continue . . . because they have faced so much failure for so many years in all kinds of conditions . . . they are proud of a 50.

Some of York's comments, however, suggested that the "leeway" in his grading system was influenced, not only by what he determined to be in the best interest of lower achieving students, but also by the behaviour of the more capable. He described a situation that was bothering him about an "extremely intelligent" student who was devoting more time to video games and sleeping late than to course assignments and attendance.

We are right at mid-term in the course I'm teaching, so we just completed mid-term reports. And there is a student who, he's playing games, I've taught him in other courses . . . so on his mid-term, I gave him 1%. And I told him why. He knows I don't have rigid marks – it's not like if you get 80% in the second half and 1% in the first half, you've failed, right? So he's still got – I know I he can do it . . . But there was a little bit of power and punishment in that 1% . . . Because I have no other way or hold over him . . . he's 18 . . . So his mark probably would have been a 25 or 35, if you look strictly at the mode. In five weeks he completed two assignments. I'm not going to give him an 80 for the two assignments he did extremely well because that would totally give him, it would mislead – even though he would know – he would read that as license to continue the way as he is doing. Of course this 1% may not make any difference to him, because he may not be there.

The course of action that York described taking in this story conflicts sharply with the principle that classroom assessment should be criterion-referenced, to which he had indicated he subscribed to in his earlier responses. It is also inconsistent with his recommendations for fair assessment, where he stressed that teachers should maintain a constructive orientation and not "punish" inattentive students. This was clearly a circumstance in which he felt that using a mechanism designed to report on achievement to communicate about behaviour was warranted. Although he recognized that absenteeism was not always avoidable because many of the students in his program had children and jobs, his knowledge about this particular

student led York to believe that ignoring unproductive behaviour in the assessment process was not in the student's best interest.

### *Summary of York's Case*

As a teacher in a program that aimed to help young adults finish secondary school, York faced challenges in maintaining the constructive orientation that felt was ideal in classroom assessment. Many of the strategies that he recommended for vignette teachers involved the use of assessment for learning, but he also explained that external factors (e.g., time constraints, student behaviour) sometimes impacted the process and made it less constructive. Overall, his responses suggested that he saw fairness as a balance between knowing as much as possible about students and their learning, and reflecting on biases and assumptions in assessment interactions. In discussing fair classroom assessment, York drew on all elements of practical wisdom, including his own experiences as a learner and his understanding of a particular program. Most frequently, his recommendations for fairness were based on the considerable pedagogical knowledge that he had acquired through his teaching experiences and his familiarity with classroom assessment policies. While his recommendations for fair assessment were often aligned with current assessment principles and policies, York was willing to adapt his grading practices when he believed that it was in the best interest of his students, especially in terms of the consequences for their futures.

### Tanar: Walk with the Student

#### *Tanar's Portrait*

The quotations given in Tanar's case portrait are from his responses to the background questionnaire (08 May 2008), the vignette questionnaire (29 May 2008), and his

interview transcript (12 June 2008).

*Teaching Assignment.* During the 2007-2008 school year, Tanar was the teacher-librarian in a medium-sized Catholic secondary school (under 1000 students), and he taught Grade 12 Writer's Craft online for his school board. As the teacher-librarian, Tanar assisted students and teachers with resources in the school library and on the school Web site. His school was located in an older suburb on the boundary of a large urban area, and it offered an arts program (dance, visual arts, music and drama), a gifted program, and the Advanced Placement program. The OSSLT pass rate in 2008 for students ( $n = 215$ ) in this school was 91%, which was higher than the provincial rate (84%). As an online teacher, Tanar's responsibilities were similar to those of a classroom teacher, except that the communication among students, and between the students and the teacher took place in a virtual environment. Students in Tanar's school board were able to take Writer's Craft online for a variety of reasons. For example, it may not have been offered in their school, or they may have been in an elite athletic or arts program that required travel. Tanar explained that teaching this particular course was an interesting experience because he had been a member of the team that developed it for his school board.

*Experience and Education.* Tanar had been teaching for 20 years, and he had been a teacher-librarian for four of those years. His teaching experience was entirely at the secondary level, and mostly in Grade 10 English (about 16 years), but he had taught all secondary grades over the years. In addition to English, Tanar also had experience teaching Business, Accounting, and Media Studies. He preferred teaching English, however, and considered himself a specialist based on a combination of experience and education.

Tanar graduated from a university in Ontario with a Bachelor of Arts and a Bachelor of Education in 1987, and a Master of Arts (English) in 1994. He had continued his professional development with Additional Qualifications courses in English (Honours Specialist, 1990), Special Education (Gifted, 1997), and Librarianship (Part 1, 2001; Part 2, 2002). Tanar had also gained considerable experience writing materials for English Language Arts for his school board and the OME. He was a member of the working committee that revised the learning expectations for the secondary English curriculum, and because of his subject expertise, he was chosen to develop several online courses.

*Views about Teaching.* On several occasions while he was discussing teaching, Tanar noted that over the years he has continued to learn from his colleagues and his students, and he believed that teachers “never know it all, there’s always a new wrinkle.” Tanar recognized the influence of his early school experiences on his views about teaching. He related that because he was the eldest son of Hungarian immigrants, he took responsibility not only for tutoring his younger brother in English, but also for communicating with teachers about his brother’s reading difficulties. In taking on this teaching role as a young student, Tanar initiated an interweaving of teaching and learning that continued to characterize his professional career. Tanar also emphasized the human aspect of teaching, noting that “we’re not machines, we’re not automatons, nor should we ever be.” He felt quite strongly about maintaining humane relations with students, explaining that “I’ve had students who have confided in me . . . it’s part of your legitimate role as a teacher. If we can’t even do that for them then why are we in the business.” Tanar seemed to enjoy assessment interactions with students, and he recalled different strategies, such as conferencing and

course evaluations, that he had used over the years. Remembering one story-writing assignment in particular, where he'd tried using cassettes to record oral feedback for students, he noted that these kinds of assessment activities had been "one of the most rewarding parts" of his teaching career.

*Ethics of Classroom Assessment.* Of the nine items on the ethics of classroom assessment included in the Background Questionnaire, Tanar identified two as ethical, four as dependent on circumstances, and three as unethical (see Table 4). During the interview, Tanar elaborated further on both of the items in the ethical category. While he felt that there was nothing "ethically inappropriate" about a teacher grading essays tests while knowing students' names (IDENTITY), Tanar didn't think it was "good professional practice." He explained that teachers often become privy to considerable information about their students, and that this could affect the fairness of an assessment because teachers are "humans." He suggested that student numbers, rather than names, be used for summative writing assignments, noting that he followed this practice himself. In discussing the item about a teacher including only positive comments on report cards (POSITIVE), Tanar shifted his position somewhat. While he had marked this as ethical on the questionnaire, he noted during the interview that teachers' comments "can't be all sunshine and roses," adding with a laugh that "you can't really say anything positive about 'does not turn up for class'." However, he concluded that teachers should at least try to maintain a positive orientation in their report card comments.

The three items that Tanar felt were clearly unethical involved a teacher who included surprise items on an exam (SURPRISE), another who deducted points for wrong answers on a

test to discourage guessing (GUESSING), and a student who received a zero for not returning a form requiring a parent's signature (ZERO). All the items that Tanar identified as unethical are practices that could introduce elements into the assessment process that would be irrelevant either in relation to the learning expectations, or to the purpose of the assessment.

During the follow-up interview, Tanar elaborated on two items that he had identified as dependent on circumstances. For the item about a teacher weighing homework heavily in final grades (HOMEWORK), Tanar gave the example of a student who was not able to attend class due to an illness, but who had previously demonstrated the ability to "handle the level that's required." Tanar argued that in such a case, including homework as evidence for the final grade would be an "appropriate decision." For the item about a teacher creating learning activities similar to the provincial assessment (SIMILAR), Tanar expressed his beliefs about the purpose of testing, specifically that "the whole idea of testing is to have students show what they've learned . . . a test is not designed to fool students . . . it's not designed to punish students." As a result, he felt that teachers should "try to prepare students for the test," and he described a variety of strategies to assist in this process, including direct instruction, group activities, and self-assessment. These nine items suggest, most importantly, that Tanar was alert to the purposes of assessment. They also indicate that while he believed in using strategies to avoid bias (e.g., masking students names), he felt that teachers should take any special circumstances into account in the process of classroom assessment.

#### *Recommendations for Fair Classroom Assessment*

Tanar shared his practical wisdom through his recommendations for fair writing assessment. He discussed varied aspects of fairness, emphasizing some more than others. For

example, in his responses to the vignettes, Tanar repeatedly stressed the importance of transparency, and he recommended that learning expectations and assessment criteria should always be “clearly indicated . . . before the work commences.” In contrast, he played down the relationship between student involvement and fairness. Although he believed strongly in the benefits of self-assessment for learning, he indicated that he was “not sure” if involving students in assessment necessarily made it any fairer. He explained that student self-assessment was incorporated into the curricular materials he had co-authored for his school board, but he avoided linking it directly to fairness, saying instead “I think it’s healthier.”

Of the many aspects of fairness that Tanar discussed, two were of prime concern. First, he believed that the assessment process should be reflective. Tanar felt that knowledge about a student was “inevitably going to sometimes effect the way that you perceive” their writing, but he recommended that teachers make an effort to “not to let a personal clash influence assessment.” What he emphasized as being most important for teachers to assess fairly, was that they “walk with the student . . . put yourself in the student’s place.” He explained that he was inspired by a colleague who once told him that “when I teach, I imagine myself walking through the course with the student, not leading, not following, walking side by side” and he added that “this is really the approach that I hope I take on my best days, and I hope that other educators can do this too.” Although he recognized that knowledge about students could fuel bias, he nonetheless aimed for a stance in classroom assessment that involved getting to know students.

The second aspect of fair assessment that Tanar felt strongly about was that it required a constructive orientation. Many of his recommendations, such as praising student

work and speaking to students about problems privately, suggested that he valued positive and respectful assessment interactions. In addition, Tanar felt that this should be modelled for students, and he advised teachers to be explicit in showing students how to provide constructive feedback for peer assessment because “they have to know that the language they write can sometimes wound, and they have to choose their words carefully.” Underlying Tanar’s emphasis on the importance of a constructive orientation was a commitment to the use of assessment for learning. In explaining what he felt was most important for fair assessment, he stressed that it was a “holistic thing” that involved more than testing.

Students can learn from the whole evaluation process because evaluation really is a teaching tool . . . if we can . . . get them to reflect on where they went wrong, to think critically about how they might learn things a bit better next time, then we’re teaching them with evaluation. If it’s simply ‘how did you do on the test?’ . . . ‘oh, I got . . . 80%, I’ve learned nothing, let’s go for the next test.’ I don’t think we’re really benefiting anybody here with that approach.

Tanar associated fair assessment with assessment for learning to the degree that his description of an unfair assessment was one in which the student’s learning was compromised. He described a situation that he had recently encountered where a student submitted the same essay for two courses simultaneously. While Tanar wanted to give the student a zero, his principal disagreed, and Tanar was forced to accept the assignment. He felt that this was not only unfair to the other students in the class, but also problematic in the long term for the student as it taught him that what he had done was acceptable.

The unfair assessment situation that Tanar described hinted at the tension between consistency and differentiation in classroom assessment. Several of his responses to the vignettes suggested that he valued consistency in assessment in terms of the parameters being

the same for all students. For example, he recommended that students have the “same opportunity to learn,” and that they receive the same amount of written feedback on an assignment. However, he also noted that there were circumstances where assessment should be differentiated in the best interests of a student.

There are, from time to time, individual cases which require us to think as humane individuals, to adjust assessment . . . and that’s fair too . . . because how can you assess someone who’s lost a mother in a car accident the same way that you can assess someone who’s simply going through the normal ups and downs of a high school situation.

Responding to a question about the tension between differentiating and being consistent in assessment, Tanar stated “I think it’s an essential tension, and I think it should be there. It will always be there . . . there is no great solution to this . . . a lot of assessment boils down to . . . being judicious.” In brief, Tanar’s interpretation of fairness in classroom assessment involved multiple aspects that centred on students’ learning and teachers’ reflection.

#### *Elements of Practical Wisdom*

Tanar’s responses seemed to draw on all elements of his practical wisdom, but most frequently he expressed himself in terms of his moral beliefs, his experience as a learner and as a teacher, and his understanding of his educational context.

Tanar felt strongly about following his own moral beliefs in making assessment decisions in particular situations. He related that he once had a student whose father had a heart attack and died at the family dinner table, and that this had occurred toward the end of the school year when the students were working on their culminating assignments (i.e., final summative tasks). He believed that although this assignment was a major component of the students’ final grades for the course, it would have been “absurd” to insist that this student

complete it. He stated emphatically that “if that’s the kind of system we’re in, I don’t want to be a part of it.” Although he drew on different elements of practical wisdom to discuss fairness, the stories Tanar shared about students indicated that his assessment decisions and interactions were often guided by his moral beliefs.

In giving his recommendations for teachers to assess fairly, Tanar relied most frequently on his own experiences. He related one event in particular that he remembered well from his experience as a learner.

I had a Grade six teacher who I think was possibly one of the worst – certainly in the top three worst days of school for me. I’d written a story. It wasn’t a good story, and there were certain parts of it that were not appropriate. She read the entire story to the class and then shot me down completely. And I was – I know I was trying to be cool, I was trying to . . . fit in with a certain group, but for her to do that I thought was just completely inappropriate. I mean, she should have talked to me . . . she should have disciplined me in private, not in front, not humiliated me in front of the entire class.

Tanar followed this childhood story with a direct link to his teaching practice, saying that “the thing is that this has happened in my classroom. It does, it’s inevitable, it’s going to happen, and I think making a big circus out of it doesn’t work.” He stressed that classroom interactions, even when they arise from some kind of problem, should be “turned into a learning opportunity.” Tanar indicated that he had come to realize this, not only through his experience as a learner, but also as he gained experience as a teacher and developed his “craft” over the years. In discussing his growth as a teacher, Tanar was aware of the interplay of the elements of his practical wisdom, explaining that “it’s experience, but it’s personal beliefs, morals, childhood experience, whatever you want to call it, but it’s also experience in the classroom.”

With his years of classroom experience, and his involvement in curriculum

development, Tanar was quite familiar with the course of reform in Ontario's educational system, and his expressions sometimes suggested disdain for politically mandated change. For example, he referred to the current "push" for a particular approach that was the "latest buzzword," and he described some materials he been required to read as "pulp." Nonetheless, Tanar seemed proud of his work with the provincial committee, and he clearly approved of certain features in the revised curriculum, especially the emphasis on student metacognition. He expressed considerable understanding of the particular educational context in which he worked, and this was evident in his practical wisdom about fairness in classroom assessment.

### *External Factors*

Tanar identified relatively few factors that might influence the fairness of classroom assessment. At the classroom level, he noted that "class sizes can make things difficult because . . . if you have lots of students . . . you can't assess as much, so even right off the bat you have an issue." He also identified students as a factor that could influence the fairness of classroom assessment in his responses to almost every vignette. He wrote that the "maturity level of an entire class can make peer assessment challenging" and he mentioned "class dynamics" several times, including "personal issues between students" and the "relationship between the teacher and the student." He felt that the attitude and behaviour of individual students, as well as the "nature of the class" could impact a teacher's ability to follow his recommendations for fair assessment. Despite frequently identifying students' attitudes or behaviour as a potential impediment for the teachers in the vignettes, Tanar's responses suggested that he was comfortable managing these challenges. After recommending that the teacher in one vignette share the assessment criteria with the students, for example, Tanar

noted that “the student in question . . . may challenge the mark. However, if the teacher has explained . . . this should not be an issue.” Rather than viewing students’ attitudes or behaviour in a negative light, Tanar suggested that they could be a constructive source of information. He explained that at the end of every course, he asked his students to complete an anonymous evaluation form to help him “learn to be a better teacher.” He used this activity, not only to gather information about his teaching practice, but also to model an openness to feedback for his students.

At the school level, Tanar’s few references to administrators tended to be less than flattering, but he voiced strong praise for his colleagues. For example, he volunteered that one of his online course collaborators was “a great guy, and he’s very creative as well.” Nonetheless, Tanar identified colleagues as an external factor for just one vignette, and he did not directly link colleagues with the fairness of classroom assessment. However, he noted the influence of colleagues on his growth as a teacher at several points. He explained that he was “lucky enough to start [his] career” with a colleague who offered guidance and inspiration, and he felt that his “guiding principles” had evolved from his interactions with his students and colleagues at his current school. With this, Tanar suggested that while colleagues might not have an immediate impact on the fairness of classroom assessment, they could influence the development of a teachers’ practical wisdom about fairness.

#### *Summary of Tanar’s Case*

For Tanar, teaching and learning were endlessly interwoven; he learned as a teacher, and he taught as a learner. Approaching assessment this way, he felt that for classroom assessment to be fair, it had to serve learning. He emphasized the importance of a reflective

assessment process that balanced between knowing students and avoiding any bias that might result from that knowledge. He was also aware of the need to balance consistency with differentiation in classroom assessment when it was warranted by individual circumstances. Although all the elements of practical wisdom were evident in Tanar's recommendations for fair classroom assessment, he most frequently expressed himself in terms of his moral beliefs, his teaching and learning experiences, and his understanding of the educational context. A strong link between teaching and learning was also evident in the external factors that Tanar identified. Essentially, he saw fair classroom assessment as a human endeavour that could be influenced by teachers and students, and thus required reflection to ensure that it supported learning.

#### Lucy: They've Got to Earn It

##### *Portrait of Lucy*

The quotations used in Lucy's case portrait are from her responses to the background questionnaire (30 April 2008), the vignette questionnaire (02 May 2008), and her interview transcript (27 May 2008).

*Teaching Assignment.* Lucy taught Grade 10 and 11 English in congregated gifted classes, and Grade 12 Writer's Craft during the 2007-2008 school year. Lucy's public school was located in a suburban neighbourhood on the outskirts of a major urban area. The school had a large number of students (over 1000) in its regular, alternative, gifted and International Baccalaureate programs. Of the students who took the OSSLT for the first time in 2008 (n = 317), 92% passed, which exceeds the provincial pass rate of 84%. Data for this group of students indicated that many learned to speak a language other than English as their first

language (69%), and few spoke English as their primary language at home (34%). However, relatively few (13%) of these students were identified as English language learners, or as having special learning needs (13%). Considered with the advanced programs offered at Lucy's school, these statistics suggest that the student population was academically strong overall, and only a small percentage needed additional educational support.

*Experience and Education.* Lucy had been teaching for a total of 40 years, of which 21 years were in Ontario. She started her teaching career in South Africa when one of her university professors recommended her for a position teaching Business English in a centre for continuing education. Following a year in England, Lucy returned to a South African university to teach in a teacher education program until 1984. When she moved to Ontario, Lucy taught initially at a college, and then began teaching at the secondary level in 1988. Over the ensuing years, Lucy's primary teaching position was as a secondary school English teacher, but she also continued to teach English as a Second Language at two adult colleges.

Lucy graduated from a university in South Africa with a Bachelor of Arts and a Teachers' Higher Diploma in 1968, and she earned a second Bachelor of Arts from another South African university in 1976. When she first moved to Ontario, she continued her professional education and obtained qualifications in Intermediate and Senior English (1987), English as a Second Language (Part 1, 1987), and Guidance (Part 1, 1987). In 1990, Lucy became an Honour Specialist in English, and she took two further Additional Qualifications courses on Computers in the Classroom (Part 1, 2004) and Guidance (Part 2, 2004). Over the years, Lucy attended many professional development workshops, and she remembered three in particular that related to curriculum development, drama, and teaching the Holocaust.

Through her career, Lucy had always taught English in some form or another, and she considered herself a specialist in teaching English based on her experience and education. In 2007, Lucy was recognized with an award of distinction from her school board, and she also received an award of excellence from the Ontario Secondary School Teachers' Federation. In the last year before her retirement (2008), Lucy was the recipient of a major award for achievement in teaching from her school board, which was one of the largest in Ontario. As she said, it "was a crowning moment . . . to leave teaching . . . to end my career with it, that was great." Even with her career officially coming to a close, though, it seemed that teaching would still be part of her life. Moments after stating that she was looking forward to retirement, Lucy mentioned that she was planning on tutoring students the following year.

*Views about Teaching.* Lucy was particularly proud of a culminating assignment that she had just developed for one of her Grade 10 English gifted classes. In her department, the final assignment for this course was usually an essay, which Lucy did not find sufficiently engaging or challenging for her particular class. She explained that she thought it was "terrible to just haul out old lessons, and just slap them down, and keep going . . . you're not doing anything for anybody." Instead, she was having her class work in groups to develop multimedia responses to a novel they had just read about the Holocaust, and she said that her students were "jumping for joy" about having this alternative. In describing her teaching practice, Lucy noted that it was "so important to be ahead of . . . students and keep up," and she emphasized this in describing her use of technology.

I'm the oldest in the department, but I probably use it more than anybody else . . . because I see it as such an amazing tool for teaching. I teach through my web page . . . we're so fortunate to use it . . . and it's terrible not to avail yourself to what there is now to benefit the students . . . I make a conscious

effort to.

Lucy believed that her students appreciated her, not just because she tailored projects to engage them, but more because of her content knowledge and high standards. She explained that the gifted students like her courses because they “trust [her] knowledge,” and she relayed that she had overheard a student describe her as “tough, but fair.” Lucy thought that most of her students would agree with this appraisal, and she emphasized this point with a story about a former student who had recently visited the school. When the student learned that Lucy was retiring, she stopped by her classroom to thank Lucy for having done “so much” for her. In essence, Lucy wanted to be remembered, at the end of a “wonderful career,” as a teacher who strove, with a combination of engaging activities and high standards, to engage her students in genuine learning while demanding their best effort.

*Ethics of Classroom Assessment.* Of the nine items on the ethics of classroom assessment that were included in the Background Questionnaire, Lucy identified four items as ethical, three as depending on circumstances, and two as unethical (see Table 4). The items that Lucy felt were clearly unethical were about a student receiving a zero for not returning a form with a parent’s signature (ZERO), and about a teacher helping a student use a test response form accurately (ANSWER). The items that Lucy indicated were ethical were about teachers weighing homework heavily in calculating final grades (HOMEWORK), creating learning activities similar to tasks on the provincial assessment (SIMILAR), grading essays tests where the student is identified (IDENTITY), and including only positive comments on students’ strengths when writing report cards (POSITIVE).

In looking over her questionnaire responses during the follow-up interview, Lucy

discussed two of the items that she had identified as dependent on circumstances. Regarding the practice of adding a few surprise items to exams (SURPRISE), Lucy thought she wouldn't "go with" the practice herself, noting that "the students really hate that." In reference to changing a student's grade based on the work completed (CHANGE), Lucy commented on how teachers are sometimes pressured to adjust senior students' grades. She describe a recent case that she found objectionable because she was asked to raise a final grade by 5% so that a student would qualify for a scholarship. Lucy explained that she did not give grades "lightly," but she would consider a smaller adjustment in similar circumstances, if she felt the student had earned it. Lucy's responses to these nine items suggest that she considered the purpose of the assessment in determining whether it was ethical, and she was also aware of the impact of assessment decisions for students. Some of her responses, such as including homework in final grades and writing positive comments, also seemed to be reflective of a school environment in which many students routinely experience academic success.

#### *Recommendations for Fair Classroom Assessment*

In her recommendations for fair writing assessment, Lucy discussed many aspects of fairness. She felt most strongly about three, which essentially related to transparency, consistency, and the appropriateness of classroom assessment.

Several of Lucy's vignette recommendations indicated that she felt transparency was necessary for fair assessment. Although she preferred to use writing models rather than rubrics as a means of sharing information about a writing task, she repeatedly noted that the expectations should be "clear right at the outset." When asked directly what was most important for teachers to consider for fair classroom assessment, Lucy identified the need for

teachers to understand the purpose of an assessment as a prerequisite for transparency.

I think that the first thing is you have to know your parameters. You have to know what, on any given task, what are you measuring here. And that's got to be made – you have to know that yourself before you can give out the exercise. So, what do I want the students to learn from this? What's the best way to assess that? I think that that's critical, and fair means that must be communicated to them because otherwise they'll grope in the dark. If they don't know what your expectations are, you can't be fair.

With this recommendation, Lucy indicated that sharing assessment information with students requires teachers to reflect beforehand about the learning expectations and assessment task.

While Lucy thought that teachers should begin by articulating the expectations and criteria, she placed more weight on the importance of consistency for fair assessment. She noted, for example, that she marks student work “to one standard and one standard only. I do not have a sliding scale for students because of how they have been identified. It is essential that one is consistent.” She raised this point several times, stressing finally that consistency was “the most important thing.” Although this suggests that treating students equally was paramount for Lucy, other comments indicated that she was not at all opposed to differentiated assessment. She described a situation where she had modified an assessment task for a student with Asperger's syndrome, and she said, “I don't see that as being inconsistent. I'm just making allowances for those differences.” Lucy pushed this point further, explaining how she saw differentiated assessment for particular students as being consistent. “It's consistent in the way, in my philosophy, and that is that I will always make allowances for you.” Lucy relied on information about particular students in her assessment process, and she made it clear that her perception of their effort was a determining factor in the degree of flexibility she maintained.

Fairness for me is knowing that the student has done the job honestly and with a maximum effort in order for me to grant them any – give them the advantages that they might not have had in the first place. I've got a class of wonderful students who – this is the deadline and they'll come and say to me, 'I can't make it.' Because I know how hard they've worked, and their usual effort, I'll say, 'that's fine.' I would never do that to some other students. So it really is based on my sense of how honest they are, the sincerity of their need, and their performance.

Lucy felt that her emphasis on effort made assessment meaningful for students. She said that she saw “too many teachers just giving away marks that mean nothing, and the students don't gain anything from that.” She explained that she had never lowered her standards, and that she didn't “just give” students their grades “under any circumstance.” Lucy thought that her students knew that she would “never give them anything gratis,” and that they had to “really earn it.” She believed that anything less would be a “disservice” to her students.

Lucy also felt that differentiating assessment to ensure that it was appropriate for the type of class made it fairer. This was evident in the rationale she gave for changing the assessment tasks for her gifted classes from those designed for the regular English classes.

It's much more fair because I'm not just looking at something that they know how to do. And I think I'm evaluating for more skills, and I think it's much more – I keep talking about the gifted students, but because they've got so many levels of learning, and I think it's fair to be challenging them at that level.

As such, fairness for Lucy involved assessment that was either appropriately designed for a group of students, based on her knowledge of learning theory, or appropriately differentiated for individual students based on her understanding of their particular needs or characteristics.

### *Elements of Practical Wisdom*

In her recommendations for fair writing assessment, Lucy seemed to draw on all the elements of practical wisdom. Most frequently, though, her practical wisdom about fairness

was expressed as moral beliefs and ethical knowledge, and knowledge of learners.

Lucy remembered that when she first came to Canada, she was really surprised by the high grades students received. In her experience as a student in South African classrooms “to get a 70 was really good, 60 was good.” She related a story about her first years of teaching in Ontario where she worked with a department head who had a sign over his desk that said, “When all else fails, lower your standards.” She stressed that she was proud of never having lowered her standards, and she explained that “I like to think, and I really, really do believe . . . that I have a set of values . . . and I’ll never compromise my values because otherwise I think I am failing in my job as a educator.” As she discussed her teaching philosophy, Lucy expressed the belief that “in order to gain something, you’ve got to work for it. It’s certain that you can’t just open your hand, you’ve got to work for it. You’ve got to carry logs.” This moral belief manifested itself not only in Lucy’s expectations for her students, but also in how she approached her teaching. In addition to the hours she invested in developing new units and marking student writing, Lucy also devoted her time to assessment for learning. She described, for example, a situation where a student had “written terrible stuff,” but after their meeting “he did it over at his own choice. I helped him with it, and he gave me, really, a much better piece of work.” She explained that when she knows that a student “can do something that is good” she will “work hard to get them to do that.” In essence, Lucy applied her belief in the value of hard work to all aspects of her practice.

Lucy expressed considerable knowledge about learners, which she tended to support with examples of particular students drawn from many years of teaching experience. In her recommendations for the vignettes, she made several general statements about students, such

as “often inattentive students are inattentive because they are having difficulties following,” and “even when students have been identified as gifted, they are not necessarily good writers.” Lucy brought these statements to life during the interview with stories about her students. She described a student in one of her current classes as a “math whiz” who “doesn’t like the creative group stuff, he likes something logical,” and she explained that to assess his writing fairly, it was best to “give him an expository essay . . . for him to write something about his emotions – he hates it.” Lucy also described another student that she’d had years before who “tried so hard, but she was weak.” She remembered that when she talked to this student about her achievement “tears started running down her face,” but the student had returned to visit her in later years, and had said that she appreciated the honest feedback.

Lucy indicated that her teaching experience informed the recommendations she gave for many of the vignettes. In one response she noted that she had “enough experience to know” when to re-teach a lesson, and for another where she felt it was important to reduce students’ writing errors, she noted that “experience has shown me that simply teaching grammar does not really do this.” However, she explicitly referred to her teaching experience far less frequently during the interview, and it seemed that her past experience had simply become part of her current knowledge. Sometimes, when she did refer to the past, it was to express professional growth. For example, when she was talking about differentiating assessment she said, “there might have been a time that I’d say, ‘you’re doing it this way and that’s it,’ but I changed from that a long time ago.” She explained that at this point in her career, she was “very secure in the fact that I can do that and not either compromise myself or lose face.” This confidence, however, did not preclude ongoing learning for Lucy. A few

weeks shy of her retirement, she remembered the “wonderfully inspiring teachers” she’d had as a learner, and she noted “even now . . . I was thinking how much I still could learn from my students, from my colleagues.” Overall, Lucy was confident in expressing her practical wisdom about fairness in classroom assessment, and she seemed to enjoy sharing it through stories about the students she had encountered during her forty years as a teacher.

### *External Factors*

Lucy identified external factors that might influence fair classroom assessment at all levels. She referred to time constraints at several points, noting the amount of time required for planning lessons and marking student writing. She also indicated that time constraints influenced how she designed an assessment task.

So they could do that as the oral component, but I don’t have time for that, so they’re going to write, put together – their going to work in pairs – either a newspaper, magazine or scrapbook where they have to get pictures.

In designing this task, Lucy eliminated one component of the assessment and cut the number of final products in half by pairing students to ensure that the assessment was feasible within a busy end-of-year timetable. The majority of Lucy’s comments about factors that could influence fairness tended to focus on students within the classroom environment. On several occasions, Lucy noted differences in the attitudes of “non-academic” students and students in the gifted program. For example, she wrote that her response to one vignette was “based on the fact that I know for certain that they hardly ever read what the teacher writes anyway - unless they are enhanced students.” She also explained that the difference in the way these students approached assessment information lay in their expectations for learning.

The academic students want to get the most out of it, they want to work hard, they want to feel that they’re learning something. There’s some students that

don't care if they learn nothing . . . if they're having a good time in class . . . that's fine.

In all her comments about students, however, Lucy's tone was positive, and even the above observation was offered with laughter that suggested acceptance. When she spoke specifically about the students in the gifted program, Lucy's comments were often glowing. For example, when she described the alternative assessment activity for one class, she said,

Even though this is going to be more work for them, they'd rather do it this way because that's the way they are – but those are gifted kids. Which student is going to do more and be happy about it? It's just that they need to have that variety, to use those creative skills that they so excel in.

In identifying students as a factor that could impact the fairness of an assessment, Lucy was sometimes more concerned about the teachers' knowledge about students than the students' actual behaviour. When she explained that her experiences with peer assessment had always been very constructive, for example, she also recommended a strategy for teachers to build on students' attitudes and avoid a negative assessment process.

I've never had to say, 'Don't laugh at each others mistakes'. . . they take it very seriously when you're doing that sort of thing. I think what probably gets around that . . . is having them work in groups where they are comfortable with the others in the group, where there is respect for the other people in the group, and there's friendship. I let them choose their own groups for peer editing because otherwise that might lead to that sort of situation. And they value one another's opinions far too much, and their friendship.

In brief, Lucy identified a range of external factors that might influence a teachers' ability to follow her recommendations for fair assessment, but she seemed to view few as major hindrances. With many years of teaching experience behind her, Lucy had a wealth of knowledge to draw on and considerable expertise at hand.

### *Summary of Lucy's Case*

As an English teacher in a school environment with several advanced program options and many academically strong students, Lucy's high standards for student learning were not out of place. Lucy worked hard to engage and support her students, and she expected a similar effort from them. This approach to teaching was reflected in her recommendations for fair writing assessment. She noted that the first step toward fairness was for teachers to be aware of the purpose of an assessment, and to communicate the expectations clearly to the students. While she emphasized the importance of consistency, Lucy indicated that this did not mean that differentiated assessment was unfair. On the contrary, she felt that differentiating assessment was consistent with her teaching philosophy, and she was confident that her students understood that she would always accommodate their needs. A caveat that accompanied Lucy's association of differentiated assessment with fair assessment was that student effort had to warrant any allowances made. This view of fairness was informed by Lucy's moral belief in the merit of hard work. While she drew on all the elements of practical wisdom in her recommendations for fair assessment, Lucy expressed herself most frequently in terms of moral beliefs and ethical knowledge, and knowledge about students, the later of which was often supported with stories about particular students. In identifying external factors that might influence fair assessment, Lucy didn't view many hindrances as significant. Most of her comments related to students in the classroom, but rather than interpreting students' attitudes as a problem, Lucy tended to focus on the knowledge and strategies that teachers could use for fair classroom assessment.

## Amada: Never Settle for Mediocrity

### *Portrait of Amada*

The quotations used in Amada's case portrait are from her responses to the background questionnaire (30 May 2008), the vignette questionnaire (14 June 2008), and her interview transcript (28 June 2008).

*Teaching Assignment.* Amada was a retired teacher who tutored elementary and secondary students, and taught as an occasional teacher at an adult learning centre during the 2007-2008 school year. The learning centre was located in a densely populated and culturally diverse area that had amalgamated with a large city through urban expansion. The students that Amada tutored were also from this area, and she had been working with several of the same students for multiple years. Some of her students were actually the children of students that she had taught before her retirement. Although she also tutored in French, Amada explained that because most of her students "didn't grow up here," her tutoring usually focussed on helping them improve their English writing. Some of her students were preparing to take the essay component of the General Education Diploma exam, and one was aiming to move from Grade 10 applied English to a Grade 11 academic class in order to apply to a university program the following year.

*Experience and Education.* Amada had been teaching for a total of 40 years, 38 of which had been in Ontario. She began teaching at a private school in Jamaica after she graduated from high school in 1958. She taught English, Latin and Mathematics there for two years before moving to Canada. She graduated with a Bachelor of Arts from a university in Ontario in 1964, and she completed a professional education program in 1969. Amada

pursued graduate studies and professional development concurrently, earning an Honour Specialist in English (1973), an Honour Specialist in Spanish (1973), as well as a Master of Arts in Spanish (1974). She continued her education through the next three decades with Additional Qualification courses in English as a Second Language (Part 1, 1985), French as a Second Language (Part 1, 1986), Adult Education (1990), and Reading (Part 1, 2004). Amada also earned a Master of Education in Ontario in 1986, and she completed two Principal qualifications courses in 1988 and 1989.

Amada explained that it was a “fluke” that her first teaching assignment was in a city in northern Ontario. It was simply the first job offer she received after a long series of interviews. Although she wasn’t initially planning on staying in Canada, she met and married a “German-Rumanian guy . . . from Niagara Falls” who was also an English teacher in northern Ontario. Over the years, Amada taught Spanish, French, Mathematics, Consumer Studies, Personal Life Management, Co-operative Studies, and Literacy Skills in addition to regular English courses. She also taught English as a Second Language and Spanish to Francophone students in a university outreach program. Amada was the literacy leader for her school in the 1980’s, and she then became the co-ordinator of an alternative secondary school program. She held that position until 1997 when she was sufficiently frustrated with the impact of funding cuts in education that she retired and moved to central Ontario. Amada subsequently taught English as a Second Language part-time in a continuing education program, but when her husband died, she decided to return to teaching full-time. From 2000 to 2005, she taught English (including Literacy) and Spanish courses at a secondary school in her neighbourhood before retiring for the second time. In keeping with the choices she made

through her long career, Amada indicated that she preferred teaching Spanish, English Literacy, and English as a Second Language, and she considered herself a specialist in these areas based on her experience and education.

*Views about Teaching.* Describing her development as a teacher, Amada noted that “it didn’t take very long” before she began to assume leadership roles. She started as the Assistant Head of Modern Languages before becoming a literacy leader and program coordinator. Working with teachers who were thirty years her junior in the English Department in her last secondary school, Amada related that she found herself to be “the person who sat them down and taught them grammar.” As such she viewed herself as having been a mentor to other teachers for much of her teaching career.

It was abundantly evident from Amada’s academic history that she valued education. She explained that during her childhood in Jamaica, secondary education was not a “given” for all adolescents.

It was a privilege because your parents had to pay for it. If they didn’t have money, you couldn’t get a high school education. And social mobility in Jamaica depends on education. When I was growing up it was a British-type society, class society. And so for you to get anywhere you had to get more education, or else you married somebody who had money.

Amada recognized the potency of her experience, noting that it was “always in the back of my mind” and that it “has always guided my teaching a lot.” Although she was multilingual, and had raised her own children to be bilingual, Amada felt compelled to encourage her adult students to prioritize English.

Being a minority person, I know that . . . I’m dealing with adults, they’re doctors, they’re lawyers, accountants, whatever, and they’re working in factories right now just for the fact that they don’t have enough . . . English. So it’s a different education for them, I’m telling them, “Look, you’ve got to

stop speaking your native language, and go make some Canadian friends, and you've got to practice that English so that you can get better at it because unless you do, you're not going to get out of the ghetto you're in.

Amada felt an "urgent" calling at this point in her life to "teach them how to wade through the waters, and what they need to do to make themselves better." This comment is indicative of a key aspect in Amada's understanding of her role as a teacher. She stressed the importance of learning to learn, noting that it had "always been my goal as a teacher to remove myself from the process." She added that when students become "their own teachers, their own guides, when I'm not needed anymore, then I know I've done my job." In sum, as a teacher Amada was driven by a social imperative to help others, especially newcomers to Canada, in becoming independent learners and upwardly mobile members of society.

*Ethics of Classroom Assessment.* Of the nine items on the ethics of classroom assessment that were included in the Background Questionnaire, Amada initially identified one as ethical, two as dependent on circumstances, and six as unethical (see Table 4). The only item that Amada felt was clearly ethical was about a teacher creating learning activities with tasks that are similar to those in the provincial assessment (SIMILAR). Amada elaborated on both of the items that she felt would depend on the circumstances. She explained that the first, where a teacher announces she will deduct points for wrong answers on a test to minimize guessing (GUESSING), would depend on the students. If they demonstrated little knowledge about a subject they should have studied, then she felt it would be ethical to penalize them for guessing. For the item about a teacher knowing the names of students while grading their essay tests (IDENTITY), Amada commented that in her experience, word processors and student numbers were used to mask writers' identities. She noted, however,

that in teaching any English course teachers quickly became familiar with the students' writing styles.

Amada identified the six other items on the questionnaire as unethical, and she discussed three of these further during the follow-up interview. She was emphatic about the item on surprise questions (SURPRISE), saying "I would never add surprise items to a test. I think kids deserve fairness, and if you tell them you're going to do a, b, c, then you should reflect a, b, c." Amada also reiterated that including only positive comments on report cards would be unethical (POSITIVE), explaining that it would "paint an overly rosy picture" for parents. Amada did reconsider the ethics of one item. She thought there were circumstances in which it would be ethical to change a student's final grade (CHANGE). The example she gave was of an assignment being handed in late due to a serious event in the student's life, such as a parent's death. Then she felt that an adjustment could ethically be made to the student's final grade. Although Amada initially identified most of these nine items as unethical, her comments indicated that she was aware of circumstances that would shift the ethics of classroom assessment. Overall, she tended to select the responses that would be supportive of student learning, which suggests a constructive orientation to assessment.

#### *Recommendations for Fair Classroom Assessment*

Amada shared her practical wisdom through her recommendations for fair writing assessment. She discussed several different aspects of fairness, and she explicitly referred to being fair in different ways. For example, when she stated that it is "not fair to constantly mark for everything" in students' writing she was in the process of explaining that the number of learning expectations for any one assignment had to be feasible for the students. In

contrast, when she said that she was “always wanting to be fair” in her comments and marking, Amada was describing the negative consequences of inaccurate feedback, and asking a hypothetical student, “What would you have me do? To tell you that it’s good, and then you go off thinking you’re such a super writer, and then end up going to university and all of a sudden you’re in the F category?” A third way in which Amada used the word fair raised the issue of treating students equally in the classroom.

As a teacher I have always wanted my students to view me as fair in my practices because they talk a lot among themselves, and so I’m cautious to make sure I’m not giving one more attention than the other, or praising one more than the other. I try to be as democratic as I can . . . to be equal.

While Amada recognized the role of student perceptions in classroom dynamics, she also saw the need for differentiating assessment. She explained that she tried “to be consistent, but sometimes . . . depending on the circumstances and students you’re dealing with you might have to use some form of differentiation.” She added that she did not view differentiation as “a sign of inconsistency,” and she did not seem concerned about it in terms of fairness.

When asked directly what she thought was most important for fair classroom assessment, Amada quickly articulated several points, which in essence, involved appropriate assessment criteria, tasks and tools, and a constructively-oriented assessment process. Amada felt strongly about the importance of teachers’ knowledge in developing appropriate assessments. She believed that a teacher not only needed a “high level of mastery of his or her subject material” (i.e., content knowledge), but also a “certain modicum of common sense in dealing with young people.” She explained that she saw fair assessment as being able to “balance . . . knowledge of where the student is, what the student is capable of doing, and what it is they can realistically demand from them at this point in time.” In describing her

approach to peer assessment, Amada gave an example of how she used her knowledge about students' abilities to support fair assessment as the dynamics of the classroom evolved.

Kids are sometimes reluctant to mark their peers' work, but you have to teach them how . . . . So what I do is, I know my students strengths and weaknesses, and so when I am putting pairs together then I put people of similar – initially– of similar backgrounds that will be accepting – this one doesn't spell and that one doesn't spell, so they can help each other through that. At a later stage, though . . . when they are a bit more comfortable with accepting criticism from people that they know to be better writers than they are, then I'll put a much stronger writer . . . with a weaker writer.

Amada also noted, in emphasizing the importance of teachers' knowledge, that a teacher should “have at his or her disposal a wide range of assessment tools because the same tool can't be used for everything.” From Amada's perspective, teachers' knowledge, in which she included content knowledge, knowledge about classroom assessment, general knowledge of students, and an understanding of particular students, was paramount for fair assessment.

The second aspect of classroom assessment that most concerned Amada for its fairness was that it be constructively oriented to support learning. This was evident in several of her recommendations for the vignette teachers. For example, in response to the vignette about peer assessment, Amada wrote that the teacher should “speak to the individual student about his/her lack of respect” and “address the whole class to ensure they understand peer-editing techniques.” For the vignette on writing errors, she explained that she would first “congratulate students on progress made in journal writing so far, then establish new expectations for future journals.” Amada's understanding of assessment as an ongoing process was clear in many of her comments. For example, her description of giving students feedback involved much more than a unidirectional sequence.

I think they learn more because then I am back to them assessing. I have

assessed it, this is what I've said, now they need to read over what I have said. And then they have to internalize it, and assess what I have said and how it relates to what they have done . . . and then when they usually come back to see me, they have to tell me . . . what are your plans, how do you intend to improve, what's the next step . . . it's like a cyclical process. They start here and they move forward a little and then we've got to move the bar again and start the cycle over again.

Amada clearly saw fair classroom assessment as an ongoing interaction that focussed on students' learning, and she emphasized that teachers needed to be responsive in this process.

As the student grows, then you move the yardstick a little bit more . . . the teacher should never be contented with whatever effort that is given by the kid. Praise him, praise her for whatever she's done, but always keep saying . . . there's a little more that you can do . . . so never settle for mediocrity.

From Amada's perspective, the fairness of classroom assessment rested squarely in the teacher's hands; it depended not only on teachers' knowledge, but also on their ability to respond constructively to student learning.

#### *Elements of Practical Wisdom*

Amada's recommendations for fair classroom assessment involved, to some degree, all the elements of practical wisdom. Her sense of professional ethics was heard in several statements, such as "it's my job to teach as many times as needed using different approaches if necessary until students grasp the material adequately." Her understanding of the particular educational context in which she worked was also apparent in her description of the changes that had taken place over the past few decades. Most often, though, Amada expressed herself in terms of knowledge about English writing and pedagogy, and in relation to her experiences as a learner and a teacher.

When Amada discussed the recommendations she had given for fair classroom assessment, she used examples that displayed her knowledge of the elements of style and

mechanics in writing. Explaining why she thought it was important to show students exemplars of good writing, for instance, she said that it helps them “move up to a higher level of expressing themselves . . . rather than writing in straight simple sentences or compound sentences, to do that in combination.” Amada explicitly linked this type of content knowledge with fairness in classroom assessment.

That is extremely important because how can you assess adequately if you don't know the criteria you are applying, or what constitutes that criteria . . . they have to be certainly more than familiar with their subject area, especially in English, if you're assessing writing.

During the interview, she also emphasized the need for teachers to develop their knowledge of educational theory more generally.

One of the things I tell them is they can't just be contented to just go teach, teach, teach. They have to read educational philosophy; they have to get some idea of theory; they have to have some understanding of why they do what . . . so introducing them to simplified forms of some of the models of teaching . . . the philosophical underpinnings of this approach, or that approach.

A similar pattern was seen in regard to technical knowledge. Amada displayed her own knowledge of pedagogical strategies in her recommendations, and she elucidated the need for other teachers to develop this knowledge. For example, Amada noted in response to one vignette that “that “rubrics are a clear outline of expectations for assignments,” and she later noted the need for teachers to develop their knowledge of varied assessment tools. Amada stressed the importance of teacher knowledge repeatedly in her responses, and her thoughts in this area seemed to be strongly influenced by her leadership experiences in education.

Amada clarified her responses on several occasions with references to her experience as a learner. Amada attended a private school in Jamaica in the 1950's, which she remembered as being much more stringent in its academic demands than the system she

encountered when she moved to Canada.

Here they struggle with one Shakespeare. When we did English in Jamaica we had – it was almost like a historical approach to English. You had to know all the periods of English, through the Romanticism, the Victorianism, all the way through. And when you started doing Shakespeare, it wasn't a question of one Shakespeare, you had to do a representative historical play, a tragedy, a comedy . . . so it was a much broader thing. And then when you went to write essays, it wasn't any of the stuff we can regurgitate. You'd get things like 'There's nothing new under the sun. Discuss' . . . very abstract topics that you had to put your own application to. . . so we did a lot more . . . critical thinking. There was no such thing as your teacher coming in to class taking Shakespeare and going through line by line with you. When you came to class, you should already have read the play, know the content, and your teacher was attacking from the critical level. So when I came here I saw what passed for teaching, I found the standards very low.

Amada realized that her experience as a student in a rigorous educational program influenced how she taught English Language Arts. When she first started teaching in Ontario, she was determined to get her students “beyond just memorizing stuff” to the more critical approach she had experienced, and she pushed them to improve with her feedback. She related that when her students complained that she was never satisfied, she told them that “I don't want you sitting there feeling that's the best you can do. Because there is always – every time you reread something you're bound to find some way you can improve it.” In this light, Amada's experiences as a learner informed the practical wisdom she offered as a teacher about fairness in classroom assessment.

Amada indicated that most of her recommendations for the vignette teachers were informed by her “past experience in assessment of writing.” She explained that as a teacher, fairness dilemmas became less of an issue as her career progressed, and experience is clearly beneficial in this way. However, with the advantage of hindsight, Amada also identified a tension in the process of gaining experience. She noted that while “you can't step in the same

river twice because the water is flowing” and you have to keep moving “with the tide” of educational change, but she also felt teachers should “have that constant in their belief system to be successful,” so they shouldn’t “just bend” with every tide. Ultimately, Amada explained that developing practical wisdom was a process that required “exposure” to good mentors, ongoing professional development, and classroom experience until “you get this amalgam together, and then you evolve your own personal philosophy that guides whatever it is that you do.” Amada’s practical wisdom about fairness in classroom assessment was, just as she described, an amalgam that drew on the beliefs, knowledge, and experience she had gained over the course of a long teaching career.

### *External Factors*

Amada identified a variety of external factors that might hinder or facilitate fair classroom assessment. At the classroom level, she noted that class size could be a problem, especially in adult English as a Second Language programs with over 50 students. She related that “when you’re dealing with huge classes, particularly when you have to mark at night, you get bleary eyed, and sometimes you get to the point where you wonder . . . am I reading this . . . and assessing it correctly?” At the school level, Amada noted that colleagues could be helpful in terms of sharing techniques, but she explained that even if she consulted a colleague, she would stay with what she felt was “right” for her own practice.

Most of Amada’s responses relating to external factors involved those at the system level. She identified the impact of time on fair assessment at several points, noting that a shortage of time could influence the methods and tools that a teacher chose. However, her comments also hinted at the autonomy teachers have in their usage of time. For example, she

believed that peer assessment was sufficiently valuable in the process of learning to write that teachers should “take time” to show students how to assess fairly. Amada also identified time as an issue in relation student understanding of criteria, explaining that “for them to know what makes good writing, they have to be taught, and the teaching process is not a one day job or a one week job. . . it’s a gradual thing.” For Amada, fair assessment meant that the aspects of writing to be assessed (i.e. criteria) needed to be introduced over time so that the task was feasible for students and the process constructive for their learning.

In her responses to seven of the eight vignettes, Amada indicated that a teacher’s ability to follow her recommendations for fair writing assessment would be impacted by the teacher’s knowledge. Given the content of the vignettes, it is not surprising that she focussed on assessment literacy, noting for example, that the fairness would depend on the teacher’s “knowledge of assessment techniques” or “confidence level in assessment practices.” Amada also associated fairness with the consistency of teachers’ content knowledge.

If you took two teachers and you gave them the same essay that you were using to evaluate, you’d be stunned at the variety of marks that teachers would give because it seems that the standard by which they mark is so . . . shifting. I think that’s because their criteria for what is good work is not very well established . . . in their minds.

Although teachers’ knowledge is part of their practical wisdom, and hence, not considered an external factor in this study, Amada repeatedly identified it as such. She saw teacher knowledge as the key to fairness, and in her final recommendations for fair assessment, Amada emphasized the importance of teachers’ ongoing professional development.

#### *Summary of Amada’s Case*

Looking back over her questionnaire responses during the interview, Amada voiced

the realization that her “whole life has been teaching.” Amada learned to value education as a child, and she not only pursued higher education herself, but devoted her career to sharing her knowledge with others. Amada discussed many aspects of fairness in her recommendations, but she felt most strongly that fair assessment depended on teacher’s knowledge and ability. In this she included content knowledge, pedagogical knowledge specifically about classroom assessment, general knowledge of students, an understanding of particular students, and the teacher’s ability to respond constructively to student learning. Amada drew on all elements of practical wisdom in giving her recommendations for fair classroom assessment. Most often, she expressed herself in terms of knowledge about English writing and pedagogy, and she drew on her experiences as a learner and a teacher. The value Amada placed on education, and on teachers’ knowledge in particular, permeated all her responses. While she recognized the potential influence of a wide range of factors on the fairness of classroom assessment, Amada placed the most emphasis on the need for professional development to ensure fair classroom assessment.

## CHAPTER SEVEN: CROSS-CASE RESULTS

*We must try to understand teachers' actions and reactions from their perspective in the classroom, because what may look like foolishness to an observer in the back of the room may look like the only route to survival from behind the teacher's desk.*

Shulman (1987/2004, p.264)

In this chapter I summarize the most salient aspects of the information gathered to provide a cross-case description of the results. I have organized it into three sections that correspond directly to the research questions for this study. The first relates the participants' practical wisdom about fair classroom assessment, the second describes the elements of practical wisdom at play in their recommendations, and third reports on the external factors they identified as influential in the process of fair classroom assessment.

### Practical Wisdom about Fairness in Classroom Assessment

Looking across the cases, different aspects of fairness in classroom assessment were highlighted in each of the participants' recommendations. I have organized this section according to the six that emerged as the most prominent in the context of this study: offering students opportunity, reflecting about assessment interactions, sharing expectations and criteria, establishing a constructive environment, maintaining consistency in the assessment process for a group of students, and ensuring the appropriateness of an assessment for individual students. I report on each of these aspects by summarizing the participants' responses to relevant vignettes, and by highlighting any issues or themes that emerged in the practical wisdom they offered about fairness. For clarity in presenting these results, the word *teacher* refers to the fictional teachers in the vignettes, and the word *participant* refers to the teachers who responded to the vignettes as part of this study.

### *Offering Students Opportunities*

The aspect of fairness illustrated in Vignette A, opportunity to learn, was clearly recognized by all of the participants. Although several indicated that they would speak with the inattentive students about paying attention in class, none of the participants suggested that these students be excluded from the additional lesson. As such, providing a student with an additional opportunity to learn, based on information gathered through interactive assessment for learning, was unanimously recommended by the participants regardless of the reason for the need (e.g., inattention or acceptable absence). Most of the participants suggested that the teacher work with all three students in a small group. Several also noted that the students' inattentive behaviour might be indicative of learning difficulties, and York stressed that the reason for the students' behaviour should be investigated.

Recommendations relating to the provision of opportunity to learn were also given for Vignette G. Although this vignette illustrated a different aspect of fairness (transparency), it was set at a similar point in the interactive process of assessment for learning. Most of the participants felt that the teacher should use her observations about the students' writing to plan further instruction. Lorena and Kevin recommended "mini-lessons" before journal writing, and Lucy suggested a "class lesson on the most common errors." In a similar vein, Tanar thought that the teacher should address the errors as a "sidebar" when the task was taken up again. In essence, the participants recommended that the teacher use interactive assessment information to adjust her teaching, and provide further opportunity for students to learn about writing conventions.

While offering students opportunity to learn was an aspect of fairness that figured in

all the participants' vignette responses, during the interviews they focused more on the need to provide students with opportunity to demonstrate learning. Lorena and York both discussed the benefits of having multiple opportunities to demonstrate learning. They felt that for classroom assessment to be fair, it should not be based on the results of a single assessment, and it should provide as a complete a picture of the learner and learning as possible. As such, they stressed the importance of seeking information from multiple sources to fully understand each learner. Lucy and Kevin also discussed opportunity to demonstrate learning, but in relation to quality rather than quantity. They felt that for an assessment task to be fair, it should be tailored to engage or challenge a group students. Essentially, Lucy and Kevin were concerned with the alignment of students' learning needs and the opportunities they had to demonstrate their learning. This is different from the assessment interaction illustrated in Vignettes A and G in terms of timing. There the teacher used information gathered during interactive assessment for learning to provide students with further opportunity to learn, whereas Kevin and Lucy both described planning culminating assignments based on their knowledge of a group of students. In summary, the participants all recommended that information gathered during interactive assessment for learning be used to provide students with opportunity to learn, but only in response to the vignettes. The examples they gave from their own practices focused on offering students multiple or tailored opportunities to demonstrate learning as part of the assessment of learning. However, even with a summative function, assessments that were tailored for a group were thought to be fairer because they were challenging or engaging, and thus provided students with an opportunity to learn.

### *Reflecting about Assessment Interactions*

Two vignettes were developed to elicit responses relating to teachers' critical reflection about their assessment practices. Vignette B, which illustrated a moment during a planned assessment where a teachers' ideas about writing style were challenged, generated more discussion than any of the other vignettes. Most of the participants' recommendations for fair assessment acknowledged that different writing styles are suitable for different purposes. However, only two participants suggested that the teacher in this vignette should reflect on his own assumptions about the criteria for good writing. The other participants all assumed that the teacher had intentionally assigned a writing task that required a concise style, and that the issue was simply a matter of clarifying this criteria. One participant explained that when he experienced a similar situation, where he realized that an assumption on his part had led to his students' misunderstanding his expectations, he had adjusted his teaching accordingly. Overall, though, this vignette prompted more responses about other aspects of fairness than about the targeted issue of teacher reflection on bias in assessment. Similar results were seen with Vignette F, which involved a teacher reacting to a gifted students' average writing. This vignette aimed to elicit responses relating to the impact of stereotypes on fairness in classroom assessment. Although almost all of the participants commented on their experience working with gifted students, and most noted that gifted students are not necessarily gifted writers, none explicitly recommended that the teacher in this vignette reflect about student stereotypes. As such, the vignettes that aimed to elicit responses about teachers' reflection as an aspect of fairness in classroom assessment generated fairly little in the way of explicit discussion on the topic.

Despite the limited focus on teacher reflection in the participants' responses to Vignettes B and F, it was an aspect of fairness that was raised by several participants at other points. York noted the possibility of teacher bias in judging the students' degree of attentiveness in response to Vignette A, and he recommended that "personal feelings must be kept at bay." Tanar expressed similar ideas about avoiding bias in assessing student writing. For example, he suggested that masking students' names on their essays helps ensure the "personal doesn't get in the way." When the participants were asked to elucidate what was most important for fair assessment, York and Lorena both discussed teacher reflection. York volunteered that the interview process itself had been beneficial for him because it was "good to . . . reflect on your own practices because it's so easy to get into habit and just do the same thing." Lorena noted the need "to be aware of . . . biases in order to assess fairly" in several of her responses, and she stressed the benefits of reflecting about assessment individually as a teacher, and with the support of colleagues. In sum, while teacher reflection was not explicitly recommended for fair classroom assessment in response to the target vignettes, several of the participants indicated that reflection played a role in their own assessment practices, and they considered it to be an important aspect of fair classroom assessment.

#### *Sharing Learning Expectations and Assessment Criteria*

The primary fairness issue in two vignettes involved students knowing about the learning expectations and assessment criteria for a task. For Vignette C, where a group of teachers developed a rubric for a cross-curricular unit, the participants all recommended that the rubric be shared with the students. They differed considerably, however, in their responses to one teacher's concern about the effect of the rubric on student's writing. Lorena

emphasized the need for clear criteria to guide student learning at the beginning of a task, whereas Lucy suggested that it could be given at a later point in the project for students to check their work. Tanar and York felt that the teacher might have a reasonable concern about the rubric, and they suggested that it could be modified to ensure that it was not overly prescriptive. In other words, they questioned the quality of the assessment tool, and recommended that it be changed if it was not well designed. They also offered two further strategies to ensure that the rubric did not limit student writing. Tanar wrote that the teachers could explain to the students that the purpose of the rubric was to guide rather than prescribe, and York suggested that the students could be involved in creating the rubric. Both of these strategies would improve the transparency of the assessment process. The same issue comes up in Vignette G, but at a later point in the assessment for learning cycle where the students were already engaged in a task (journal writing) and the teachers' expectations had not been articulated. Again, the participants unanimously recommended that the expectations and criteria be clearly communicated to the students, but they differed in whether they felt it would be fairer to give feedback for clarification during or after the writing activity.

The need to clearly articulate the learning expectations and assessment criteria for students was also expressed repeatedly in the participants' responses to other vignettes and during the interview. Their recommendations for Vignette B, for example, were strongly focused on the clarity of the expectations and criteria even though it was designed to elicit responses about assessment bias. As well, for Vignette D, where students are involved in peer assessment, several participants recommended that the teacher clarify her expectations and criteria. All of the participants' comments referred to when students should receive this

information. They said it should be when the teacher “gives an assignment” (Amada), or “beforehand” (Lorena), from the “very beginning” (Kevin), “up front” (York), and “before the work commences” (Tanar). Lucy also felt that the expectations and criteria should be made clear at the “outset” of a task, but she preferred writing exemplars to rubrics. The similarity of the participants’ recommendations, and the emphasis they placed on the importance of sharing expectations and criteria suggests not only that this may be an accepted practice in Ontario, but also that English teachers may consider it to be a key aspect of fairness in assessing writing for learning.

The participants’ responses also suggested that while sharing and clarifying learning expectations and assessment criteria makes assessment fairer, it is not sufficient as a lone strategy. Amada, for example, also stressed that to be fair the expectations and criteria had to be feasible, and she recommended that they be scaffolded gradually to allow students to improve their writing over the duration of a course. Several participants also felt that fair classroom assessment required criteria that were relevant for the task at hand. In responding to Vignette G, for example, half of the participants questioned the relevance of the teachers’ expectations for journal writing (Lorena, York and Tanar), whereas the other half assumed that mechanical errors should be addressed regardless of the form of writing (Kevin, Lucy, Amada). This is very similar to a pattern I saw in my analysis of the 2002 SAIP data. About half (51.4%) of the experienced specialist teachers in Ontario (n = 105) disagreed with the statement that “in order to write well, students need to know the basic rules of English grammar and syntax” (CMEC, 2002c, p.4). In summary, the common interest that the participants in my study shared in using assessment for learning is reflected in their

association of fairness with the quality (i.e., feasibility and relevance) of learning expectations and assessment criteria. However, the controversy more generally voiced by English teachers about what should be valued in writing, and hence what is relevant for assessing writing, was also evident here.

The question of purpose cropped up at several points when the participants discussed the expectations and criteria for fair classroom assessment. For example, York asked what the assessment was “for” in relation to one vignette, and he qualified several of his recommendations for specific purposes. Lorena also explained that teachers should ask themselves what they were “looking for” when assessing student writing. However, the participants’ recommendations also showed some assumptions about purpose. This was particularly evident in their use of different frames of reference in responding to Vignette F, where a teacher is surprised by the quality of a student’s writing during an initial assessment. York was the only participant who noted the timing in relation to the assessment purpose, stating that “it is early, so this should really be a diagnostic phase of assessment,” but several participants also indicated that the assessment should be ipsative, or student-referenced. For example, Amada felt the student’s writing should be seen as a “snapshot of where he/she is currently,” and she recommended that the teacher work with the student to “establish new goals to be attained in the next piece of work.” Similarly, Lorena thought that the “assignment ought to be used as the springboard for teaching and encouraging the writer to produce work of a higher caliber.” The other participants all recommended that the students’ writing should be judged in reference to the learning expectations (i.e., criterion-referenced) for the assessment to be fair, and their language was more consistent with assessment of

learning. For example, Tanar referred to marking and Lucy referred to a scale. In a few of their responses, the participants' comments about the purpose of an assessment were more general, and not focused specifically on expectations or criteria. Tanar was particularly emphatic in explaining that for classroom assessment to be fair, teachers must understand that it goes "beyond simply writing a test," and should be used for learning. Overall, the notion that classroom assessment should be constructive and oriented to learning, and that the students should understand this purpose, was clearly heard across the participants' recommendations for fair assessment.

### *Establishing a Constructive Environment*

Vignette D illustrates an assessment interaction between peers in a classroom where a constructive environment has not been established. All of the participants recommended that the teacher in this vignette address the whole class about this situation, and several recommended that he speak privately with the negative student first. Although several participants suggested that the assessment criteria or process needed to be clarified, none explicitly discussed the issue of the student's assumption. When students are involved in assessing writing, their biases or assumptions about what is valued can impact the process. Instead, almost all of the participants focused on the issue of respect within the peer assessment process, and the teacher's role in fostering a constructive environment. For example, Lucy recommended giving the students "very clear guidelines" to follow during the session, and Lorena was particularly emphatic about the teacher's role, stating that "fostering a supportive environment is paramount to a class community. Being proactive in the future will set up the classroom community to be respectful and everyone will be aware of the

common expectations.” In reaction to the statement in this vignette that the teacher might avoid peer assessment in the future, most participants recommended that he persist with the practice. Lucy and Amada offered strategies to improve the peer assessment process, such as giving students a mark for the quality of their editing, or creating compatible pairs of students based on previous assessment information, and Lorena felt that avoiding peer assessment after this event would give the students the wrong message.

The importance of a constructive environment for fair classroom assessment became a thread that wound through the participants’ responses to several vignettes and also emerged during the interviews. Several participants’ comments stressed the need for the teacher to actively establish expectations and routines, not only to maximize productivity, but also preempt negative situations that could impact the fairness of an assessment. For example, Kevin explained that new teachers needed to learn to “talk and walk” and “always be circulating around” to maintain a constructive learning environment. While some of these strategies might not have an immediate or direct impact on the fairness of a particular assessment, they help create an atmosphere that is conducive to fair classroom assessment. This was especially evident in Lorena’s responses because she saw a constructive environment as prerequisite to assessment for learning.

If there’s no relationship in the classroom, I don’t think learning can happen. I really don’t. And that doesn’t necessarily mean there has to be a positive relationship where everybody loves each other. It’s just you have to have some kind of common understanding of what my expectations are . . . of you as a student, and what you expect of me as a teacher . . . so having some kind of a relationship with the students before we can actually get them to produce for us . . . and that includes having a safe space as far as the physical environment in the classroom. We have to have a place that is comfortable and welcoming, and that they feel they belong. . . . They need to feel that they’re part of the class – each student needs to feel part of the class, and that they’re valued.

When asked directly if the kind of learning environment described above would make classroom assessment fairer, Lorena replied in the affirmative, saying, “I think it will,” and she explained that it would influence how she looked at students’ work.

At some point in discussing the learning environment, all of the participants alluded to their relationships with students and the methods they used for communicating with students. While Amada and Lucy both felt that their students appreciated their use of technology (i.e., email and website) to facilitate teacher-student communications, York emphasized the importance of speaking directly with students, especially if the fairness of a decision was under question. Kevin and Tanar both related that they used modeling as a means of communicating with students. Kevin seemed to use the technique primarily to establish a constructive environment, which from his perspective involved a balance between fun and discipline. He explained, “I try hard to model what I expect. It’s okay to joke, but . . . at the end of the day, you really have to get your stuff done so that’s why I usually try to do a whole lot of routines.” Tanar’s use of modeling was more directly related to fair classroom assessment. He described a quick exercise he did prior to peer assessment, where he showed his students two examples of feedback, and he said that “the kids always choose the right one because the wrong one is so blatantly awful, but . . . then we talk about that . . . what’s appropriate.” Basically, this strategy was a direct and overt approach to establishing the type of classroom environment that would facilitate fair peer assessment. Another strategy that Tanar used was to have his students “fill out an evaluation form” for his courses. Although he was not required to do this, the process allowed him to model a reflective use of assessment information that he benefited from, and that he felt would benefit for all learners.

### *Maintaining Consistency of an Assessment for a Group of Students*

I developed one of the vignettes to explore the issue of consistency specifically in relation to assessment for learning because classroom-based research indicates that teachers often associate fairness with consistency in grading student work (as discussed in Chapter 3). Vignette E describes a teacher who is engaged in a planned assessment for learning event where she is concerned about giving her students different amounts of written feedback. Of the eight vignettes, this one elicited the least discussion from the participants. A few suggested that the teacher might even out the feedback for better consistency across the group, but only if it was essential and time permitted. Overall, the participants' responses focused more on how the teacher should give her students feedback. York noted that to be fair, the teacher should provide the same amount of feedback to all her students, but that it would be more efficient to use a checklist or rubric, and more helpful to give students oral feedback. Kevin also felt that conferencing individually with students would be more effective than written feedback. Additionally, Kevin, Lucy, and York all noted that there would be no point in providing a great deal of written feedback because it was unlikely that the students would apply it to their writing without further interaction. Amada pointed out that regardless of the amount or medium of the feedback, what really mattered was for the students to "know their strengths and weaknesses." In summary, the inconsistency of the teacher's written feedback did not generate strong reactions about fairness, and the practical wisdom that the participants shared in response to this vignette centered on the quality, or effectiveness of the feedback for students to know how to proceed with their projects.

In order to elicit the participants' practical wisdom about consistency as an aspect of

fairness in assessment for learning, it was necessary to probe beyond Vignette E. What emerged during the interviews was that the participants' concerns about consistency were often related to students' perceptions. For example, Lucy, who had no qualms about differentiating assessment for groups of students or individuals with special learning needs, repeatedly commented on the need for teachers to assess student work consistently. She felt that it was important because "otherwise they [the students] lose faith in you." Both Amada and Kevin explained that they tried to treat students as equally as possible because the students shared their perceptions about a teachers' fairness with each other. Describing the influence of feedback from students on his decision-making, York also made reference to students' perceptions of fairness.

It [feedback from students] does influence in the sense of you want to be fair to some kind of abstract ideal of what fairness is, but it has to be really fair, and that helps. . . . fairness is something that students feel or intuit. And you can't reason fairness into them.

York seemed to indicate that what is thought to be "really" fair in classroom assessment is that which is felt by those in the classroom. This is seen in the situations that Kevin and Tanar described, where they disagreed with the principals' decision about a student's grade. They both voiced concern about the unfairness of the outcome for the other students in their class. Tanar felt that passing the student who submitted a single essay for two courses was a "slap in the face to all the other students who didn't do that," and Kevin thought that it wasn't fair for an absentee student to pass "when everybody else had put in their . . . eighty classes." In both of these situations, the equal treatment of students during the assessment of learning process was disrupted, and fairness was perceived to be lacking by the participants on behalf of the other students in the class.

As Kevin and Tanar related their stories about an unfair assessment event, it was clear that they did not enjoy having their assessment decisions overruled by administrators. There seemed to be, however, more at hand than a matter of dented authority. Both Kevin and Tanar expressed concern about the long term impact on the individual students involved. Tanar stated that what was “even more important” than fairness to the group was that “this student has learned nothing . . . unfortunately this student will likely do this again, but maybe in university . . . it’s a life consequence for him.” Kevin also felt that the outcome in his case was counter productive for the development of the student’s learning habits.

I think it’s giving them [the student] the wrong message that as long as the job gets done, it doesn’t really matter . . . when you do the work . . . and I don’t think that’s right . . . If you don’t show up for work, and all of a sudden you walk in on the day it’s supposed to be done and say, “here’s the report” I’m sure they’re not going to say, “yeah, great.” And they may be happy to have their thing, but, “where the heck were you. . . to me that’s a big fairness issue.

In both these situations the participants felt that negative consequences would result for these students because they were not treated in the same way as the other students in their group, and the differentiation was not warranted by the circumstances.

#### *Ensuring the Appropriateness of an Assessment for Individual Learners*

Long term consequences resulting from an assessment were also of concern for Amada and Lucy, but they focused more on the need for assessment interactions to be appropriate for an individual student. For both, this meant that communication had to be honest, rather than a matter of platitude. Amada stated at several points, in relation to report card comments and in giving students feedback on their writing, that teachers’ should address students’ weaknesses as well as their strengths. Lucy explained that her goal in giving students feedback was to make them “aware of the fact that, ‘Yeah, I’m not perfect at this,

but I'm getting better at it, and it's being recognized.” Asked about the impact of honest feedback on students' self-esteem, Lucy replied that students did not “suffer” when “real” feedback was given constructively by a trusted teacher. To support her position, she related a story about a student she had taught years before who had struggled in English. Even though the student cried when Lucy gave her honest feedback about her writing, she had returned as an adult to visit. According to Lucy, this student was thankful for the feedback because it was appropriate, and thus meaningful, for her learning.

Another way that the appropriateness of an assessment was discussed was in terms of accommodations or modifications. Vignette H was developed to elicit responses about differentiating a portfolio assessment to make it more appropriate for an individual learner. The participants were unanimous in recommending that the teacher in this vignette talk to the student, and most recommended that she find out why the portfolio assignments were missing. Almost all of the participants felt that the student should complete all of the work assigned, and they did not address the possibility that the student might have already met the learning expectations. There was very little agreement among the participants about how the teacher should go about getting the student to complete the work. Amada and Kevin recommended some form of penalty, depending on the reason the student might give for not having done the work, whereas Tanar referred to the OME policy that discourages late penalties. Lorena suggested that the student should be involved in negotiating a plan for completing one of the missing assignments, and York felt that the student should be given a “chance to express what they need” because some assignments might be “more important than others.” In this variety of suggestions, only two participants (Lorena and York) hinted at

the possibility of modifying the portfolio requirements for this student. This lack of emphasis on differentiation is similar to the responses for Vignette F, which described a gifted student with average writing skills. Although most participants mentioned their experience with gifted students, only two noted that modified expectations, such as an individual education plan or a congregated group plan, might be appropriate in that situation.

The participants' responses to Vignette F and H, where the students are not struggling or below average academically, contrast with their recommendations for Vignette A, which involved two inattentive students. There the possibility of learning disabilities was noted, and the participants seemed to have no difficulty with differentiation. Two connected issues seemed to emerge here, the first of which relates to timing. Vignette A was situated very early in the assessment for learning cycle, and the additional opportunity to learn could be considered as a form of differentiated instruction. Kevin described a similar situation in his practice, where he gave a student who had just been suspended, and who confessed that she did not understand how to tackle essay writing, an individual lesson so that she could work on the assignment at home. In keeping with the participants' responses to Vignette A, Kevin did not seem to view this additional instruction as a form of differentiation. In contrast, when he described a situation later in the assessment process, where students were presenting completed projects in class, Kevin felt that any form of differentiation had to be warranted. He explained that he would modify an assessment process if it would result in "psychological trauma" for the student, but the need had to be clear.

The second issue connected to this relates to the perceived degree of need. When Lucy and Lorena discussed differentiating assessment they also gave examples of when they

deemed it to be appropriate. Lucy recounted a situation where she modified an assessment task for a student with Asperger's syndrome, and Lorena recommended that a student who had physical difficulty writing be accommodated with a computer. These two examples involve students who clearly require differentiation to ensure that classroom assessment is fair in terms of appropriate opportunity, and they involve the types of modification and accommodation suggested in OME assessment policy. When the need for differentiation was not mandated, or as clearly evident in the vignettes, several participants noted that effort was a determining factor. Lucy and Kevin explicitly stated the role of student effort in their understanding of fair classroom assessment, and a perceived lack of student effort was involved in some of the unfair assessment situations that the participants recounted from their practices. Overall, the participants seemed to view assessment accommodations and modifications as fair as long as they deemed the student to be clearly worthy or in need of different treatment. The challenge for fairness in classroom assessment, then, is that it depends on the teachers' understanding of students, and on their reflection in determining what might be in students' best interest.

### Elements of Practical Wisdom at Play

The participants in this study were asked to articulate how their practical wisdom informed their recommendations for fair classroom assessment. First they indicated which elements of practical wisdom they used in responding to each vignette, and then we discussed these elements more generally in relation to fair classroom assessment during the interviews. In this section, I report on the elements of practical wisdom that I observed in their responses, as well as those that the participants explicitly identified. I have organized this section to

show, as clearly as possible, how the different elements of practical wisdom came into play across the cases. However, it should be kept in mind that these elements are intertwined in individual networks, and they are used to varying degrees and in different combinations by each participant in response to different situations.

### *Knowledge about Learners, Content and Pedagogy*

The knowledge that participants expressed in relation to fair classroom assessment basically related to three areas: learning and learners, content, and pedagogy. All of the participants in this study had completed a teacher education program and taken professional qualifications courses in which they would have gained theoretical knowledge about learning and learners. They all drew on this type of knowledge in their responses to almost all the vignettes (A, B, D, E, F, G, H), but the basis of this knowledge appeared to be more practical than theoretical. In responding to Vignette F, for example, Lorena, Kevin, Lucy and York all stated explicitly that their knowledge about gifted students was drawn from their teaching experience, and only Tanar identified special education courses as his source. Often their recommendations for fair assessment were based on general knowledge about students (e.g., type of feedback preferred by students) that had been gathered through years of working with particular groups (e.g., grades, types) of students. York sums up this type of knowledge as knowing where students are “coming from, where they’re at” and “what works for students.” While this knowledge was sometimes offered simply as a matter of good practice, it appeared to be central in how the participants avoided unfair assessment situations. Lucy seemed to be especially adept in this way. For example, in reference to establishing the constructive environment needed for fair peer assessment, she stated, “I know the way they think, and that

helps a lot.” Comments that involved general knowledge about learners were often supported with examples of particular students the participants had taught. In discussing fairness in classroom assessment, then, the participants usually referred to general knowledge about learning and learners that was gained through experience rather than theoretical study.

In addition to their professional qualifications, all of the participants had completed an undergraduate degree in the Arts in which they would have acquired knowledge about English Writing as a subject. However, of the different types of knowledge they expressed, content knowledge was the least evident. Vignette B and G, which both illustrated a situation with a teacher preferring a certain writing style, generated a few references to content knowledge. Lorena, for example, voiced a fundamental principle in the study of literature and composition in noting that “in art some people prefer cubism to impressionism - neither are wrong. The same applies to writing - as long as the point is made and the argument stands with evidence.”

Although content knowledge was not frequently heard, when it was expressed, it was often intertwined with pedagogical knowledge as the participants explained their thoughts about how to teach and assess writing. All of the participants expressed this type of knowledge in at least one response. Amada offered pedagogical content knowledge more than the other participants, with comments such as “students become better writers by writing, errors made provide teachable moments” and “put a better writer with a weaker writer, it works.” However, Amada also stressed the importance of content knowledge as a precursor for assessment, noting that teachers’ need a strong grasp of the content matter “because you have to diagnose . . . student errors . . . there are x number of things you think

make perfect writing, so you look at all the kids you have and zero in on the things that they need to improve.” Explaining how she developed her knowledge about assessing student writing, Amada said that it took “awhile, maybe within the first five or ten years, and then as I went along teaching it just got more and more refined.” Amada’s combining of knowledge about content matter and pedagogy illustrates how theoretical and technical knowledge work together as part of a teachers’ practical wisdom. Her comments also suggest that this interworking of different types of knowledge evolves with experience to become the pedagogical content knowledge that gives teachers confidence in assessing student writing.

Of all the elements of practical wisdom that the participants drew on in discussing fair classroom assessment, pedagogical knowledge was heard most frequently. Very often it was expressed without reference to the content matter, and focused specifically on assessment methods or strategies. While pedagogical knowledge informed the participants’ responses to almost all the vignettes, it was most evident for Vignette C, which was about sharing a rubric with students for a cross-curricular unit. There the participants’ familiarity with rubrics, and the emphasis on the importance of clear criteria in their recommendations suggested that they were aware of current classroom assessment principles. Often this type of assessment knowledge was stated in the abstract, such as York’s recommendation that “criteria should be clear and helpful so students can demonstrate their learning.” The participants’ responses this vignette and others suggest that they valued knowledge about classroom assessment, and several explicitly identified teachers’ assessment literacy as a factor that could influence the fairness of an assessment. When they were asked about their assessment literacy, though, participants did not refer to any specific theories or authors.

York and Tanar, who were particularly familiar with current assessment ideals, were the only two who indicated how they had acquired their knowledge about classroom assessment. York identified professional development workshops and assessment policy documents as the sources for his assessment knowledge. Tanar thought that he'd probably heard about assessment for learning in a workshop, but his interest in meta-cognition and student self-assessment seemed to stem from his experience writing curriculum documents for his school board and the OME.

### *Teaching Experience*

The participants in this study had between 13 and 40 years of teaching experience, most of which involved teaching English Language Arts in Grades 7 to 12 in Ontario's publicly funded educational system. This wealth of experience permeated the practical wisdom they offered about fairness in classroom assessment, with each of the eight vignettes eliciting references to teaching experience from at least three participants. Frequently, their experience as teachers gave the participants insight into what to expect, or what would work in different situations. A theme that emerged as they discussed their teaching experience was the notion of pre-mandated practice, where several of the participants noted that something they had done on their own later become a desirable practice. Amada explained that she developed variety in her teaching strategies to match the variety she encountered in students' learning styles, and that she was "doing that stuff before they actually gave it a name." A practical need also influenced Tanar's early adoption of a strategy that is now recommended. He explained that he was "doing conferencing with kids way before it become fashionable because I had to - they simply could not read what I was writing so it was sheer dire

necessity.” With his elementary experience, York had been using rubrics and involving students in developing assessment criteria for many years, which is a practice that is now recommended in the revised secondary English curriculum documents.

It’s interesting because special education is finally catching up with what we’ve been doing in elementary since I started teaching . . . I noticed more workshops are talking about it – almost as if it’s new – this is what you should be doing. And, not to sound arrogant or anything, but elementary teachers have been doing a lot of these best practices for years.

In discussing how teachers’ develop practical wisdom, Lorena made a similar observation, but rather than practical need or previous experience, she felt that some practices were natural, or intuitive for teachers.

I’m constantly surprised, even after this many years of teaching, that a lot of things I do innately, and then find out that there’s either research backing it up . . . or it’s called something . . . and wow, this is something I feel.

While these comments could simply be a reflection of the circular relationship between research, policy and practice, they were made in reference to assessment practices that had been tried by the participants as a means of supporting student learning.

When they discussed fairness in classroom assessment, the participants often drew on examples from their teaching experience, but most did not explicitly link their experience with making fair decisions. Amada was an exception in making this connection in her response to a question about whether she faced fairness issues or dilemmas similar to the vignettes in her practice.

When I first started out teaching that was much more frequent. In my latter stages of teaching, ‘Oh, what do I do here?’ - a split second of deliberation, and I have an answer. But when I first started out, it would take hours. Sometimes I would even go to sleep thinking about the problem. It’s a journey. . . there’s a lot of trial and error, and mistakes, and wailing and gnashing of teeth, and wrecks, and what not. But then after a certain number

of years you get to a comfortable zone where you don't even think any more. It's just so internalized that it occurs.

Amada's words might give hope to less experienced teachers who struggle with assessment dilemmas, but they also raise a question about teachers' ongoing reflection in practice. While new teachers do not have the breadth or variety of classroom experiences to draw on for fair assessment, very experienced teachers may have repeated similar experiences to the point that they stop reflecting on fairness issues. However, this was not evident in this study. What was clear is that teaching experience informed teachers' practical wisdom about fairness to a great degree, but it was not the only element at play.

### *Personal Characteristics*

Gathering information from participants on three occasions did not permit the familiarity that would be necessary to fully understand their characters. However, the information they volunteered about themselves provided some evidence about their personal characteristics, and the influence of character on thinking about fairness. In particular, some of the participants referred to roles outside of the classroom, their cultural backgrounds, and their own experiences as learners in the process of discussing fair classroom assessment.

Kevin noted that becoming a parent had changed how he felt he should treat his students. Lorena also referred to parenting, first when discussing the fairness of differentiated assessment, and a second time when she explained the professional responsibility that she assumed in assessing student learning. Although Amada did mention her family when she explained her beliefs about learning languages in Canada, her recommendations for fair classroom assessment seemed to be more influenced by her leadership roles. Having worked with much younger teachers before she retired, Amada emphasized the importance of

professional development for fair assessment, and she was also the only participant who mentioned the need for teachers to understand educational theory.

Five of the six participants mentioned their own cultural background in discussing classroom assessment. Tanar described his parents as Hungarian immigrants, Lorena referred to being an Italian Catholic mother, York mentioned his schooling in England, Lucy remembered becoming a teacher in South Africa, and Amada talked about attending school in Jamaica. In Tanar and York's cases, cultural background in itself didn't seem to influence their understanding of fairness. They referred to it only in situating a story about an unfair assessment that they had experienced as a learner, and that they did not wish to reproduce as a teacher. As such, cultural background indirectly influenced their practical wisdoms about fairness in classroom assessment, via their experiences as learners. For Lucy and Amada, though, the link between cultural background and practical wisdom about fairness was more direct. For example, in explaining why she felt that it was fair to expect students' best effort, Lucy stated, "that's my South African background. It's a very rigid system there." Both Lucy and Amada had completed their secondary schooling at a time and in educational systems where high standards for English Language Arts prevailed, and they both referred to the influence of this experience on their teaching and assessment practices on several occasions.

Despite their varied backgrounds, a commonality that was clearly evident as a characteristic across all six participants was their personal commitment to learning. This was evidenced not only in their higher education, but also in their extensive engagement in professional development through workshops, conferences, and additional qualifications courses. All of the participants were sufficiently confident in their practices, and concerned

about teaching as a profession to have volunteered on a teachers' network to mentor other teachers and share their practical wisdom.

### *Moral Beliefs and Ethical Knowledge*

Moral beliefs and ethical knowledge are similar, and they may even involve the same issues. The difference between them is in how they are held and expressed, with moral beliefs held personally and expressed individually, and ethical knowledge held collectively and expressed professionally. All of the vignettes generated at least one response that drew upon a moral belief or ethical knowledge, and some elicited both. Vignette A, which was about a teacher deciding whether to provide inattentive students with an additional learning opportunity, provides a good example. York related that he had found himself in a similar situation as a teacher, where he countered the temptation to “punish [a] child for not being attentive” with the belief that it was “not right.” In contrast, Lorena and Amada responded to the same vignette in terms of their professional obligation to support student learning, which expressed their sense of the ethics of practice, rather than their personal morality.

One vignette in particular, Vignette H, elicited far more responses involving moral beliefs and ethical knowledge than the other vignettes. It involved a teacher deciding what to do about the incomplete work in a student's portfolio. Fairness for the class was given precedence over fairness for the individual in this situation, which seemed to be because the participants identified the issue as one of work ethic rather than academic need. Amada, Lorena, Kevin, Lucy and Tanar all thought that the student should complete all the work, and Lucy explained her reasoning for this.

The student cannot be allowed to do less simply because he has the ability. It sets a very poor example for the rest of the students, and they are very aware

of students who get away with doing less. They resent it.

Kevin expressed a similar idea, but he also suggested that the teacher was responsible for taking action, through phone calls home and a penalty system, to ensure that the student completed all the work.

The moral beliefs and ethical knowledge underlying the participants' responses fell essentially into two categories. The first related to work ethic, or the belief in the value of hard work. Kevin's focus in this area was on completion; he felt strongly about all students completing all assignments. Lucy's interpretation of hard work was somewhat different in that she emphasized the need for intellectual challenge. Both associated work ethic with fair assessment, but for Lucy it involved the quality as well as the quantity of work completed. The second category of moral belief underlying participants' recommendations for fair assessment related to the quality of classroom relationships, or how teachers and students treat each other. All of the participants discussed the importance of respectful and constructive relationships in the classroom for supporting student learning and fair assessment. Lorena emphasized this most frequently with statements about her role and the need for trust, such as "I am not in the classroom to ambush or 'catch' students for lack of knowledge" and "my role as a teacher is to encourage excellence. If I take on a punitive role, then I might discourage the student from further developing their skills." For Lorena, the nature of the teacher-student relationship was clearly a matter of professional responsibility that required explicit attention. In her response to Vignette H, for example, she wrote that "trust, relationship, support are important factors to foster and communicate to the student." Tanar also stressed the teacher's role in nurturing respectful and constructive relationships,

and he felt that teachers needed to be proactive, through modeling and direct communication, in how peer relationships developed. He explained that teachers should help students “realize that language has power, and it unfortunately sometimes has the power to upset others.” In essence, while the participants all seemed to hold similar moral beliefs, particularly about the value of hard work and the importance of treating others well, they diverged in expressing them in personal terms or as professional ethics.

#### *Understanding of Particular Contexts, Situations and Students*

In discussing fairness in classroom assessment, the participants drew on their understanding of particular contexts, situations, and students to varying degrees. Changes in the Ontario educational system came up during all of the participants’ interviews, either in reference to a political figure, a key date, or particular policies. The influence of these are discussed further in the next section on external factors, but what is important to note here is that the participants did refer to their understanding of the educational system when they discussed fairness in classroom assessment. Their understanding of the particular school environment in which they worked was also evident in their responses. For example, Kevin explained that when military families arrive from other provinces, their children are faced with new learning expectations, and they are not always prepared for EQAO’s standardized assessments. He felt that it wasn’t “really fair” to expect secondary students to take the OSSLT when they haven’t studied the same English curriculum. The role that the participants’ understanding of particular environments played in their practical wisdoms about fairness is especially evident in contrasting their interpretations. For example, with her understanding of teaching English to gifted students in an advanced program, Lucy

associated fairness with high standards and intellectual challenge, whereas York's understanding of teaching in a "high needs" program meant that he saw fairness as helping less academically successful students meet the requirements for an OSSD.

The role of understanding particular classroom situations in the participants' practical wisdoms about fairness was clearly seen in their responses to the nine items on the ethics of classroom assessment that were included in the Background Questionnaire (see Table 4). The participants initially selected the 'depends on circumstances' option for a third of their responses. The only item that none of the participants thought would depend on circumstances was the one about a student receiving a zero for not returning a form with a parents' signature. During the interview, the participants elaborated on the circumstances that would influence the ethics of different items, and in some cases they thought of particular situations that resulted in a change of response. Amada and Kevin, who both indicated on the questionnaire that changing a students' grade would be unethical, later described specific situations where it would be ethical. Lorena was especially sensitive to the particulars of circumstance, and she explicitly stated at several points that the ethics "would depend," and that she would "really need to know what the situation" was in order to decide whether an assessment practice was ethical or unethical.

All of the participants gave examples that involved particular students when they discussed their practical wisdom about different aspects of fairness. Kevin and Lucy both related short stories about their interactions with particular students by repeating the dialogue as they remembered it, and giving the students voices. Most of the participants also used examples of specific students when they discussed differentiating assessment. Amada and

Lorena were the most direct in linking fair assessment with understanding particular students. Amada saw this as part of being a teacher, and she explained that her “relationship with each student involves knowing their capabilities.” Lorena stressed this element of practical wisdom in her recommendations.

Part of the fairness thing to me is to know the students. You have to know who each of the learners are, so by knowing who the students are I can take those things into consideration. The idea is to understand who all the students are, and having them strive for excellence. Everybody’s striving for their own excellence.

Understanding particular students played an important role in Lorena’s practical wisdom about fairness, and for many of the other participants it figured into their recommendations for fair assessment, especially if there was any possibility of differentiated assessment.

In summary, all of the elements of practical wisdom were used by all of the participants to discuss fairness in classroom assessment. There were differences in the degree to which participants used the different elements in their responses, but this may be simply a matter of expression. Lucy and Amada, for example, both had four decades of teaching experience to draw on, but they expressed their practical wisdom differently. Lucy referred more frequently to her knowledge about learners and pedagogy, while Amada shared her content knowledge more often, and she was more explicit in including her experiences as a teacher and a learner. Amada and Lucy’s wealth of teaching experience did seem to advantage them in that they were both very articulate and confident in sharing their practical wisdom, whereas those with less experience were sometimes more hesitant or less fluid in their expression. For example, Lorena expressed uncertainty about topics that fell outside of her own teaching experience (e.g., multiple choice tests). Kevin initially struggled when he

tried to articulate his ideas about how structure and routine made classroom assessment fairer, but he did subsequently make his point. The interview became, for other participants as well as Kevin, a process of working through ideas about fairness in order to articulate their practical wisdom. All of the participants also commented on how different elements of practical wisdom worked together as a “combination” (Kevin), a “blend” (Lorena), or an “amalgam” (Amada), and they seemed to agree that it developed over time with education and teaching experience. Overall, both the practical wisdom that the participants shared in their recommendations for fair classroom assessment, and their thoughts on how they came to these recommendations, appeared to be informed by a dynamic inter-working of various forms knowledge, understandings, beliefs and characteristics that continued to develop through learning and experience.

#### External Factors of Influence

One of my interests in this study was to determine which external factors might influence the fairness of writing assessment in English Language Arts in Grades 7 to 12 in Ontario. The participants were first asked to identify which factors would make it easier or more difficult for a teacher to follow their recommendations for fair assessment in response to each vignette, and then during the interviews, we discussed these and any other external factors that they felt might influence fairness. In this section, I have organized the participants’ responses according to the concentric levels used earlier, beginning with those that are closest to the practice of assessment for learning.

#### *Classroom Level Factors*

The participants discussed four factors at the classroom level that could impact the

fairness of classroom assessment: class size, student ability, student behaviour and parents. All the participants identified class size at least once, but none discussed it at great length. Kevin identified class size most often, and he saw three ways in which it could impact the fairness of classroom assessment. He explained that he'd had large classes (35 students) "way back when," and now that his classes were all smaller (11 to 24 students), he could interact one-on-one or with small groups more "effectively" rather than focusing on "keeping them inline, and making sure everybody's safe and working." Kevin felt that it was also easier to "keep track" of fewer students during peer-assessment to ensure that their feedback was constructive, and he also noted that smaller class sizes made individual writing conferences feasible and less time-consuming. This last comment was made in reference to Vignette E, which was about a teacher writing feedback on students' assignments. It generated the most information about class size from the participants, and several described strategies that they had tried to manage feedback for a class. York and Kevin both recommended giving more oral feedback, and Amada described a systematic process that involved going through a class set one criterion at a time. Tanar related that he'd once tried giving each of the 37 students in his Writer's Craft class individual feedback recorded on cassettes, but it ended up taking even more time than usual, so he did not recommend it for other teachers. Basically, smaller class sizes allow classroom assessment to be tailored more specifically to the needs of each learner. As York noted in his program, "our class sizes are small, so we can do a lot of individual accommodation." In brief, the participants identified class size as an influence on the fairness of classroom assessment, but they seemed to accept it more as a matter of course than concern.

Students were identified more often as a factor in fair assessment than any other factor at any level. All of the vignettes elicited at least a few responses about students, and Vignettes A and D generated recommendations relating to student behaviour or attitudes from all of the participants. Although they all agreed about the need to provide students with the opportunity to learn for Vignette A, most indicated that the behaviour of the two inattentive students could hinder a teacher's ability to implement their recommendations for fair assessment. For Vignette D, where the peer-assessment process was less than constructive, all six participants mentioned either the behaviour of individual students, the social dynamics between students, or the maturity of the group as a factor for fair assessment.

Despite the frequency of references to students' behaviour as a factor of influence, negative comments about students were rare, and they tended to relate to the challenges that some participants' faced, such as absenteeism in York's program, or attitudes about academic work in Kevin's applied classes. Most often, the participants' recommendations regarding the link between student behaviour and fair assessment focused on the need to be proactive in establishing expectations and communicating with students. For example, Lucy explained that she keeps a "sharp eye" on the process of peer-assessment, which usually results in students being "very productive." The importance of the student-teacher relationship was also mentioned repeatedly by the participants, particularly Lorena, Tanar and York. When asked directly if student behaviour could impact a teacher's ability to assess fairly, Tanar replied, "yes, I think that's inevitable, unfortunately . . . we try not to, and I certainly try not to let a personal clash influence assessment, but . . . sometimes it can't be helped." In these situations, and overall in discussing student behaviour as a factor in fair assessment, the

participants emphasized the teachers' professional responsibility for creating and maintaining a constructive classroom environment.

The participants also identified students' academic ability as a factor that could influence the fairness of classroom assessment, but to a lesser degree than student behaviour. The relationship between students' academic ability and fairness came up in reference to specific assessment methods with certain groups of students. Lorena, for instance, noted that younger students who use rubrics for the first time "need support in understanding the expectations." Amada felt that self-assessment with ESL/ELL students was a "different ball game" than with students who "grew up in the system," and she explained that "it takes longer" because ESL/ELL students need to gain confidence in their English language ability before they can assess their writing. Several participants also identified the range of students' academic abilities in one class as a factor that might hinder fair assessment, but Kevin noted that it could also be the reverse. He explained that "if students do have different abilities it can often be a good thing as some students just need a little coaching and they're back on track, enabling the teacher to spend more time with those who need it." In discussing students' academic abilities, concern tended to center on fairness for students in the lower range, which is evidenced in Lorena's comment.

If something happens that this child is not keeping up within the range of the class, and my grade 7 classes would run from . . . a grade 5 to a grade 8 . . . and I'm in an inner city school, so there was a wide range in there. But if they're falling off the end scale . . . I've got to find out why. And to be fair in my assessment, I need to do that.

In keeping with the views they expressed about students' behaviour, the participants went beyond identifying students' abilities as one factor, and their responses tended to concentrate

on their responsibilities for ensuring what they felt would be a fair assessment process.

Parents were identified as a factor that could be influential, but they garnered very little of the participants' attention. Amada and Lorena commented on the pressure that some parents place on their children, but neither explicitly related this to the fairness of classroom assessment. Parents figured into Kevin's responses more often than the other participants, and he wrote that he was a "very firm believer in keeping parents in the loop." He explained that he used several methods to communicate with parents, including an "electronic bulletin board" that he had created. In some comments, Kevin expressed disappointment about the lack of reinforcement from parents regarding students' work habits, but the link with fairness seemed tenuous. However, in the recommendations he offered explicitly for fair assessment, he stressed the importance of clearly communicating with parents, indicating that he thought the classroom assessment process should be transparent for parents as well as students

#### *School Level Factors*

Two factors at the school level that influence teachers' assessment practices are colleagues and school leaders. The participants in this study discussed colleagues more frequently than school leaders in relation to the fairness of classroom assessment. All of the participants identified colleagues as a factor for either Vignette B or C, or both. For Vignette C, where a group of teachers developed a rubric for a cross-curricular unit, Kevin, York and Tanar felt that the teacher's ability to follow their recommendations for fair assessment might be impacted by her colleagues' ideas. As Kevin noted, colleagues could "either make it much easier or much more difficult, depending on who your colleagues are." For Vignette B, where a teacher and his colleague's views about writing style differed, no other external factor was

identified more frequently than colleagues. While some participants felt that differences in teaching styles could result in conflicts, they also recognized that the teacher might find it beneficial to listen to his colleague. For example, in his recommendation for fair assessment, York suggested that the teacher “consider his colleague’s opinion because it is good to check our own assumptions and habits.” This resonates with the positive view of collegial support heard at other points in the study. Kevin, Tanar and Lucy all mentioned how much they had learned from mentor teachers that they had encountered in the early years of their careers. Amada’s responses differed on this topic in that she focused on the later years of her career when she was working with much younger teachers, and she took the mentor role.

A difference relating to colleagues and classroom assessment also emerged in the responses to Vignette B. Amada and Kevin both felt that it was better not to be too sensitive to colleagues’ opinions when assessing student writing, and they basically recommended that the teacher be confident in his own assessment. Kevin felt that there were “too many factors” that could “affect a student’s work from one class to the next or one year to the next” for a teacher to worry about how other teachers had previously assessed that student. Amada explained that she had developed her own confidence because “back in her day” as a new teacher, you could turn to another teacher for “a bit of advice” but there wasn’t “a lot of collaboration.” While Amada felt that new teachers should have “good mentors,” she preferred, like Kevin, to maintain autonomy in her assessment practices. In contrast, Lucy worked in an English department that had agreed to assess by levels with “very clear descriptors for each level.” Lucy felt that this would improve the fairness of classroom assessment in her school because there was more “likely to be consistency in terms of

consensus” among the English teachers. Of all the participants, Lorena placed the most emphasis on the influence of colleagues in fair classroom assessment. She mentioned colleagues repeatedly, and she felt that the collaborative assessment (i.e., teacher moderation) process used across a grade level at her school did “bring out more fairness” for the students.

Although principals are sometimes involved in the type of assessment initiatives described by Lucy and Lorena, school leaders were not mentioned very frequently by the participants in relation to fair assessment. Lorena noted the importance of a supportive school culture for fair assessment in her responses to several vignettes, and during the interview she explained, “I don’t think that would impede fairness in my world right now, but I can see that if there was a lot of friction between the administration and the teacher that might get in the way of fairness.” Kevin and Tanar had recently had negative experiences where their assessment decisions had not been supported by their principals, and they both identified administrative support in their responses about the factors that could impact fair assessment. However, they both also saw that the issue went beyond the school level. Tanar felt that the outcome in the situation he described “besmirches the integrity of what we’re trying to do as a system,” whereas Kevin conceded that his principal’s decision was consistent with the OME’s policies. In both cases, a problem in assessing learning at the classroom level was resolved at the school level by drawing on policies established for the educational system.

### *System Level Factors*

The participants identified a variety of system level factors that could influence fair classroom assessment. Report card requirements, standardized testing and teaching traditions were all mentioned by at least one participant, but on the whole, they were much more

concerned about assessment policies, professional development opportunities, and time.

Several different comments about assessment policies were generated by Vignette H, which described a teacher wondering how to deal with missing assignments in a student's writing portfolio. Tanar indicated that "adherence to ministry guidelines" would permit a teacher to follow his recommendation for fair assessment, whereas Amada noted that "personal beliefs in assessment policies" would be at play. Kevin justified his belief in the importance of work completion by referring to several OME assessment policies, explaining that "by collecting all assignments, I am able to assess many of the ways that we are encouraged to including: most recent work, most consistent work, and offering more than one chance to demonstrate understanding." While Kevin thought that these policies would be constructive, especially if a student had "started off rough and . . . gotten a little bit better further along," he also voiced disagreement on several occasions with the OME assessment policies on late and incomplete work. All of the participants expressed disagreement at some point in their responses with at least one policy, either at the local (i.e., school or school board) or provincial level. For the most part, though, the participants seemed positive about the impact of newer assessment policies. In describing how teaching had changed over her many years of classroom experience, Lucy noted that the OME achievement chart made assessing English essays less laborious.

I'll tell you what else has changed. I think the fact that we now use the four levels . . . so you just do a holistic mark. You just look at it, and say this is a level this, a level that. And then they're marked, and you know well enough what the level is. I was painstakingly saying three out of five, seven out of ten.

York was particularly enthusiastic about a current assessment document produced by his school board because he felt it was constructively oriented to student learning.

In contrast, a theme that emerged across the participants' responses related to the politics of educational change, and the participants' references to past changes in Ontario's educational system tended to be negative. Lorena began a story about a friend not liking the standardized report card with "you know when the government was, in '97 when the Torries had taken over and everything changed," and Amada explained that she retired as a program coordinator because "Mike Harris came into power and he cut all the funding." Some of the participants' comments suggested that their years of experience contributed to their perspective on change in the educational system. For example, Amada noted that as a teacher in Ontario, she'd "gone through . . . six, eight different ministers with their whatever they wanted to do in schools," and in discussing the English curriculum, Tanar explained that the "latest documents are to a certain extent about putting back into the curriculum what was taken out in the first set of revisions." While their comments suggest a sense of cynicism about change, it is notable that the negativity expressed was consistently in reference to the previous political regime in Ontario. None of the participants referred to the politicians or political parties in power, or the educational changes underway at the time of this study.

A second system level factor seen in the participants' responses was the need for professional development opportunities. With a few exceptions, the participants did not recommend professional development in explicit terms. Instead, they specified the type of knowledge that they felt the vignette teacher needed to assess fairly. While the vignettes contained a range of assessment interactions, the type of knowledge recommended did not seem to depend on the content of the vignette, and several participants clearly favoured certain types of knowledge. For example, Amada felt that knowledge about the content

matter and about classroom assessment would help the vignette teachers, whereas Lucy thought that knowledge about learners would be useful. Vignette D generated the most responses relating to the teachers' knowledge, including the need for knowledge about learners and pedagogical content knowledge about how to teach and assess writing.

However, Kevin was the only participant to explicitly recommend professional development for this vignette. He suggested that teachers should learn how to "set up their classes for success by having firm rules in place" to avoid problems and facilitate fair peer assessment.

Although the participants discussed their mentors in glowing terms, and they were generally positive about their colleagues, negative references to other teachers ran through their responses. This is not surprising given that some of these responses were elicited with vignettes that describe teachers who were poised at a moment of professional uncertainty, but the frequency and strength of feeling about other teachers' lack of competence warrants mention. For example, Kevin wrote that other teachers "stick too rigidly to unit activities," they "put too much value into labels," and they "don't focus enough on the skills that students need" in his responses to several vignettes. More often, though, it was when the participants were explaining their own practical wisdom about fairness in classroom assessment that they voiced criticism of other teachers. In justifying her own high standards, for example, Lucy noted that when other teachers are insufficiently demanding of their students, learning is devalued and the fairness of assessment is compromised. And after having discussed the need for multiple assessments, York explained that assessment was less fair when teachers equated it with testing and did not realize the variety of ways that student learning could be assessed. Lucy and Lorena were both particularly disdainful of other

teachers who used the same lessons year after year, and did not continue to develop as teachers. Lorena stressed that “teachers cannot be stagnant” and she cautioned that while teachers should not “jump on every bandwagon” they should “be aware of trends and changes.” This echoes the tension that Amada identified in teachers’ development when she explained that teachers needed to maintain their own core beliefs while keeping up with the pace of educational change.

The third system level factor that was evident in the participants’ responses was time. Many of the comments about time involved the number of students whose writing was being assessed, which was discussed as class size. Additionally, the issue of time was also raised in relation to the curriculum and course length. Kevin wrote about the push “to get through” the curriculum in several responses, and he identified one of the consequences of this pressure.

Sometimes there is simply so much that teachers feel needs to be covered that you really are pushed for time to get things in. This means that you can breeze through a topic and have the students not understand it as well as you want them to, or you really feel like you don’t have the time to go back and re-teach a lesson or a group of lessons.

Kevin’s comment suggests that time constraints can reduce a teacher’s use of assessment information to adjust teaching plans for the additional opportunities students might need to learn. York also discussed time and curriculum in tandem, and he saw that it could hinder the opportunities students might receive to demonstrate learning. Several comments by Kevin, York, Lucy and Amada suggested that time influenced the type of assessment that they selected or developed for their students, with tools and activities that consumed less time for both students and teachers necessarily being favoured.

The participants’ identification of time as a factor of influence on fairness in

classroom assessment seemed to have some relationship with their teaching assignments. The two participants who did not mention time at all were both teacher-librarians, and they had reduced teaching loads as a result. In contrast, York, who taught in an intensive ten-week program, and Amada, who taught large classes in an adult ESL program, discussed time constraints the most. In brief, time was identified by most of the participants as a factor that could hinder fair assessment. While they dealt with time constraints by using the assessment methods that they felt were most effective for the time frame of a course, their comments also suggested that aspects of assessment for learning, such as the adjustment of teaching based on assessment information and the use of multiple opportunities to elicit learning, require time that was not always available.

In summary, the participants identified a range of external factors that could help or hinder a teacher's ability to follow their recommendations for fair assessment. They were, by far, most aware of the influence of students' abilities and behaviour, but they also discussed the benefits of collegial support and the need for professional development opportunities. System policies were identified as a facilitator as well as a hindrance for fair assessment, whereas time was always discussed in the deficit. Class size and school leaders were also identified by several participants, but neither were major sources of discussion, and their influence seemed to be accepted as part of the standard operations in the Ontario's educational system.

## CHAPTER EIGHT: DISCUSSION

*"We do not learn from experience; we learn by thinking about our experience."*

Shulman (1996/2004, p.474)

In this chapter, I discuss the results of my study in relation to the literature and the concepts in my framework in order to highlight how my work contributes to existing knowledge. In organizing my discussion, I followed the same tripartite structure used to report the results, with the three following sections relating to fairness in classroom assessment, teachers' practical wisdom about fairness, and external factors that influence fair classroom assessment

### Fairness in Classroom Assessment

In sharing their practical wisdoms, the participants discussed many different aspects of fairness, and these could be taken up in a variety of ways. Because I am interested in fairness as a quality of assessment for learning, I discuss the six aspects of fairness highlighted in the cross-case results specifically in relation to this purpose of classroom assessment. As described earlier (see Chapter 2), assessment for learning involves specific practices (actions) in a dual cycle of planned and spontaneous interactions (Cowie & Bell, 1999; Tierney & Charland, 2007). In theory, assessment for learning is an ongoing spiral of interactions between a teacher and a student, but in the classroom environment there are necessarily starts and stops for specific periods of learning (e.g., secondary semester). In the following sections, I discuss aspects of fairness in an order that makes sense within one cycle of this spiral. I am not suggesting with this organization that assessment for learning is consistently manifested in exactly this way, nor am I prescribing a fixed set of steps for fair

classroom assessment. My purpose here is to situate the six aspects of fairness that emerged prominently in my study in relation to the theoretical process of assessment for learning.

### *Constructive Learning Environment*

With the recognition of the social nature of learning, the importance of a constructive environment for assessment to support learning has been increasingly emphasized in classroom assessment literature (e.g., Cowie, 2005; Pryor & Crossouard, 2008; Shepard, 2006). This is reflected in one of the standards associated with fairness in *The Student Evaluation Standards* (JCSEE, 2003), which states that “student evaluations should be constructive, so that they result in educational decisions that are in the best interest of the student” (p.67). In these standards, a constructive environment is described as one in which balanced assessment information (i.e., strengths and weaknesses) is used in a timely manner to support learning. The participants in my study recommended that teachers be proactive in establishing and maintaining a constructive learning environment as the setting for fair assessment. However, the strategies they suggested went beyond the use of balanced assessment information, which is essentially a technical concern, and they tended to focus on the quality of the human relationships within the classroom. Their references to trust suggest that their concerns about fair classroom assessment were ethical as well as technical. In the ethical standards for teachers in Ontario, the ethical standard of Trust “embodies fairness” (OCT, 2008a, p.9). In an analysis of students’ perceptions about classroom assessment in New Zealand, Cowie (2005) found that the students needed to “feel comfortable with a teacher before they disclosed their ideas by asking questions” (p.206). She concludes that the development of trust and respect in the classroom environment is central to students’

willingness to engage in assessment for learning.

Respect appears in another standard for classroom assessment that is associated with fairness, which states that “students should be treated with respect in all aspects of the evaluation process, so that their dignity and opportunities for educational development are enhanced” (JCSEE, p.45; similar to Airasian, 2005). Not only were many of the strategies that the participants recommended, such as speaking privately to students about negative behaviour or monitoring peer assessment activities, similar to the guidelines given for this standard, but they also suggest that the participants were taking the social nature of assessment for learning into account in their thinking about fairness. In some of their recommendations, the relationship between fair assessment and assessment for learning seemed to be symbiotic, meaning that *fair* and *for learning* are not the same, but they can be mutually supportive. For example, one participant’s suggestion for fair assessment involved teaching students about constructive peer feedback both directly and through modeling. This is similar to one of the practices recommended by Camilli (2006) for fair classroom assessment. It could also be described as a form of meta-discourse that includes fairness in framing the assessment for learning environment (Pryor & Crossouard, 2008). Essentially, then, a constructive learning environment is needed for fair assessment, and fairness in classroom interactions helps ensure that the learning environment is constructive.

#### *Relevant Opportunities with Purpose in Mind*

In classroom assessment theory, both the need for students to have opportunities to learn and opportunities to demonstrate learning are identified as aspects of fair classroom assessment. Although the provision of equal learning opportunity for all students involves

social and economic factors beyond a teacher's control (Weeden et al., 2002), opportunities relating specifically to the learning expectations for a course should be planned. Textbooks on classroom assessment tend to define opportunity to learn more narrowly in these terms, suggesting that the alignment between teaching and assessment contributes to the fairness of an assessment (e.g., Airasian, 2005; McMillan, 2007). When the participants in my study were presented with a situation that involved students who did not understand an assignment, they all recommended that the students be offered additional opportunity to learn. Unlike some of the teachers in the existing research on classroom assessment (e.g., Eggen, 2004; Gummer, 2000), they did not seem concerned about all the students in the class having exactly the same opportunity to learn, and their responses focused on providing learning opportunities according to need. As such, they were more concerned with equity than equal treatment in providing students with opportunities to learn.

The participants' recommendations for fair assessment related more frequently to providing students with opportunities to demonstrate learning. Using multiple and varied methods to elicit learning is one of the core practices of assessment for learning, but some of the participants' examples also involved culminating assignments that were intended for the assessment of learning. Two different interpretations of providing opportunity to demonstrate learning were evident in their examples. The first involved the use of multiple measures, which seems to be common practice among secondary English teachers in Ontario. Of those who responded to the 2002 SAIP questionnaire, for example, almost all (92 %) reported using at least five different pieces of information to compute students' final grades (CMEC, 2002a). In keeping with this, several of the participants in my study suggested that to be fair,

assessment decisions should be well informed to provide as complete a picture of student learning as possible. This interpretation of fairness is akin to the quality of reliability in classroom assessment theory, which involves ensuring a sufficiency of information for assessment decisions (Brookhart, 2003; Smith, 2003). There are many uses of assessment where it is evident that a sufficiency of information contributes to the fairness of a decision, such as in the selection or certification of students. As such, when an assessment is for purposes other than student learning, reliability and fairness are clearly related.

The relationship between reliability and fairness is not as evident when the primary purpose of assessment is to support learning. In dynamics of the classroom, where student knowledge is ideally changing, teachers often have to respond without the completeness of information that is ideal in the assessment of learning. Morgan and Watson (2002) argue that in this context, the fairness of teachers' decisions is threatened not only by the "necessary incompleteness" (p.104) of their information, but also when they do not question their initial impressions or the reasons for students' results. It may be that teachers can counter the first, the issue of incompleteness, by attending to the second, that is, by continuing to question what they do know about students. This necessitates a view of assessment information as impermanent. In describing assessment as being part of the teaching-learning cycle, Amada suggested that for the process to be fair, learning expectations should be continuously modified as the teacher comes to know and respond to the student's learning. Bulterman-Bos and colleagues (2002) make a similar point in their analysis of teachers' stories about classroom observation, where they noted that "new bits and pieces of information are continually captured, and old bits are lost (forgotten) or replaced" (p.1086). This suggests

that to make fair decisions in the momentum of assessment for learning, teachers need to update their understanding of students learning frequently, as opposed to relying on first impressions or infrequent assessments. Considering Amada's interpretation of fairness in tandem with existing literature, it seems that fairness in assessment for learning requires a shift from understanding assessment as the task of capturing a complete picture (i.e., snapshot) to seeing it as a process of refreshing a changing picture frequently (i.e. digital stream) to integrate new information and provide a real-time image of a student's learning.

The second interpretation of opportunity to demonstrate learning seen in the participants' responses related to their understanding of learning theory. Several participants expressed the idea that students were more likely to be meaningfully engaged if opportunities to demonstrate different styles or types of learning were provided. This brings to mind Gipps' (1999) explanation that when learning is viewed as a "process of personal knowledge construction and meaning making" (p.374), as it is from sociocultural and constructivist perspectives, diversity is needed in the methods of assessment. The participants in my study expected students to learn through assessment, and they felt that to be fair, assessment tasks had to be varied, or tailored, to engage students' interest in the process. Badger (1999) put forth a similar idea in a model of equitable assessment that embraced newer ideas in teaching English Language Arts, particularly the recognition of non-print literacies, the cultural context of all texts, and personal interpretation of texts. She argued that assessments should have value for students because "unlike physical phenomenon, the measurement of competency needs the willing cooperation of the person whose competency is being measured" (p.68). When students are not genuinely engaged in an assessment task, it risks

becoming a measure of their willingness to comply, rather than an indication of their learning. In this way, fairness in classroom assessment comes not only from providing students with multiple and varied opportunities to demonstrate their learning, but also by offering opportunities that are meaningful for the students' learning.

Assuming that all assessment tasks should engage learners muddies the divide between assessment purposes. While the tasks that some of the participants described were clearly intended to provide information for final grades, they were also learning activities. This blend of formative and summative functions is of note because the participants' ideas about the purposes of assessment seemed to guide many of their responses about fairness. Purpose was a determining factor in whether they felt certain assessment practices were ethical or unethical (e.g. surprise items on test), and some considered purpose in recommending the fairest decision for a particular assessment situation (e.g., diagnostic purpose at the beginning of a course). The participant's attention to purpose in discussing fairness suggests that not all assessment information can be used fairly for any or every purpose. However, the participants also seemed to feel that all classroom assessment could support learning, and that students should be oriented to this purpose of assessment. In this light, it seems that the purpose of an assessment should be considered by teachers for fair decision making, and that purpose should be also communicated to students without precluding them from using any or all assessment information for their learning.

#### *Transparent Communication about Assessment*

Clearly communicating about learning expectations and assessment criteria is one of the core practices of assessment for learning (ARG, 2002; Stiggins, 2002). The same idea is

expressed as the principal of transparency, and it is commonly associated with fairness in both theoretical (e.g., Gipps, 1999; McMillan, 2007) and empirical literature (e.g., Morgan & Watson, 2002; Szyrka, 2001). In the assessment of learning, the need for transparency appears to have been accepted for some time. For example, Brookhart (1994) noted more than a decade ago that “one measurement principle that is commonly followed is the clear communication to pupils, ahead of time, of what their grades will reflect” (p.290). In a seminal speech about assessment in a learning culture, Shepard (2000) pointed out that “access to evaluation criteria satisfies a basic fairness criterion” because “we should all know the rules for how our work will be judged” (p.60). In the context of English Language Arts in Ontario, understanding of this principle is seen in teachers’ responses to the 2002 SAIP questionnaire, where almost all (97.9%) indicated that they discuss assessment criteria with students at least a few times a month when writing tasks are assigned (CMEC, 2002a). In their responses to specific situations in assessment for learning, the participants in my study were unanimous in recommending that learning expectations and assessment criteria be shared with students. Moreover, they mentioned this aspect of fairness repeatedly in discussing their own assessment practices, and they emphasized that students should understand the learning expectations and assessment criteria from the outset of a writing assignment. In effect, the participants’ practical wisdoms about this aspect of fairness in assessing writing for learning are consistent with an accepted principle that was previously established in more general terms for the assessment of learning.

Some difference of opinion was heard among the participants regarding the use of rubrics as the vehicle for sharing assessment criteria. This is also consistent with existing

literature on teaching and assessing writing, where a debate about the benefits and drawbacks of rubrics is ongoing (e.g., Badger, 1999; Newkirk, 2004; Wilson, 2006). Nevertheless, the OME curriculum documents support the use of rubrics (see for example 2007a, p.22), and they do appear to be commonly used by English teachers in Ontario (CMEC, 2002a). While none of the participants in my study indicated that they would avoid using rubrics to assess writing, several suggested strategies to ensure that rubrics support learning without constraining student creativity. These included revising the rubric, discussing its purpose directly with students, discussing concerns about the rubric with colleagues, using exemplars first to convey the learning expectations, and involving students in developing the criteria for the rubric. Given that students' accuracy in assessing writing with rubrics is questionable (Laveault & Miles, 2002; Ross et al., 1999), all these strategies might add to the fairness of students' self-assessment with rubrics.

As a strategy to improve assessment with rubrics, having students develop the criteria is of particular note because student involvement is generally considered to be at the heart of assessment for learning (Black et al., 2003; CERI, 2005; Stiggins, 2002). One participant suggested that learning expectations are clarified when students participate in the development of criteria, and he noted that this could contribute to the fairness of an assessment. He drew a connection between transparency and learning that is similar to Shepard's (2000) thinking about the cognitive and metacognitive benefits of helping students understand assessment criteria. She explains that these benefits "speak to a different sense of fairness than merely being even-handed in evaluating students, that is, they provide students with the opportunity to get good at what it is that the standards require" (pp.61). However,

none of the participants in my study suggested that the learning benefits associated with transparency necessarily require assessment criteria to be developed by students. In summary, clear communication about learning expectations and assessment criteria is associated with fair assessment for learning, but it is not clear that student-developed criteria provides any additional benefits for fairness that cannot be achieved by working through teacher-developed criteria with students.

#### *Appropriate Assessment for Individual Learners*

As an aspect of fairness, the appropriateness of an assessment for individual learners comes up at two different points in an assessment cycle. The first relates to differentiating an assessment process or task. The types of differentiation that should be made for English language learners and students with special needs (i.e., accommodations and modifications) have been increasingly specified in the OME's assessment policy documents over the past decade. The participants in my study identified a range of situations where some form of differentiated assessment was possible. Most gave examples from their own practices where they had accommodated a student, or modified an assessment in some way, and these all related to students who had either been formally identified with special education needs, or who were dealing with extremely serious events at home (i.e., death of parent). Their stance in this regard was consistent with the idea expressed more generally by teachers in Ontario that job satisfaction is derived mainly from helping students (OCT, 2003, 2004). It is also similar to the notion of "pulling for students" that McMillan and Nash (2000) described in their study on classroom assessment practices where they concluded that most teachers aim to "give students the best opportunity to be successful" (p.9).

The participants' responses in my study varied, but were less accommodating overall when students' special needs were not explicitly identified. I see two possible explanations for this: assessment purpose and perception of learning needs. First, the purpose of an assessment may be a determining factor in whether teachers believe accommodations or modifications are fair. The participants all responded favorably to differences in teacher-student interactions that occurred early in the assessment for learning cycle, even when the student was not identified as having special learning needs. However, some expressed reservations about differentiating for individuals at the end of a unit or project. Brighton (2001) found that the teachers in her study on differentiated instruction and assessment were enthusiastic about individualized instruction, but they were concerned about fairness in the evaluation phase, and they "had difficulty grading and defending their grading practices relative to the varied levels and choices of assignments" (p.133). When assessment for learning interactions are initiated, teachers are still in their coaching role, and differentiation seems to be perceived as a means of providing diverse students with opportunities to learn according to individual needs. This suggests that in assessment for learning, equitable treatment of students is of prime importance to teachers, but when they shift to a judging role in assessment of learning, equal treatment becomes a larger concern for some.

The perception of learning needs may also explain the variation in participants' responses for situations where students were not formally identified. In their suggestions for the vignette teachers, and in the stories they offered from their own practices, the participants associated accommodations and modifications with learning difficulties, and they seemed comfortable with the idea that differentiated assessments were fair for students who faced

some type of learning challenge. In contrast, when issues arose with students who were intellectually capable but not succeeding academically, the participants were less likely to view differentiated assessment as a fair alternative. Brookhart (1993) noted a similar “double standard” (p.140) in how teachers were more lenient in the grades they awarded to below-average students than average or above-average students. These types of assessment decisions appear to be influenced by multiple internal factors. For example, in her examination of the differences in teachers’ assessment practices, Watson (1999) found that teachers’ beliefs about students’ learning behaviour, their perceptions of different learning styles, and the type of effort they valued based on their own learning experiences were all involved. This suggests that when decisions about differentiating assessment appropriately for individual students are made by teachers (i.e., when students are not formally identified with special language or learning needs), some types of students will be better served than others. This is a point of consideration for classroom assessment in Ontario, first because of the diversity of the student population, and second because of the discrepancy between the percentage of students officially identified as having special needs (OCT, 2002) and the larger percentage that teachers report in their classrooms (OCT, 2006).

The appropriateness of an assessment for individual learners was also discussed in relation to a second point in the assessment cycle, when students are given feedback. Giving students balanced and descriptive feedback is one of the core practices of assessment for learning (Black et al., 2003; Stiggins & Chappuis, 2005). In their responses about the practice of writing only positive comments on report cards, most of the participants felt that weaknesses should also be included to give students and parents some direction for the

student's improvement. The ideas they expressed on this point were not unlike Shepard's (2006) argument that feedback is most useful when it "provides guidance about what to do to improve" (p.632). Although one participant felt that the tone of feedback to students should be as positive as possible, the two participants with the most teaching experience both felt strongly that for feedback to be fair, it could not be empty or false. Again, this is similar to Shepard's advice about giving students feedback. She writes that it is a "mistake to give false praise in an effort to motivate students and boost self-esteem. At the same time, straightforward, let-the-chips-fall-where-they-may, negative feedback can undermine learning and students' willingness to make subsequent effort" (p.632). Similar ideas were expressed in the ethical inquiries by Buzzelli and Johnston (2002), Jackson and colleagues (1993) and Campbell (2003). Regardless of whether the teachers in these inquiries held uniformly low expectations, or gave students indiscriminately positive feedback, the lack of honesty had the same consequence: individual achievement was not recognized or valued. Some teachers may be reluctant to point out individual weaknesses due to a concern, as one of the novice teachers in Graham's (2005) study expressed, about damaging student self-esteem, particularly if a student has not already experienced a great deal of academic success. However, Wiggins (1993) argues that fairness involves helping students "to gain from the kind of assessment and competition that properly balances challenge and possible success" (p.174). With decades of experience informing their practical wisdom, two participants in my study seemed to agree. They felt that indiscriminate feedback was unfair because it had negative consequences for students in the long term, and they emphasized the need for students to understand their weakness as well as their strengths in order to continue learning.

### *Consistent Treatment for a Group of Students*

A pattern across the participants' responses that initially seemed contradictory relates to consistency in assessment. Unlike the teachers in the existing literature who aimed to be consistent while they assessed a set of assignments (e.g., Eggen, 2004; Ryan, 2001), the participants in my study expressed very little concern about the vignette teacher who faced a similar issue. However, they all subsequently discussed the need to treat students equally during the assessment process, and when they were asked about what would be most important for fair classroom assessment, several stressed that teachers should be consistent. Given this, and the association of fairness with consistency not only by teachers in other educational contexts (e.g., Yip & Cheung, 2005; Yung, 2001b), but also in established assessment principles (JAC, 1993; JCSEE, 2003), I wondered why the issue of inconsistent feedback had not provoked stronger responses. Some light is shed on this by Black and colleagues (2003). They observed that teachers have a "tendency to emphasize quantity and presentation of work and to neglect its quality in relation to learning" (p.11), and this was evident to some extent in my study. The participants focused more the quality of the vignette teachers' feedback (i.e., content, useful medium), and they were not terribly concerned about students receiving the same amount of information. However, when it came to discussing students' work, almost all of the participants stressed the importance of having every student complete the same amount of work. This brings to mind an example given by Shepard (2006) where teachers seemed to be "creating an elaborate system to keep track of student work, but without evaluating the quality or content of that work" (p.636). While the participants in my study seemed to be more concerned about quality than the teachers observed by Shepard or

Black and colleagues, their emphasis on work completion begs the question of why, especially when focusing on quality would be more effective for supporting student learning. It may be that it is easier for teachers to count pieces than assess quality. Alternatively, as some of the participants' comments in my study suggest, it may also be that moral beliefs underlie teachers' emphasis on having all students complete the same amount of work. The link between moral beliefs and the emphasis on work completion is considered further in the discussion on teachers' practical wisdoms.

Shulman suggests that "when we are confronted with what appears to be examples of teachers' resistance to change" (1987/2004, p.264), we should investigate further. A theme that emerged as I discussed the need for consistency with the participants in my study relates to teacher-student relationships. Katz and colleagues (2001) note that in grading, teachers need to ensure that each "child appears to be treated as fairly as every other child" (p.22), and the idea of appearance, or perception, is key here. Green and colleagues (2007) explain that trust in teacher-student relationships can be harmed "by assessments that the student perceives as unfair or unfounded" (p.1009). The participants' in my study wanted to appear fair by treating all students equally because they were aware of their students' perceptions. Research with students in several different educational contexts has shown that while a certain degree of differentiation for peers with special needs tends to be accepted, equal treatment in the classroom is highly valued (Bursuck et al., 1999; Smith & Gorard, 2006). Teachers seem to respond to students' perceptions of fairness by attempting to maintain the appearance of equal treatment even when it is not needed for the purpose of the assessment. For example, even when the teachers in Torrance and Pryor's (2001) study were giving

students individual feedback, which is inherently inconsistent, they still “recognized that any differentiation had to be seen to be fair” because the students are “aware of each other’s attainment” (p.625). Thus, it seems that the tradition of equal treatment, which has merit in the assessment of learning, spills over into assessment interactions that could focus more equitably on learning.

Discussions about fairness in classroom assessment often use a sports game analogy, but in different ways. For example, the teacher in Gummer’s (2000) case study described giving students equal opportunity to demonstrate learning as a level playing field. In contrast, a participant in my study referred to a level playing field as the use of common criteria by teachers to assess students’ work consistently across classrooms. Wiggins (1993) also makes use of the sports analogy, but he suggests that fair assessment should “ensure that everyone is learning to play the ‘same’ game so that everyone can improve” (p.174). Current materials for teachers contain statements such as “fairness is not sameness” (OME, 2007b, p.28) and “fair isn’t always equal” (Wormeli, 2006, p.6) to counter the common association of fairness with equality in education. What Wiggins points out, though, is that some degree of sameness, or equality is beneficial in terms of learning in a classroom. An example that is particularly relevant for thinking about fairness in ongoing assessment interactions is turn-taking. Jackson and colleagues (1993) identify turn-taking as one of the “just practices” (p.28) of the classroom, and Shulman (1987/2004) argues that turn-taking results in a “more equitable distribution of learning opportunities” (p.261) for students. Without some level of equality in classroom interactions, assessment for learning would be akin to a board game where some players take turns more frequently than others. The participants in my study

seemed to be conscious in how they managed the assessment process of providing students with equal opportunities while maintaining flexibility to also permit equitable opportunities. As such, it seems that teachers may be justified to some extent in aiming for consistency in their practices as a means of improving the fairness of classroom assessment.

A challenge for some teachers in this process may be to know exactly when consistency is, or is not important for fair assessment. Cowie (2005) observed that students tended to be less reserved as trust and respect in classroom relations grew over the school year, which suggests that teachers may need to vary their approach to fairness at different points. It may be that consistency is needed to develop trust at the beginning of a course (secondary) or academic year (elementary) in order to spark an assessment for learning spiral. Ideally, as the cycle progresses, the constructive nature of the learning environment becomes apparent to the students, and the teacher gains a better understanding about each student's learning, which increases not only the relevance of assessments designed for the class, but also the appropriateness of differentiation for individuals. As such, fairness in classroom assessment may require more than a choice between consistent practices where the intent is to treat all students equally, and differentiation where the goal is to provide equitable treatment for particular students. What may be needed is a gradual shift from a high degree of consistency in the tasks, procedures and interactions that occur early on in an assessment spiral to a more flexible system that actively supports individual learning as trust, respect, and understanding develop between teachers and students.

#### *Reflective Interaction in the Assessment for Learning Cycle*

Calls for teacher reflection are frequently heard in the existing literature on fair

classroom assessment, particularly recommendations for teachers' to consider their roles and biases, question stereotypes, and be sensitive to student differences (e.g. Airasian, 2005; McMillan, 2007; Speck, 1998; Tierney, 1998/2005; Volante, 2006). At minimum, teachers need to reflect about their classroom observations in order to gather "fragments of information together into a more or less meaningful picture of an individual student" (Bulterman-Bos et al., 2002, p.1086). In my study, the participants were asked to give their recommendations for two situations where teachers had made assumptions while assessing student writing. Although the participants all recognized and commented on the assumptions, recommendations relating to teacher reflection were limited. Two participants felt that the teacher in one situation should reflect on his biases, and they also stressed the importance of teacher reflection for fair assessment more generally during their interviews. None of the participants mentioned reflection in relation to the second situation. To some extent, this may have been due to timing. The vignettes described teachers who were on the verge of making a decision, which may have led the participants to suggest actions rather than reflection. Nonetheless, the dearth of recommendations for teacher reflection is notable because the participants clearly gave thought to their own biases in assessment. For example, some participants discussed the tension between knowing students and using information about students fairly in assessment decisions, and others shared reflections about the human or personal nature of assessment interactions. As such, all of the participants engaged in reflection about their assessment practices, but only two articulated the value of reflection as a strategy for fair classroom assessment.

Current classroom assessment literature often goes beyond discussions about tools or

methods to consider the people involved in the process. Several of the standards associated with fairness in *The Student Evaluation Standards* (JCSEE, 2003), for example, recommend that students be actively involved in classroom assessment. A variety of studies have shown that this can have positive benefits for students' learning (e.g., Black & Harrison, 2001; McDonald & Boud, 2003; Shepard, 2006), and it is sometimes thought to improve the fairness of assessment. For example, Speck (1998) suggests that "ethical concerns about classroom assessment" (p.28) in post-secondary education can be addressed by involving students in the assessment process. Most of the participants in my study felt that involving students in the assessment process was beneficial in terms of student learning. As students gain a better understanding of expectations and criteria through their involvement, it should, in theory, make the process more transparent, and thus fairer. However, the participants did not suggest that involving students necessarily made an assessment any fairer. If anything, it seems that engaging students in peer assessment poses additional challenges for fairness. In their recommendations, the participants strongly emphasized that it is the teacher's responsibility to be proactive in establishing a respectful and constructive environment in order to avoid any negative peer interactions that might lead to unfair assessments.

Although the participants all suggested strategies to control the quality of students' assessment interactions, only two participants explicitly discussed power. One participant acknowledged the power of students in peer assessment. He felt that teachers should help students understand the impact they had on each other with their feedback. The idea that peer feedback can have a powerful impact is supported by Cowie's (2005) work with secondary students where she found that classroom assessment, including peer interactions, shaped

students' identities as learners. Currently classroom assessment literature encourages teachers to reflect about the fairness of their assessment decisions (e.g., Airasian, 2005; Volante, 2006), but when students are actively engaged in self and peer assessment they become the assessor as well as the assessed. Thus, it seems that students, as well as teachers, should be encouraged to reflect about the values and biases that underlie their assessment decisions.

The notion of students yielding power is also consistent with the idea of students as active agents, rather than passive subjects, in classroom assessment. In her discussion about the differences between large-scale testing and classroom assessment, Brookhart (2003) explains that while the first is an externally mandated event, the latter is a cyclical process where the "scores, grades, feedback, and other classroom information fold back into and become part of the classroom environment" (p.7). In this process, a "student is not the 'subject' whose achievement is measured but an active participant in the process" (p.8). Bulterman-Bos and colleagues (2002) take this idea further in writing about observation as an assessment method, arguing that students have power in this process because they "contribute actively to what can be observed" and can "even compel teachers to take notice" (p.1085). However, there is also evidence to indicate that students may not be prepared to yield power consciously in the classroom assessment processes. For example, the young students who engaged in peer and self-assessment in Dann's (2002) case study were reluctant to challenge the teacher's decision if it differed from their own. The students' power in the assessment process was, essentially, usurped by the teacher, as the adult authority in the classroom. The participants in my study emphasized the importance of trust and respect in classroom interactions, which might have shifted the power dynamics in Dann's case study.

However, none of the participants in my study discussed assessment for learning interactions in terms of power. For the most part, they saw the teacher as the prime assessor and classroom leader in their responses to the vignettes, and this stance was also evident in many of the stories they told about their students.

The second participant who explicitly discussed power acknowledged his own use of power in grading. He explained that he had lowered a student's grade a one point because he thought it would send a strong message to the student about the need for more productive behaviour. As such, he used his power as the assessor in a way that he thought might benefit the student in the long term. Overall, the participants indicated that they gave considerable thought to the consequences of assessment for their students, and they tended to keep the best interests of their students in mind. However, none of the participants recommended that teachers reflect about power dynamics as a means of improving the fairness of classroom assessment, and some of the power issues in the literature on classroom assessment were not discussed at all. For example, several researchers who use Michel Foucault's ideas about power to frame classroom-based studies have voiced concern about tools and methods commonly used in classrooms. Brookfield (2001) discusses learning contracts, student journals, and the evaluation of participation in classroom discussion in terms of power, and Schendel and O'Neill (1999) argue that the power dynamics involved in asking students to assess their own disclosures could serve to "replicate and reify the problems of grading writing" (p.206). When students are involved in assessment, whether it be through self or peer assessment, or through strategies such as choosing assessment tasks or criteria, the relationships between teachers and students in the classroom are changed.. The participants

in my study saw reflection as an important aspect of fair assessment, but without clearly recognizing how power dynamics might influence the process, especially if these dynamics shift as students and teachers take on different roles in assessment. The participants responses hinted at some of the power issues in classroom assessment, but this is an area that could bear much further investigation. To understand teachers' reflection in classroom assessment further, I consider how the participants' practical wisdoms informed their responses about fairness in classroom assessment.

#### Teachers' Practical Wisdoms about Fairness

I described practical wisdom as a fluid network of internal elements that individuals draw on in unique ways in different contexts earlier in this dissertation (see Chapter Four). Looking specifically at teachers' practical wisdoms about fairness in classroom assessment, the idea of an internal network is still apt as the participants responses did seem to be informed by a dynamic inter-working of various forms knowledge, understandings, beliefs and characteristics that had developed through both formal learning and experience. In their expression, these various elements of practical wisdom were often tangled, like a mess of roots in an old garden bed. With the participants' willingness to reflect aloud about their practical wisdom, I was able to tease it apart sufficiently to gain some understanding of how the different elements informed their recommendations for fair classroom assessment.

#### *Types of Knowledge at Play*

All of the participants had studied English at the post-secondary level, completed a teacher education program, and engaged in further professional development through additional qualification courses and work-related activities (e.g., action research projects,

curriculum writing teams, school board workshops). Through their learning careers, they would have gained both theoretical and technical knowledge of relevance for teaching English Language Arts. Most of their expressions of theoretical knowledge related to learners and assessment rather than English writing (i.e., content knowledge). This may be because the questions I asked on both questionnaires and during the interviews focused more on assessment than writing. Even so, one participant stressed the importance of strong content knowledge for fair assessment. Morgan and Watson (2002) came to a similar conclusion when they found that differences in content knowledge affected the value that teachers attributed to students' work. They suggested that teachers with stronger content knowledge are more likely to recognize merit in atypical or unexpected responses than teachers who have less content knowledge informing their assessment decisions.

The participants in my study usually referred to, or expressed content knowledge in combination with pedagogical knowledge. This blended knowledge seems, at face value, to be better represented by Shulman's (1987) concept of pedagogical content knowledge than by Aristotle's ideas about knowledge (as discussed in Chapter Four). Because Aristotle distinguished between types of knowledge, some researchers have rejected his ideas (e.g., Clandinin & Rosiek, 2007). However, Elbaz (1980) argues that Aristotle's work has been "misrepresented and oversimplified" (p.25), and she believes that it emphasized the "inseparability" of theoretical and practical knowledge. In the examples Aristotle gave in the *Nicomachean Ethics* (trans.1925/1975) to explain the concept of practical wisdom (e.g., knowing light meats, young mathematicians), it is clear that he thought various elements were necessarily combined. To illustrate Aristotle's notion of a phronimos, Dunne (1993)

gives a description of the different types of knowledge needed by a mathematics teacher that highlights their inseparability. Dunn does not refer to Shulman's ideas, but because his example involves a teacher, it reveals the similarity between practical wisdom and pedagogical content knowledge. Essentially, they both involve the combining of theoretical and technical knowledge, which may be expressed more specifically in relation to teachers as content and pedagogical knowledge. What should not be forgotten in comparing these two ways of thinking about teachers' knowledge, though, is that practical wisdom also involves moral beliefs and ethical knowledge, and it cannot be separated from personal characteristics, experience, and the particulars of practice.

### *The Particulars of Practice*

Dunne (1993) explains that what is most important in understanding practical wisdom is not just that it combines different types of knowledge, but that it applies this combination in relation to particular circumstances.

The crucial thing about phronesis . . . is its attunement of universal knowledge and the techniques to the particular occasion, so that they are deployed in relation to 'the right person, to the right extent, at the right time, which the right aim, and in the right way . . . [which] is not for everyone, nor is it easy' [Nicomachean Ethics cited]. (p.368)

Practical wisdom is inseparable from practice. Looking at any one form of knowledge, or even a composite concept such as pedagogical content knowledge, is insufficient for understanding teachers' decisions in classroom assessment because it neglects the influence of the particular. It was clear from the participants' responses in my study that their practical wisdom about fairness was dependent on their understanding of a particular context, situation or student. This was evidenced by the frequency with which they either chose the "depends

on circumstances” option on the questionnaire, or qualified and clarified their written responses. It is also seen in the prevalence of their use of particular students and situations as examples in sharing their practical wisdom about fairness in classroom assessment.

The importance of understanding particular students and their circumstances emerged repeatedly in the participants’ recommendations for fair classroom assessment, and this raises an issue for ensuring fairness in teacher-student interactions. Kessels and Korthagen (1996) argue that because practical wisdom is perceptually rather than conceptually based, teachers “must above all be able to perceive and discriminate the relevant details” in order to “choose and justify a particular course of action” that is “appropriate for a situation” (p.19). As such, making appropriate decisions for individual students, which is an aspect of fairness that was discussed in the previous section, depends on teachers’ perceptions of students. This highlights the importance of teachers’ critical reflection about their beliefs and biases, and gives credence to the emphasis that some of the participants placed on reflection as an aspect of fairness in classroom assessment.

### *The Moral Dimension*

The practical wisdom about fairness expressed by the participants was often based on their moral beliefs and ethical knowledge. Sometimes they expressed themselves in terms of professional obligations, or actions to take in particular circumstances, and usually these types of statements were consistent with the ideals presented for teaching in the *Foundations of Professional Practice* (OCT, 2008a). For example, the obligation to investigate the reasons for a students’ lack of achievement, which was associated with fair assessment by one participant, is consistent with the view of teachers as “ethical decision-makers who exercise

responsible, informed professional judgment (p.17), and the openness to differentiation that several of the participants expressed was in keeping with the image of teachers as “responsive pedagogical leaders who are respectful of equity and diversity within Ontario’s classrooms and schools” (p.17). The participants were not always aware of the moral beliefs and ethical knowledge in their written responses, and in several instances they realized the moral underpinnings of a recommendation only as they were in the process of explaining it.

Two categories of moral beliefs and ethical knowledge were evident in the participants’ recommendations for fair classroom assessment: work ethic and human relationships. The moral imperative for helping students develop a good work ethic, which was explicit in some participants’ recommendations and implicit in others, is not stated as such in the *Foundations of Professional Practice* (OCT, 2008a). The strength of the participants’ feelings on this topic is reflected, however, in Leibowitz’ (2000) book about teaching habits of mind. She writes that despite ongoing public debate about educational issues, there tends to be “universal agreement on at least one aspect of schooling: Students need to develop and demonstrate a stronger work ethic” (p.62). She explains that this is because the “professional world outside schools requires demonstration of a work ethic and makes the need for this ethic visible to the public and employees” (p.62). It may be that teachers feel strongly about encouraging good work habits because they believe that preparing students to function successfully within the work world is part of their professional mandate. In this light, the participants’ association of fair assessment with equality in the amount of work completed by students falls under the professional standard of being committed to students and student learning, which states that “members facilitate the

development of students as contributing citizens of Canadian society” (OCT, 2008a, p.13).

While the meaning and merit of this standard offers several avenues for debate, what is pertinent to the issue of fair classroom assessment is whether teachers’ evaluation of the quantity or quality of work completed by students best serves this social purpose.

The second category of moral beliefs and ethical knowledge, relating to human relationships, is more evident in the literature on classroom ethics. For example, Colnerud (2006) notes that “it is in the human relationships that the more enduring moral values are at stake. Morals are shaped in everyday events and the teacher is a role model whether he or she is aware of it or not” (p.376). While only two of the participants referred explicitly to being role models, they all stressed the necessity of trusting and respectful relationships for fair classroom assessment. Colnerud also points out, however, that because of the “relative powerlessness and vulnerability of children” (p.380) teachers are presented with temptations (e.g., classroom management through humiliation or sarcasm) that are sometimes even an “institutional habit” (p.380). Several of the participants in my study explained that having been on the receiving end of such treatment as children, they consciously aimed not to repeat it with their students. One participant indicated that he was guided by the golden rule, which is a widely accepted moral principle that states, in various renditions, the idea that we should treat others as we wish to be treated (Strike & Soltis, 1998; Waluchow, 2003). The ethical standards in the *Foundations of Professional Practice* (OCT, 2008a) are similar to the golden rule in that they present universal principals (i.e., care, trust, respect, and integrity) to guide human interaction, but specifically in relation to students and teachers in Ontario.

Although the participants’ recommendations for fair assessment were generally

consistent with the ethical standards in the *Foundations of Professional Practice* (OCT, 2008a), none referred specifically to this document. This may be, as Dunne (1993) explains, because “practical-moral universals” (p.311), such as the ethical standards, “contain in themselves an element of indeterminateness which is removed only through confrontation with a particular case” (p.311). In defining an equitable judicial system, Aristotle explained that established laws must be “corrected” for specific cases (Dunne, p.311). He used the metaphor of the “leaden rule” (trans.1925/1975, p.133), which was a tool used by builders to measure around stonework, to illustrate the problem of applying a rigid rule in practice. Given that fairness is dependent on the particulars of practice, and that the participants’ supported their recommendations with specific cases rather than general principals, the degree to which ethical rules actually help fair decision-making in classroom assessment is not evident. In a review of research relating to ethics in teaching, Colnerud (2006) questioned whether ethical rules restrict or support teachers’ critical reflection and the development of practical wisdom. He concluded that “since the moral dimensions of teaching have been implicit in practice and the moral language of teaching is weakly developed, rules in the Aristotelian sense might provide tools for reflection” (p.372). Because the ethical standards for teachers in Ontario must be sufficiently malleable to accommodate the diversity of students and learning environments, they offer general guidelines that teachers may not connect to the particulars of practices. A more potent use for these ethical standards might be as a tool to help teachers initiate and frame reflection and collegial discussion about fairness in classroom assessment.

### *The Circularity of Experience*

In their explanations about how they came to their recommendations for fair classroom assessment, the participants in this study all referred repeatedly to experience. Several described memorable experiences as learners, and others mentioned their experiences as parents, but most often their practical wisdom about fairness in classroom assessment was informed by their experiences as teachers. For the most part, it would have been impossible to distinguish between teaching experience and the other elements of practical wisdom had the participants not articulated which elements they thought they were drawing on for their different recommendations. This is because, as Dunne (1993) notes, practical wisdom “arises from experience and returns into experience” (p.293). The participants often used examples from their teaching experience to support their responses, but this did not seem to reduce their practical wisdom to a collection of anecdotes. This may be because experience “stabilizes the succession of individual and particular impressions . . . through a process of universalization” (Dunne, p.293), which then functions as a frame of reference for new circumstances. Dunne argues that the “more experience is reconstructed in this way, the more sensitive and insightful phronesis becomes – or, rather, the more the experiencer becomes phronimos” (p.293). In this way, the relationship between a teacher’s experience and practical wisdom is circular; while experience is part of the network, it also supplies the network.

The notion of experience as a rich and endless source for the development of practical wisdom begs the question about why experience and practical wisdom are not always correlates. Aristotle’s interpreters seem to agree, despite the ongoing discussion about the

meaning of his words, that the development of practical wisdom requires more than experience (e.g., Dunne, 1993; Warne, 2006). Hughes (2001) explains that “to have practical wisdom is to be good at thinking about how to live a fulfilled and worthwhile life as a whole” (p.84), and Warne (2006) writes that the “person of practical wisdom characteristically makes excellent decisions about the best actions to perform in any set of circumstances. He arrives at these decisions through deliberation” (p.90). Although I did not attempt to assess the quality of the participants’ practical wisdoms in my study, all of their recommendations for fair assessment seemed to rest on what they thought would be best for their students. A personal characteristic that they clearly shared was their interest in teaching and learning. As research participants this was manifested as a willingness to discuss their teaching and learning experiences, and to deliberate on the fairness of assessment interactions. Shulman (1996/2004) points out that the “process of remembering, retelling, reliving and reflecting is the process of learning from experience” (p.474). Taking the circularity of experience and practical wisdom further, Dunne (1993) argues that “phronesis is what enables experience to be self-correcting and to avoid settling into mere routine” (p.292). As an element of practical wisdom about fairness in classroom assessment, experience not only informs teachers’ assessment decisions, but it is also informed by the ongoing interactions of the classroom. Thus, for teachers to develop practical wisdom about classroom assessment, experience should not be used as a crutch for unexamined decisions, but as a valuable source for enrichment and reflection.

#### *Elucidating Practical Wisdom about Fairness*

Previous studies have found that teachers have difficulty explaining some aspects of

their classroom assessment practices. For example, Dixon and Williams (2003) found that the descriptions of formative assessment given by primary teachers in New Zealand were fairly limited. McMillan and Nash (2000) also found that the teachers they interviewed, especially those with many years of teaching experience, had difficulty explaining their classroom assessment decisions. This contrasts sharply with Campbell's (2003) observation that while teachers' actions and interactions in classroom in Ontario seemed to be "largely spontaneous and habitual, some teachers nonetheless do seem able to perceive and explain them within a moral and ethical framework to an extent greater than that with which they have been previously credited" (p.39). The difference in the teachers' responses in these three studies may be a product of the researchers' approach. In the first two studies, teachers' explanations of their practices were examined in relation to existing classroom assessment theory, whereas the latter study sought to illuminate teachers' knowledge as it existed in practice. My approach falls somewhere in between in that I drew on teachers' thinking as a source to better understand an existing concept in assessment theory. None of the participants in my study had difficulty describing what they did in their assessment practices, and they were all able to elucidate their understanding of fairness as a quality of classroom assessment. Unlike the experienced teachers in McMillan and Nash's (2000) study, the most experienced participants in my study were the most fluid in articulating their thoughts.

Where the all of the participants' words became more halting in my study was when they tried to explain how they came to their recommendations for fair classroom assessment. This may be partly because we are not always aware of the connections between the elements of our practical wisdom. Unless some aspect of an experience makes it particularly

memorable (e.g., receiving hurtful feedback from a teacher, reading a text with personal relevance), it can be difficult to pinpoint the underpinnings of our beliefs, understandings and knowledge. It may also be that the elucidation of practical wisdom is necessarily limited because it is both dynamic and situated. Dunne (1993) writes that we can “never quite catch up” (p.357) with what we know because in gaining knowledge we are “always already beholden to assumptions, antecedent interests and tacit procedures which are not themselves known” (p.357). What seemed to be helpful in facilitating discussion about practical wisdom in my study was to provide particular assessment situations (i.e., vignettes) from which the participants could elucidate their practical wisdom. Being able to read and respond to these situations in writing gave the participants time to think about their recommendations for fair assessment. However, explaining their responses aloud seemed to spark additional reflection that might otherwise have remained dormant. It seems that even though practical wisdom is individual and rooted in the particulars of practice, discussing it with others helps it flourish.

#### The Influence of External Factors

When the participants in my study were asked to identify any factors that might influence the fairness of writing assessment in English Language Arts classes in Grades 7 to 12 in Ontario, they most frequently discussed student factors. To a lesser degree, they also saw colleagues, policies, professional development opportunities and time constraints as sources of influence. In considering fairness specifically as a quality of the interactions in assessment for learning, these factors essentially fall into two groups: a) the framework mandated for the interaction, and b) the people directly involved in the interaction.

### *Framework Factors*

Although the external factors that frame assessment interactions were discussed less frequently than people factors, the possibility of their influence on fairness was clearly noted. The participants referred to both negative and positive aspects of OME assessment policies in responding to the vignettes, and in discussing their own practices. One participant appreciated the emphasis on student learning in the newer assessment documents produced by his school board and the OME. Another felt that the achievement chart levels facilitated marking, and made teachers' assessment decisions more consistent, and thus fairer for students. A problem identified by Watson (1999) and Schmidt (2001), namely that teachers face conflicting assessment policies within a system, was not strongly voiced by the participants in my study. The examples and stories they offered suggested that assessment of learning was foremost in their thoughts, but there was little indication that using assessment information for both learning and grading purposes hampered fairness. Two aspects of the provincial assessment system, report cards and large-scale assessment, generated minimal discussion. One participant noted that students' strengths and weaknesses were not clearly communicated because of the limited space for comments and the language used on standardized report cards. It could be argued that the fairness of an assessment process is jeopardized if the information reported is insufficient for students to understand and use, but this was not a point raised by the participants. They also expressed little concern about the impact of EQAO's large-scale assessment program, with only one participant indicating that it would in any way affect his classroom assessment practices. This is not to say that the tension between assessment purposes does not exist in Ontario, but the participants in this

study did not identify large-scale assessment or standardized report cards as factors that would hinder or facilitate fairness in assessing student writing in the classroom.

The only aspect of assessment policy that the participants explicitly linked with fairness related to student responsibility. Several participants described recent situations in their practices where a particular student had not devoted sufficient time and effort to course work. They felt that OME policies should emphasize students' responsibilities for learning and the importance of developing a good work ethic. This is consistent with concerns expressed by a sample of mathematics teachers in Ontario ( $n = 77$ ) in a study on the extent to which they followed assessment principles and OME grading policies (Tierney, Simon & Charland, 2009), and with teachers' opinions reported in a local newspaper (Laucius, 2009). Assessment policies are currently under review in Ontario, and this issue is a matter of ongoing discussion in the educational community.

Another external factor that frames assessment interactions is time. According to existing survey data, teachers in Ontario are challenged by time constraints, particularly as many hours are spent assessing student work (CMEC, 2002a; OCT, 2004, 2006). The participants in my study also identified time as a factor that could influence the fairness of classroom assessment. Several noted that time was a necessary consideration in their choice of assessment methods or tasks. However, their comments differed from the literature on assessment reform where teachers report not having enough time to implement new assessment methods (e.g. Cheung, 2002; Hargreaves et al., 2002; Lock & Munby, 2000; Torrance & Pryor, 2001). Here the participants tended to focus on students' learning, and they felt that the quality of assessment interactions could be impacted by either the class size,

or the number of learning expectations. Several explained that this was because of the pressure teachers feel, in the current system, to move a group of students forward through the curriculum to complete a course within a specified period of time. Teachers need sufficient time to gain an understanding of students' learning in order to meaningfully support further learning. With many students and learning expectations to teach and assess, several participants indicated that they avoid more time-consuming assessment activities, such as having students develop assessment criteria. Several also explained that the amount of individualized feedback that they can give students depends on the class size. Having a large number of students reduces the "mental room available" (Bulterman-Bos et al., 2002, p. 1086) for gathering ongoing information in the process of assessment for learning. As such, teachers may move on regardless of whether their understanding of each student's learning is sufficient to continue orchestrating teaching, learning and assessment, or whether they have provided students with sufficient opportunity to learn for a subsequent assessment to be fair.

Some of the participants in my study offered suggestions for dealing with time constraints, such as matching the tool to the time frame or prioritizing certain learning expectations. Their perception of time constraints as more or less of a problem appeared to be dependent on a combination of factors in their particular learning environment. Thus, it seems likely that when English Language Arts teachers in Ontario are faced with a large number of students, many learning expectations and limited time, they will face at least two dilemmas. One is the classic choice between breadth and depth in giving students' opportunities to learn and demonstrate learning, which are identified as aspects of fairness in classroom assessment theory. The other involves a choice between adhering to OME

curriculum and assessment policies, or diverging as they believe is fair for the needs of the individual students in their care. Ultimately, the fairness of classroom assessment is influenced by the choices that teachers necessarily make in their responses to these framework factors.

### *People Factors*

Teachers are often thought of as “solo practitioners” (Cuban, 1990, p.11) who work in the isolation of their classrooms. However, current literature suggests that teachers’ classroom assessment practices are influenced by their interactions with other people, specifically students, parents, colleagues and school leaders (e.g., Hayward et al., 2004; McMillan & Nash, 2000; Schmidt, 2001). In my study, parents and principals emerged as influences to a much lesser degree than students and colleagues.

Current OME assessment policy (i.e., *Growing Success*, 2008) portrays parents in two roles: a) as part of a team helping students’ develop the learning skills that enable achievement, and b) as recipients of clear information about students’ achievement. Existing survey results indicate that most teachers in Ontario believe that supportive parents are important, but many teachers (or schools) are not communicating frequently with parents (OCT, 2003, 2006; CMEC, 2002a). Although only a few of the participants in my study mentioned parents in relation to assessment, the same two roles came up. In relation to the first, they expressed concern about parental support, but the link with fairness was not evident in their comments. For the second, they discussed the need for clear communication, and they emphasized its importance for fair classroom assessment. While this may have been to some degree, as McMillan and Nash (2000) and Schmidt (2001) note, because teachers

expect some parents to ask for justification of students' final grades, these participants also seemed to feel that communication with parents was important for classroom assessment to support student learning. With the roles assigned to parents in *Growing Success* (OME, 2008), better communication with parents would improve the transparency of the assessment process. At this point, however, there is little indication in either the existing survey results for teachers in Ontario, or from the six participants in my study, that parents have a direct influence on the fairness of classroom assessment.

The influence of principals or other school leaders on the fairness of classroom assessment was also not strongly evidenced in this study. This lack of emphasis differs from the studies done in educational contexts where system-wide assessment reform was under way (e.g., Hayward et al., 2004), and school leaders played a key role in the success of the initiative. It also contrasts with research on other classroom practices in Ontario, such as Giddens' (2001) dissertation on factors that affect how students' with disabilities are integrated in regular classrooms. The teachers she surveyed in one large urban school board (n = 884) ranked "supportive school administration" and "good working relationship with parents" (p.73) as the second and third most important factors (after class size) that helped them integrate students with disabilities. Three of the participants in my study noted that principals might influence the fairness of an assessment, but one explained that she had not experienced this herself. The other two related particular events where their decision about a student's final grade was over-ruled by the school principal, and they felt that the decision was neither fair for the student's long-term learning, nor for the other students in the class. In both of these circumstances, the principals decided the outcome of a special case (i.e.,

duplicate essay, medical absence), but they were not otherwise involved in the teachers' assessment practices. This is consistent with the role that principals are assigned in *Growing Success* (OME, 2008), where they are referred to only in relation to making assessment decisions about or for students outside the mainstream (i.e. modifications for students with special learning needs, credit recovery for low-achieving students, etc.). Two participants in my study were involved in assessment strategies that might have been initiated by one of their school administrators (i.e., department use of same criteria, collaborative assessment across a grade), but neither one identified principals or school leadership as a factor for fair assessment. Although it has been suggested that principals can facilitate collaborative practices that would be beneficial for fair classroom assessment (OME, 2007d), there does not appear to be any type of research to date on the principal's role, or the influence of principals on teachers' assessment practices in the Ontario educational context.

Teacher-student relationships are central in assessment for learning. However, little is known about the impact of students on the quality of classroom assessment. Teachers in the United States and Ontario have indicated that their classroom practices are influenced by a range of student-related factors (CMEC, 2002a; McMillan & Nash, 2000). The participants in my study clarified more specifically that student-related factors could influence the fairness of assessment for learning in the context of English Language Arts in grades 7 to 12 in Ontario. Like the teachers in previous studies (CMEC; McMillan & Nash), the participants in my study identified students' abilities as a factor that could hinder and facilitate classroom practices. The connection with fairness seemed to be limited to two areas. First, the participants discussed the overall ability of a group of students in regard to the choice of

assessment method. For example, one participant explained that the difficulty of a peer assessment activity has be aligned with the general level of the class. Second, although one participant noted that students of differing abilities could help each other, most saw a wide range of student abilities in one class as a hindrance for fair assessment. This was because it made the process of ensuring that each student's learning was appropriately assessed and supported more difficult.

When the participants discussed students in relation to fair assessment, they most frequently focused on attitudes or behaviour. A few indicated that they considered the maturity of students in choosing assessment methods, which is similar to one of the veteran teachers in Hand and Prain's (2002) study in Australia. Students' behaviour also played a role in the participants' mode of communicating assessment information. Several explained that they give feedback orally because students don't bother reading written feedback. The Grade 6 teachers in Uchiyama's (2004) dissertation on formative assessment in mathematics classrooms in the United States voiced the same sentiment. They thought that written feedback took too much time, especially when "students did not read or use the feedback anyway" (p.171). These commonalities with teachers in Australia and the United States suggest that student behaviour is a factor that influences teachers' assessment practices across contexts. However, there was also evidence that the participants' practical wisdoms about fair assessment were influenced by differences in student behaviour related to their specific learning environments. This is highlighted by comparing two participants' responses. The participant who taught at a secondary school where most students were in advanced programs (i.e., gifted and IP) believed that it was ethical to weigh homework heavily in final

grades, and she emphasized the importance of intellectual challenge in designing assessment tasks. In contrast, the participant who taught in an alternative program where many of the students were struggling academically and working outside of school hours, he felt that it was unethical to weigh homework heavily in grades, and he identified absenteeism as one of his greatest challenges in assessing student achievement. These two teachers faced very different sets of circumstances within the same system, or educational context. This suggests that while some responses to student behaviour appear to be universal, teachers' practical wisdoms are also influenced by the characteristics of particular groups of students, especially when similar types of students are clustered in distinctive programs. Furthermore, it appears that teachers' practical wisdoms about fairness are influenced by the needs of individual students within these programs. For example, several of the participants mentioned that certain practices (e.g., peer editing duos) worked well with some students and not others. In this light, it can be seen that teachers' assessment practices need to fluctuate in response to particular students, or groups of students. This suggests that when we compare teachers' assessment practices, the differences may not remain constant. Consequently, the endless call for greater consistency in teachers' practices (e.g., from Finklestein in 1913 to Morgan and Watson in 2002) should be approached cautiously to ensure that the flexibility needed for assessment to be fair for individual learners remains possible.

A notable feature of the practical wisdoms voiced by the participants in this study is that they tended to focus on the teachers' responsibilities when they discussed students' behaviour. Many of their recommendations for the vignette teachers indicated that they considered the direct and proactive management of student behaviour (e.g., discussing and

modeling respect) as part of a teacher's professional responsibility, and they emphasized the importance of a constructive environment for fair classroom assessment. However, there was very little in their recommendations about how the vignette teachers might go beyond managerial functions to reflect about the interactions in classroom assessment. The influence of student behaviour on teachers' decisions was not clearly recognized, or perhaps not articulated, by the participants. In the same way that some of the teachers in Zoeckler's (2005) study adjusted grades according to students' behaviour (i.e., boosts for cooperative students but not for others), several of the participants indicated that their perception of a student's behaviour (i.e., effort) influenced the degree to which they were willing to differentiate an assessment. Their own convictions about the merit of hard work were sufficiently strong that they did not seem to reflect critically about the quality of the information on which their perceptions were based. For example, in evaluating students' posters, maps, portfolios, or any other product, a teacher cannot judge the effort that went into the production without being well-informed about the students' abilities and habits. Buzzelli and Johnston (2002) write that teaching is a moral activity because "teachers are engaged in changing the behaviour of others to attain prescribed ends" (p.8). The participants were clearly aware of this aspect of teaching in discussing fair assessment. However, they focused on the teacher's responsibility for managing or changing student behaviour to the extent that they did not see also see the need for teachers to reflect about the accuracy of their perceptions about student behaviour, and how those perceptions inform classroom assessment interactions and decisions.

Despite this apparent "moral blind spot" (Buzzelli & Johnston, 2002, p.125) in their

recommendations for fair assessment, the participants repeatedly emphasized the importance of teachers' professional learning. In contrast to the assessment reform literature, where secondary teaching is sometimes described as being inherently conservative (e.g., Hayward et al., 2004; CERI, 2005; Ryan, 2001), the participants' in my study seemed to be open to a range of approaches in their discussion of assessment practices. The only mention of teaching traditions came from the two participants who had previously taught at the elementary level. They both observed that secondary education has traditionally been less oriented to individual learning needs, and they felt that their own teaching practices benefitted from their experiences with elementary students. Considering that the participants in this study were purposefully selected from a network of teachers who had volunteered to help others, their interest in professional learning is not astonishing. What is notable, however, is the degree to which they valued ongoing learning, and the emphasis they placed on the importance of teachers' knowledge for fair classroom assessment.

The participants' responses were replete with suggestions that "other teachers" might not be current, or demanding, or simply knowledgeable enough to assess student writing fairly. It is interesting that the incompetent teachers remained nameless, but the colleagues or mentors who had been a positive influence in their careers were either identified or situated. This raises a concern related to Campbell's (2003) point that teachers in Ontario are not able to be openly critical because of a "longstanding and prevailing professional belief that 'ethical' teachers do not interfere in the business of other teachers" (p.85). Teachers in Ontario seem to enjoy working collaboratively with colleagues (OCT, 2006), and there may be benefits for fair assessment from practices such as teacher moderation (CERI, 2005; Little,

Gearhart, Curry & Kafka, 2003). One of the participants in this study felt that in addition to improving the quality of assessment decisions that are made during collaborative assessment sessions, critical reflection about assessment practices with supportive colleagues could enhance the fairness of teachers' individual practices. In a similar vein, Morgan and Watson (2002) note the benefits of collegial dialogue about assessment, but they point out that exploring differences can be more helpful than achieving consensus because it brings about an "awareness of the possibility of alternative interpretations, of the inadequacy of the evidence available, and of the need to seek other perspectives. Dialogue with colleagues can thus provide teachers with questions with which to interrogate their own judgments" (p.102). A working environment where collegial discussion about assessment is encouraged might lead, not only to more frequent communication among teachers, but also to more supportive interactions about some of the more sensitive aspects of assessment for learning. However, teachers need to be able to discuss fairness issues openly and honestly, and without threat of ostracization, if this benefit is to be realized. From this it seems that in relation to fair classroom assessment the needs of teachers and students may not be all that different. Just as a constructive learning environment makes fair assessment interactions possible, a supportive work environment may help teachers in making fairer assessment decisions. In the next chapter, I return to some of the issues discussed in order to consider the implications for classroom practice and teacher education, and the possibilities for further research.

## CHAPTER NINE: CONCLUSIONS

*The whole point of stories is not 'solutions' or 'resolutions' but a broadening and even a heightening of our struggles – with new protagonists and antagonists introduced, with new sources of concern or apprehension or hope, as one's mental life accommodates itself to a series of arrivals: guests who have a way of staying, but not necessarily staying put.*

Robert Coles (1989, p.129)

The purpose of this study is to understand teachers' practical wisdoms about fairness in classroom assessment. I was particularly interested in thinking about fairness in assessment for learning as it is, in principle, the primary purpose of classroom assessment. I set my research in the context of writing assessment in Ontario's publicly funded educational system, first, because of my own teaching experience, and second, because teachers tend to find writing assessment challenging. Before offering my closing thoughts in this last chapter, I identify the limitations and benefits I see for my work, the implications for classroom practice and teacher education, and recommendations for further research.

### Limitations and Benefits of the Research Project

Choices are necessarily made between competing goals, such as depth, breadth and feasibility, in the design of a research project. Although these choices may create limitations, they are not necessarily negative, and some may be seen as acceptable design consequences. I am aware of limitations in my work in four areas. These relate to my use of secondary data, the selection of participants and the use of their practical wisdom, the method of gathering information, and the independent nature of the project.

#### *Using Secondary Data*

Three limitations relate to the data from the SAIP Writing Teacher Questionnaire

(CMEC, 2002b) that I used as part of this study. First, it is now relatively old data for an area of classroom practice that is in transition. Evidence of a shift from a traditional assessment paradigm to assessment in a learning culture (Shepard, 2000) is seen to some degree in comparing OME policy documents from a decade ago to *Growing Success* (2008). As such, the results of the same questionnaire completed this year, with the same teachers responding, might be different than what they reported in 2002. Second, because the purpose of the 2002 SAIP was to assess student writing achievement, the sampling design focused on the selection of students. Consequently, the teachers who completed the questionnaire were not a random sample. However, the number of respondents in Ontario was sufficiently large ( $n = 415$ ) for the trends and patterns in the data to be informative. Third, none of the existing data focused specifically on fairness in assessment for learning. However, as Gorard (2002) points out, secondary data analysis is still helpful for understanding context and building further knowledge, and I did glean useful information from the process. By using secondary data in conjunction with information from a variety of sources, I was able to understand and describe the facets of the educational context that were relevant for my study. This description may also be useful for others who wish to study classroom assessment in Ontario.

#### *Using Teachers' Practical Wisdoms*

The participants in my study were experienced specialist English teachers who had volunteered as mentors for other teachers. As a result, their practical wisdoms cannot be considered representative of the practical wisdoms of all teachers in Ontario. However, the purpose of this study was not to report generally on teachers' thinking, but to use the practical wisdoms of certain teachers to better understand a particular quality in practice, and

the participants were thus purposefully selected. Patton (2002) points out that with a small number of purposefully selected participants “it becomes important to reconsider how design constraints may have affected the data available for analysis” (p.563). Gaining access to participants through a teachers’ network meant that I was working with volunteers from a pool that I developed based on years of teaching experience and professional qualifications. Had I been working with experienced specialist teachers who were nominated as “experts” or at least “moving pioneers” (Black et al., 2003, p.28) in the implementation of assessment for learning, the information I gathered may have focused more specifically on assessment for learning. However, the practical wisdoms about fairness in classroom assessment expressed by my participants reflect, as a result, a blurring of purposes that may be more typical in English Language Arts classrooms in Ontario than might have been revealed had the participants been nominated for their expertise in assessment for learning.

Relying on teachers’ practical wisdoms as a source of information for research opens the possibility of another limitation, which Shulman (1987/2004) refers to as one of the “embarrassments of practice” (p.264). Although he is a strong advocate for treating practical wisdom with respect, Shulman cautions that not all practitioners are wise, and “wise practitioners vary” (p.265). I believe that two assumptions need to be addressed here. First, as Shulman notes, we should not assume that “what is” and “what ought to be” (p.265) are one and the same. He recommends that attention be paid to variation and external factors in studying teachers’ practical knowledge. I explicitly dealt with external factors with my third research question, and in my analysis I considered issues raised by one participant as attentively as I did the commonalities voiced by all six. The second assumption I wished to

avoid in my work is one that Larabee (2003) points out, which is the assumption that insiders' knowledge is superior to other forms of knowledge in educational research. Approaching research from the perspective of critical pragmatism, I believe that contributions to knowledge about practice from both inside and outside the classroom are valuable. With this study, I aim to bring experienced specialist teachers' practical wisdoms into a theoretical dialogue that is underway among researchers about classroom assessment. The reason for this was not to canonize the words of the participants, but to consider their practical wisdoms in context, and in conjunction with existing literature. As teachers' practical wisdoms have been undervalued in educational research, I believe this work does make a unique contribution to the literature on fairness in classroom assessment.

#### *Eliciting Information for Research*

A limitation in using self-reported information in research is its lack of congruence with action (Wilks, 2004), and this is significant when the purpose of a study is to describe what occurs in a classroom. As this was not my intent, I did not observe teachers' practices in their classrooms while gathering information for this study. Consequently, I do not make generic claims about how fairness is actually realized in classrooms. When I undertook this work, many teachers in a variety of studies had already expressed concerns about their ability to assess student learning, and fairness issues in classroom assessment had emerged repeatedly in educational research (see Chapters 2 and 3). Observing practices in English Language Arts classrooms in Ontario might have resulted in similar findings, but I doubt that it would have led me to a better understanding of the concept of fairness, which was the purpose of my study. I believe that by calling on experienced teachers to share their practical

wisdoms about aspects of fairness that had been identified in the literature, this work builds on an existing body of research and offers additional insight.

However, asking participants to respond to specific issues illustrated in vignettes can lead to the same problem that occurs with closed response options on a survey. As Wilks (2004) explains, this creates a “circularity in the limited range of predetermined responses available” (p.83) to the respondents. Also, the effectiveness of vignettes as elicitation tools depends heavily on how carefully they have been crafted (Hughes & Huby, 2004). With these problems in mind, I included three different types of response opportunities in this study. The participants were asked to select a response to short items (i.e., 1-2 sentences), write open responses to longer vignettes on fairness issues (i.e., 4-6 sentences), and then elaborate orally on these and any other fairness issues. As Brookhart suggested might be the case for this type of research (in her critique as a discussant), some of the participants were willing to change their responses (e.g., from unethical to depends on circumstances) to the short items as we discussed them during their interviews. They also elaborated further on their responses to the vignettes, often with specific examples, but without altering the essential idea that they had originally expressed. As such, it seems that longer vignettes are more reliable as a means of eliciting information and generating discussion about the ethical dimension of classroom assessment. Based on my experience in conducting this research, though, I believe that follow-up interviews are absolutely essential with vignette research. The participants’ written responses were relatively short in comparison to the stories they told and the rich discussion we had during the interviews. However, I doubt that the interviews would have sufficed on their own because the written vignette questionnaire provided participants with the time and

privacy to think about the complex, and sometimes sensitive issues involved.

### *Contributing to a Dialogue*

Unlike collaborative projects that are undertaken in educational research with teams of researchers or partnerships within a research program, doctoral research is usually conducted by a sole researcher. Consequently, the instrument design, interviews, coding, analysis, and writing are all done by the same person, from the same perspective. While this may be beneficial for internal coherence, it ultimately results in a singular interpretation. To lend credibility to my interpretation, I began by building on existing ideas about fairness in the literature, and I engaged with other colleagues and researchers, as well as my participants in the research process. With this in mind, I do not cast my findings as permanent knowledge. Rather, I present them in a manner that is in keeping with my philosophical approach to knowledge as an informed contribution to an ongoing dialogue in the research literature.

### Practical Implications

This work has implications for the practice of fair classroom assessment, as well as for the initial education and professional development of teachers.

### *Fair Classroom Assessment for Learning*

When the participants in my study shared their practical wisdoms, they discussed many different aspects of fairness. The six aspects that emerged the most strongly in the context of this study involved establishing a constructive environment, offering students relevant opportunities, sharing learning expectations and assessment criteria, ensuring the appropriateness of an assessment for individual students, maintaining consistent treatment for a group of students, and reflecting about assessment interactions. Essentially, this view of fair

classroom assessment highlights human interaction, rather than methods or tools, which suggests that fairness is not a quality that can be achieved by fixed means for all situations. How then can teachers know if their assessment practices are fair? One approach that seems especially helpful for thinking about the practice of fair classroom assessment comes from Lane and Silver's (1999) ideas about large-scale performance assessment. They suggest that fairness should be determined as a matter of degree, like validity, with multiple sources of evidence. This means that no single practice can guarantee fairness, and that multiple aspects of an assessment should be considered to improve the degree of fairness. Consequently, the goal for teachers is not simply to develop a fair method or tool for classroom assessment, but to consider what aspects of an assessment practice make it fairer, and which practices might be fairest in different circumstances. Biesta & Burbules (2003) write that the "results of educational inquiries provide educators with a wider range of alternatives from which to select, in dealing with individual situations. That is all they can do. They enrich the educator's ability to judge" (p.80). I believe that it is important to emphasize that the aspects of fairness discussed in my study should not be thought of as items on a checklist, but as avenues for educators to considering as they aim for fairer classroom assessment practices.

Fairness is a quality that appears to be universally desirable for classroom assessment, but little is known about how it is manifested in different educational contexts. It may be that the aspects of fairness emphasized by the participants in my study resonate with other teachers, and will be transferable to other educational contexts. Some aspects of fairness that the participants in my study discussed, such as sharing learning expectations and assessment criteria with students, may be beneficial for the fairness of an assessment regardless of the

educational context in which it occurs. However, given the influence of external factors, it may be that some aspects of fairness are more important in some contexts than others. For example, ensuring the appropriateness of an assessment for individual learners may be less of an issue in learning environments where students are clustered according to their level of achievement rather than their date of birth (e.g., belt colours in martial arts). It may also be that some aspects of fairness are more important at different times because of the dynamic nature of classroom assessment. Specifically, it seems that some degree of equal treatment is needed to establish a constructive learning environment with students in a classroom, and students may be more likely to perceive differentiation for individual learners as fair if trusting and respectful relationships are nurtured over time. Ultimately, the implication of my work for classroom practice is that fair assessment interactions require much more than the implementation of externally mandated policies or principles; they require thoughtfulness. Decision making in classroom assessment interactions is rooted in the particulars of the circumstance (Birmingham, 2003; Dunne, 1993), and making fair decisions involves teachers in a reflective process that draws on their knowledge, beliefs, understandings and experience. In other words, fairness in classroom assessment depends on teachers' abilities to develop and use their practical wisdoms.

The connection between fairness and practical wisdom could be construed to mean that teachers' are on their own in making fair decisions. However, this project was undertaken with the belief that practical wisdom can be shared to some extent even though it is situated in the particulars of practice. Larabee (2003) points out that "where similarity exists, there is the possibility of finding practices that teachers can adopt or adapt to meet

their own pedagogical needs” (p.20). He rejects the “Lone Ranger” approach to teaching, and he embraces a “community of practice” that draws on multiple perspectives. This idea is similar to Earl’s (2004) suggestion that classroom assessment is improved with teacher moderation because it brings the “collective wisdom of all the people in the group to the exercise” (p.42). In this light, it seems that while fairness in classroom assessment depends on the particulars of an interaction, teachers do not have to make decisions in isolation, and the development and use of practical wisdom for fair decision making in classroom assessment can be supported within a school or system.

#### *Teacher Education and Professional Development*

In their research on classroom assessment practices in the United States, McMillan and Nash (2000) found that teachers’ assessment decisions were informed by past experience of what “worked best for them and their students” (p.21), and the influence of teacher education or professional development was notably “absent” (p.21). The participants in my study also seemed to have learned about classroom assessment primarily through experience. However, they also mentioned a range of professional activities, such as action research, curriculum writing, and assessment workshops, from which they had developed their knowledge. This may be reflective of a growing interest in the Ontario educational system in the pedagogical potential of classroom assessment. Although teacher education programs are still not mandated to provide pre-service teachers with instruction on classroom assessment (OCT, 1996), an additional qualifications course on student assessment and evaluation is under development by the OCT (2008b). At this point, though, there seems to have been very little emphasis on the ethical dimension of classroom assessment in either the initial teacher

education or professional development programs in Ontario. Hargreaves and Goodson (2003) note that the “ethical or moral knowledge that teachers need to inform their professional judgments” (pp.xi) has been generally overlooked in teacher education. This oversight may be due in part to the belief that moral-ethical knowledge is too personal or contextual to be taught, or that practical wisdom is acquired only through experience. For example, in writing about the relationship between human intelligence, knowledge, and wisdom, Sternberg (2000) notes that “one cannot teach particular courses of action that would be considered wise, regardless of circumstance” (p.636). Nonetheless, he concludes that “we need to turn our attention in schools not only to the development of knowledge . . . we need to turn our attention to the development of wisdom” (p.646). Put in the context of my work, this emphasizes the idea that teacher education and professional development programs should not focus on establishing the fairest course of action for each and every assessment situation, but on helping teachers develop the practical wisdom needed for fair decisions in the dynamics of classroom assessment.

The issue that arises for teacher education and professional development, given that practical wisdom is rooted in the particulars of practice, is ironically, one of technique. Specially, how can practical wisdom about fairness in classroom assessment be developed and shared? Although this was not a question asked at the outset of this study, it is one for which I can make several initial suggestions. We could begin, as Buzzelli and Johnston (2002) recommend, by acknowledging the beliefs and experiences that individuals “bring with them into teacher education programs” (p.132). For example, Whittington (1999) explains that she realized how key concepts became more meaningful for the teachers in her

educational measurement course when she had them write about an unfair assessment that they had personally experienced. Kessels and Korthangen (1996) discuss a similar approach in their effort to help pre-service teachers develop practical wisdom by examining issues encountered during field placements. They argue that the focus should not be on identifying solutions, but on helping pre-service teachers “become aware of the salient features of the experience” (p.335) so that they hone their perceptions and build further knowledge.

However, what these approaches all assume is that it is always possible for teachers to consciously identify a relevant issue from practice for discussion. Although they aim to help teachers understand what they already know, these approaches may also have the effect of focusing discussion on the more common aspects of assessment of learning, rather than some of the less obvious, but equally important fairness issues in the spontaneous interactions of assessment for learning.

A technique that might build on the approaches suggested by Buzzelli and Johnston (2002), Whittington (1999), and Kessels and Korthangen (1996), but that might prove more effective for encouraging practical wisdom about fairness in classroom assessment is the use of cases. Shulman (1996/2004) is a strong advocate of case methods in professional education because they are a way of “parsing experience so that practitioners can examine and learn from it” (p.465). Sternberg notes, furthermore, that studying cases “can help students to develop wisdom” (p.636). Cases may be particularly useful for discussing fairness in classroom assessment for the same reasons that vignettes are effective tools for research on practical ethics. Well-written vignettes can be engaging and sometimes “therapeutic” (Hughes & Huby, 2004, p.41), and they can be sufficiently complex to encourage discussion

about ethical issues where multiple responses are possible (Wilks, 2004). Kitchen, Cherubini, Smith, Goldblatt and Engemann (2008) engaged in a collaborative professional learning project in Ontario, and they report that the “participants strongly endorsed using case studies and acknowledged the benefits of this methodology. Perhaps the most transformative aspect of the sessions occurred when participants embraced perspectives other than their own” (p.7). The challenge in using cases about fair classroom assessment for the development of practical wisdom in teacher education and professional development programs in Ontario would be to produce cases that were current and relevant for a variety of learning environments. This would require ongoing collaboration with teachers in practice.

#### Recommendations for Further Research

Recommendations for further research emanating from this work fall basically into three groups, which relate to fairness as a quality of classroom assessment, the development of teachers’ practical wisdom about fair classroom assessment, and the external factors that impact fair classroom assessment.

##### *Fairness as a Quality of Classroom Assessment*

I began this study by looking at many different interpretations of fairness because it was not already clear, based on the existing literature, which might be most important in using classroom assessment for learning. One aspect of fairness that appears to be well accepted in both theory and practice is that information about the assessment process should be shared with students. This is a core practice in assessment for learning, and it is also an aspect of fairness in standardized testing. The participants in my study all felt that sharing learning expectations and assessment criteria would enhance the fairness of an assessment.

However, we still know very little about how this aspect of fairness is practiced in different classroom environments. When and how do teachers usually share information about learning expectations and assessment criteria, and to what degree do students understand and use this information? To capture the quality, effectiveness, and frequency of this practice, researchers might want to consider designs with complementary methods where classroom artifacts (e.g., rubrics, assignment instructions, course outlines) are analyzed in tandem with responses from students and teachers.

At this point, it would also be beneficial to focus more narrowly on developing a better understanding of the aspects of fairness that appear to be most challenging for teachers and students. Existing empirical literature shows a strong tension between equity and equality in classroom assessment. The participants in my study described accommodations or modifications they made to ensure that an assessment was equitable for a particular student, but they also stressed the importance of equal treatment among a group of students. While this may be, as one participant pointed out, an essential tension in classroom assessment, teachers seem to receive very little guidance about equitable assessment practices.

Assessment policies in Ontario indicate that accommodations and modifications should be made for students who are formally identified as exceptional learners, but students' needs are not always clearly defined. How fair are current assessment practices for the students who are not formally identified, such as atypical learners or the many English language learners in Ontario? Further research would certainly be helpful for understanding how teachers balance between differentiating for individual learners and maintaining the perception of fairness for a class. Given the variety of classroom environments in Ontario, teachers may be

approaching the tension between equality and equity in a myriad of ways, and making vastly different decisions with significant consequences for students. The question of feasibility also needs to be raised in investigating this tension in classroom assessment. While current literature encourages the idea that fairness is not sameness (e.g., OME, 2008; Wormelli, 2006), how many individual education plans are possible in a publicly-funded educational system? It is not evident that the degree of differentiation that is currently being promoted as educationally desirable is even feasible for students and teachers in Ontario.

The most important avenue I see for further research on fairness as an assessment quality relates to classroom dynamics. The role of power in peer and teacher-student relationships could certainly bear more investigation, especially as very little is currently known about the effects of student characteristics (e.g., gender, language, cultural background, special needs) on assessment interactions. Students' perceptions would be particularly helpful in this area. Some research about classroom assessment has involved students' perceptions (e.g. Brookhart, 2001; Bursuck et al., 1999; Cowie, 2005), but this could be expanded to focus more specifically on fairness issues in assessment for learning. Further research is also needed on teachers' perceptions of students in assessment interactions. For example, the tendency of teachers to 'pull' for some students but not others in classroom assessment raises serious concerns about what values or biases might underlie those decisions. When does knowledge about students' characteristics, learning careers, or their personal lives enhance the fairness of an assessment, and what are the circumstances where the reverse might be true? These questions could be approached from a variety of perspectives, but a critical lens would shed light on power dynamics. Existing work in critical

pedagogy could be used to frame inquiry into what students and teachers bring into play as they engage in, or resist assessment interactions. While my study touched on some these, more could be revealed by focusing specifically on what is often hidden in assessment interactions. Borrowing from Jackson and colleagues (1993), we need to pay more attention to the “elaborate amalgam of shared understandings, beliefs, assumptions, and presuppositions” that are “deeply embedded”(p.16) in classroom assessment. I believe that research designs involving simulations would be most productive for highlighting the elements and dynamics of this amalgam. Simulations could be either written vignettes or short recordings focusing specifically on power dynamics in assessment interactions. Again, these would need to be carefully crafted and reviewed to ensure that they elicited information about the amalgam at play in the interaction.

#### *Development of Teachers' Practical Wisdom about Fair Assessment*

The practical wisdom that the participants shared in this study seemed to draw on an internal network of different types of knowledge, beliefs, understanding, and characteristics that continued to develop through experience and learning. Although moral beliefs and ethical knowledge were embedded in their recommendations for fair classroom assessment, the participants sometimes became aware of this only as they tried to help me understand why or how they had come to a particular recommendation. In other words, explaining their thinking generated deeper reflection about what lay underneath it. While this implies that reflecting aloud with colleagues is beneficial, the logistics are less obvious. Under what conditions are colleagues comfortable discussing their moral beliefs? How can relationships between colleagues be systematically encouraged to permit honest reflection? Some of these

issues were tackled in a professional development project recently undertaken by a group of researchers in collaboration with the OCT (Kitchen et al., 2008). They brought beginning teachers and mentors together for case-based inquiry using the OCT's ethical standards (2008a) as a framework. A similar project could be designed for research purposes to focus specifically on the ethical dimension of classroom assessment. However, I believe that it would be important in this type of endeavour to maintain a critical stance in considering how the use of existing standards might spark or shutter reflection and discussion among colleagues about fairness in classroom assessment.

McLaughlin (1999/2007) writes that the value of reflection in teaching has an obvious and "intuitive appeal" (p.357) that makes it difficult to contest. Nonetheless, he argues that we need to "go beyond 'the reflective teacher' as a slogan" (p.365), and he suggests that the Aristotlean concept of practical wisdom provides fertile ground for thinking about teaching. As noted earlier, calls for teachers to reflect about their assessment practices are often seen in the literature, and a few of the participants in my study felt that reflection was an important aspect of fairness in classroom assessment. Following Mclaughlin, I believe that the concept of practical wisdom is ideal as a framework for reflection. Reflecting about practical wisdom takes teachers beyond technical issues to consider other forms of knowledge, beliefs, personal characteristics, or experiences that might be informing assessment decisions and interactions in the classroom. Consequently, further research on how teachers develop practical wisdom is needed, especially in relation to fair classroom assessment. A starting point would be to determine current practices in teacher education program. To what extent is classroom assessment included in the curriculum for teacher education, and does it involve

discussion about the ethical dimension as well as the techniques of assessment? How are pre-service teachers supported in the development of the professional judgment required for fair assessment decisions? These questions could also be tackled from a variety of perspectives, but beginning with a broad look might permit information about the benefits and drawbacks of different approaches to be shared within the teacher education community.

### *External Factors Impacting Fair Classroom Assessment*

The participants in my study were asked to identify any factors that might hinder or facilitate the fairness of writing assessment in English Language Arts classes in Grades 7 to 12 in Ontario. They described student-related factors the most frequently, but they also felt that colleagues, assessment policies, professional development opportunities and time constraints could influence the fairness of classroom assessment. As time constraints appear to be ubiquitous, we need a clearer picture of how teachers juggle time in their assessment practices. I wonder what teachers might do differently if they had more time to assess student learning? The narratives of purposefully selected teachers would provide insight about the impact of time on the quality of assessment decisions and interactions, especially in different classroom environments and educational contexts.

The results of my study varied somewhat from the existing literature, in terms of the factors that influence classroom assessment, because parents and principals were mentioned much less frequently than students. This makes sense given that I focused on teachers' decisions about students and students' writing, and information about external factors tends to come from research more broadly focused on the implementation of assessment programs. However, there are many questions relating to the people, beyond the teachers and students

directly involved in classroom interactions, who might influence the fairness of assessment. To what extent do collaborative assessment practices, such as teacher moderation, improve the quality of assessment decisions? What role do principals need to play for these types of initiatives to be successful? How could principals become more active in supporting fair classroom assessment? At this point, communication with parents about classroom assessment also seems quite limited. How do parents interpret the concept of fairness, and to what degree do parents recognize the benefits of assessment for learning? It would be interesting to elicit responses from students, parents, principals, and teachers in a variety of classroom environments using the same vignettes to explore the differences in their responses. Ultimately, if the potential of assessment for learning is to be realized in practice, we need to engage in further research to better understand how to support the fairness of the interactions and decisions in the process.

### Closing Thoughts

Through this research project I have aimed to develop a better understanding of fairness in classroom assessment. Although fairness is a desirable quality in educational assessment, guidance for classroom practice is still limited, especially in relation to the use of assessment for learning. Fairness is particularly complex in assessment for learning because it involves human relationships that develop in varied and changing circumstances. As such, fairness cannot be prescribed in universal terms for all classrooms. By this I do not wish to suggest that discussion within educational communities about the ethical dimension of assessment is unnecessary. On the contrary, I believe that ongoing discussion is absolutely critical to encourage fairer classroom assessment. While it has long been evident that

teachers' formal assessment decisions can have major social, emotional and academic consequences for students, the practice of assessment for learning may have similarly significant effects. If assessment for learning is to meet its pedagogical potential, the fairness of assessment decisions by students and teachers, and the interactions between them must be considered. By highlighting experienced teachers' practical wisdoms, I hope that my research illuminates possibilities for other educational stakeholders in thinking about fairness. Fairness in classroom assessment rests in teachers' hands, but ultimately it concerns us all.

## References

- Airasian, P. W. (2005). *Classroom Assessment: Concepts and Applications* (5<sup>th</sup> ed). New York: McGraw-Hill.
- Allal, L., & Pelgrims Ducrey, G. P. (2000). Assessment of- or in- the zone of proximal development. *Learning and Instruction*, 10, 137-152.
- Anfara, V. A., Brown, K. M., & Mangione, T. L. (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher*, October, 28-38.
- Ariogul, S. (2006). Exploring of teacher knowledge base: a qualitative study of English as a foreign language (efl) teachers' practical knowledge in Turkey. *Proquest Digital Dissertations*, UMI No. AAT3223068.
- Aristotle (trans. 1925/1975). *The nicomachean ethics of Aristotle*. (D. Ross). London: Oxford University Press.
- Assessment Reform Group (2002). Assessment for learning: 10 principles. Retrieved 17 April 2008 from <http://www.assessment-reform-group.org/publications.html>
- Badger, E. (1999). Finding one's voice: A model for more equitable assessment. In A. L. Nettles and M. T. Nettles (Eds.), *Measuring up: Challenges minorities face in educational assessment* (pp.53-69). Norwell, MA: Kluwer Academic Publishers.
- Barnes, S. (1985). A study of classroom pupil evaluation: The missing link in teacher education. *Journal of Teacher Education*, 36 (4), 46-49.
- Barter, C., & Renold, E. (1999). The use of vignettes in qualitative research. *Social Research Update*, 25, 1-7. Retrieved 08 February 2008 from <http://sru.soc.surrey.ac.uk/SRU25.html>
- Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85 (5), 536-553.
- Berliner, D. C. (2004). Describing the behaviour and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24 (3), 200-212.
- Bernstein, R. J. (1985). *Beyond objectivism and relativism: Science, hermeneutics and praxis*. Philadelphia, PA: University of Pennsylvania Press.
- Bernstein, R. J. (1986). *Philosophical profiles: Essays in a pragmatic mode*. Philadelphia, PA: University of Pennsylvania Press.

- Biesta, G. (2007). Why 'what works' won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory*, 57 (1), 1-22.
- Biesta, G., & Burbules, N. C. (2003). *Pragmatism and educational research*. Lanham, MD: Rowman & Littlefield Publishers, Inc.
- Birmingham, C. (2003). Practicing the virtue of reflection in an unfamiliar cultural context. *Theory into Practice*, 42 (3), 187-194
- Black, A. L., & Halliwell, G. (2000). Accessing practical knowledge: How? Why? *Teaching and Teacher Education*, 16, 103-115.
- Black, P. (2001). Dreams, Strategies and Systems: portraits of assessment past, present and future. *Assessment in Education*, 8 (1), 65-85.
- Black, P. & Harrison, C. (2001). Feedback in questioning and marking: The science teacher's role in formative assessment. *School Science Review*, 82 (301), 55-61.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, England: Open University Press.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5 (1), 7-74.
- Black, P. & Wiliam, D. (2003). 'In praise of educational research': Formative assessment. *British Educational Research Journal*, 29 (5), 623-637.
- Black, P., & Wiliam, D. (2005). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices. *The Curriculum Journal*, 16 (2), 249-261.
- Black, P., & Wiliam, D. (2006a). Assessment for learning in the classroom. In John Gardner (Ed.), *Assessment and learning*, pp.9-25. Thousand Oaks: Sage Publications.
- Black, P., & Wiliam, D. (2006b). Developing a theory of formative assessment. In John Gardner (Ed.), *Assessment and learning*, pp.81-100. Thousand Oaks, CA: Sage Publications.
- Blanchard, A. W. (2005). Four eleventh-grade English teachers' attitudes about writing and their teaching of writing in the context of a state-wide direct writing assessment. *Proquest Digital Dissertations*, UMI No. AAT 3166972.
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York; McGraw-Hill.

- Bol, L., Stephenson, P., O'Connell, A., & Nunnery, J. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research*, 91 (6), 323-330.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36 (2), 81-89.
- Borko, H., Davinroy, K. H., Bliem, C. L., & Cumbo, K. B. (2000). Exploring and supporting teacher change: Two third-grade teachers' experiences in a mathematics and literacy staff development project, *Elementary School Journal*, 100(4), 273-306.
- Borman, K. M., Clarke, C., Cotner, B., & Lee, R. (2006). Cross-case analysis. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research*, pp.3-31. Mahwah, NJ: Lawrence Erlbaum Associates; Washington, DC: American Educational Research Association.
- Brackenreed, D. G. (2006). English teachers' perceptions of standards tests, the efficacy of testing accommodations and appropriateness for students with special needs. *Proquest Digital Dissertations*, UMI No. AAT NR20280
- Bredo, E. (2006). Philosophies of educational research. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research*, pp.3-31. Mahwah, NJ: Lawrence Erlbaum Associates; Washington, DC: American Educational Research Association.
- Bright, R. (2007). *Write through the grades: Teaching writing in secondary schools*. Winnipeg, MA: Portage & Main Press.
- Brighton, C. (2001). Internal factors that influence teacher change: Teachers' beliefs and conceptions. *Digital Dissertations*, UMI No. AAT 3003985.
- Briscoe, C., & Wells, E. (2002). Reforming primary science assessment practices: A case study of one teacher's professional development through action research. *Science Education*, 86 (3), 417-435.
- Bromley, K. (2007). Assessing student writing. In J. R. Paratore & R. L. McCormack (Eds.), *Classroom literacy assessment: Making sense of what student know and do*. New York: The Guilford Press.
- Brookfield, S. (2001). Unmasking power: Foucault and adult learning. *Canadian Journal for Studies in Adult Education*, 15(1), 1-23.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30 (2), 123-142.

- Brookhart, S. M. (1994). Teachers' grading practices: Practice and theory. *Applied Measurement in Education*, 7(4), 279-301.
- Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education: Principles, Policy & Practice*, 8 (2), 153-169.
- Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational Measurement: Issues and Practice*, 22 (4), 5-12.
- Brookhart, S. M. (2004). Classroom Assessment: Tensions and intersections in theory and practice. *Teachers College Record*, 106 (3), 429-458.
- Bullough, Jr., R. V. (2008). *Counternarratives: Studies of teacher education and becoming and being a teacher*. New York: State University of New York Press.
- Bulterman-Bos, J., Terwel, J., Verloop, N., & Wardekker, W. (2002). Observation in teaching: Toward a practice of objectivity. *Teachers College Record*, 104(6), 1069-1100.
- Bursuck, W. D., Munk, D. D., & Olson, M. (1999). The fairness of report card grading adaptations: What do students with and without learning disabilities think? *Remedial and Special Education*, 20, 84-92.
- Butterfield, S., Williams, A., & Marr, A. (1999). Talking about assessment: Mentor-student dialogues about pupil assessment in initial teacher training. *Assessment in Education: Principles, Policy & Practice*, 6 (2), 225-239.
- Buzzelli, C. A., & Johnston, B. (2002). *The moral dimensions of teaching: Language, power, and culture in classroom interaction*. London: Routledge Falmer.
- Camilli, G. (2006). Test fairness. In R.L. Brennan (Ed.), *Educational Measurement*, 4<sup>th</sup> ed., pp.221-256. Westport, CT: National Council on Measurement in Education, American Council on Education and Praeger Publishers.
- Campbell, E. (2003). *The ethical teacher*. Maidenhead, England: Open University Press.
- Carr, M., & Claxton, G. (2002). Tracking the development of learning dispositions. *Assessment in Education: Principles, Policy & Practice*, 9(1), 9-37.
- Centre for Educational Research and Innovation. (2005). *Formative assessment: Improving learning in secondary classrooms*. Retrieved 18 March 2005 from the Organisation for Economic Co-operation and Development Web site: <http://www.oecdbookshop.org/oecd/display.asp?sf1=identifiers&st1=962005021P1>

- Cherryholmes, C. H. (1988). *Power and criticism: Poststructural investigations in education*. New York: Teachers' College Press.
- Cheng, L. Fox, J. & Zheng, Y. (2007). Student accounts of the Ontario Secondary School Literacy Test: A case for validation. *The Canadian Modern Language Review*, 64 (1), 60-98.
- Cheng, L., Klinger, D. A., & Zheng, Y. (2007). The challenges of the Ontario Secondary School Literacy Test for second language students. *Language Testing*, 24 (2), 185-208.
- Cheung, D. (2002). Refining a stage model for studying teacher concerns about educational innovations. *Australian Journal of Education*, 46 (3), 305-322.
- Choi, S. (2007). Elementary teachers' beliefs and practical knowledge about teaching Science as inquiry: The effects of an inquiry-based elementary Science course. *Proquest Digital Dissertations*, UMI No. 3272579
- Clandinin, D. J., & Connelly, F. M. (1996). Teachers' professional knowledge landscapes: Teachers stories - stories of teachers - school stories - stories of schools. *Educational Researcher*, 25 (3), 24-30.
- Clandinin, D. J., & Murphy, M. S. (2007). Looking ahead: Conversations with Elliot Mishler, Don Polkinghorne, and Amia Lieblich. In D. J. Clandinin (Ed.), *Handbook of narrative inquiry: Mapping a methodology*, pp.632-650. Thousand Oaks: Sage Publications.
- Clandinin, D. J., & Rosiek, J. (2007). Mapping a landscape of narrative inquiry. In D. J. Clandinin (Ed.), *Handbook of narrative inquiry: Mapping a methodology*, pp.35-75. Thousand Oaks: Sage Publications.
- Clark, C., Chow-Hoy, T. K., Herter, R. J., & Moss, P. A. (2001). Portfolios as sites of learning: Reconceptualizing the connections to motivation and engagement. *Journal of Literacy Research*, 33 (2), 221-241.
- Clarke, A., & Robertson, A. (2001). Lifting a corner of the research rug: A case for meta-interviews in qualitative studies. *Teaching and Teacher Education*, 17, 773-782.
- Coles, R. (1989). *The call of stories: Teaching and the moral imagination*. Boston, MA: Houghton Mifflin Company.
- Colnerud, G. (2006). Teacher ethics as a research problem: Syntheses achieved and new issues. *Teachers and Teaching: Theory and Practice*, 12 (3), 365-385.

- Colton, D., & Covert, R. W. (2007). *Designing and constructing instruments for social science research and evaluation*. San Francisco, CA: Wiley/Jossey-Bass.
- Connelly, F. M., & Clandinin, D. J. (1986). On narrative method, personal philosophy, and narrative unities in the story of teaching. *Journal of Research in Science Teaching*, 23 (4), 293-310.
- Connelly, F. M., & Clandinin, D. J. (2006). Narrative inquiry. In J.L. Green, G. Camilli, & P.B. Elmore (Eds.), *Handbook of complementary methods in education research*, pp.375-385. Mahwah, NJ: Lawrence Erlbaum Associates; Washington, DC: American Educational Research Association.
- Connelly, F. M., Clandinin, D. J., & He, M. F. (1997). Teachers' personal practical knowledge on the professional knowledge landscape. *Teaching and Teacher Education*, 13 (7), 665-674.
- Connolly, M. R. (2005). Exploring cases of practical wisdom (phronesis) in postsecondary teaching. *Proquest Digital Dissertations*, UMI No. AAT 3167788.
- Council of Ministers of Education, Canada. (2002a). *Responses to Teacher Questionnaire/ SAIP 2002 - Writing Ontario Anglophones*. Toronto, ON: Author.
- Council of Ministers of Education, Canada. (2002b). *SAIP Writing Assessment 2002*. [Data file]. Toronto, ON: Author.
- Council of Ministers of Education, Canada (2002c). *SAIP Writing Assessment III: Teacher Questionnaire*. Toronto, ON: Author.
- Council of Ministers of Education, Canada (2009). *Programs & Initiatives: Assessment*. Retrieved 26 June 2009 from [www.http://www.cmec.ca](http://www.cmec.ca)
- Cowie, B. (2005). Student commentary on classroom assessment in science: A sociocultural interpretation. *International Journal of Science Education*, 27 (2), 199-214.
- Cowie, B., & Bell, B. (1999). A model of formative assessment in Science education. *Assessment in Education: Principles, Policy & Practice*, 6 (1), 101-116.
- Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative reserach*. Upper Saddle River, NJ: Pearson Education.
- Creswell, J. W. & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice*, 39(3), 124-131.
- Crocker, L. (2003). Teaching for the test: Validity, fairness, and moral action. *Educational*

*Measurement: Issues and Practice*, 22 (3), 5-11.

Crooks, T. (2001, September). *The validity of formative assessments*. Paper presented at the annual meeting of the British Educational Research Association, Leeds, England.

Cross, L. H., & Frary, R. B. (1996). *Hodgepodge grading: Endorsed by students and teachers alike*. Paper presented at the Annual Meeting of the National Council on Measurement in Education. New York, April 9-11.

Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher*, 19 (1), 3-13.

Dann, R. (2002). *Promoting assessment as learning: Improving the learning process*. London: RoutledgeFalmer.

DeLain, M. T. (1995/2005). Equity and performance-based assessment: An insider's view. In S. J. Barrentine & S. M. Stokes (Eds.), *Reading assessment: Principles and practices for elementary teachers*, pp.52-55. Newark, DE: International Reading Association.

Delandshere, G. (2001). Implicit theories, unexamined assumptions and the status quo of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 8 (2), 113-133.

Delandshere, G., & Jones, J. H. (1999). Elementary teachers' beliefs about assessment in mathematics: A case of assessment paralysis. *Journal of Curriculum and Supervision*, 14 (3), 216-240.

De Waal, C. (2005). *On pragmatism*. Belmont, CA: Thomson Wadsworth.

Dixon, H., & Williams, R. (2003). Teachers' understanding and use of formative assessment in literacy learning. *New Zealand Annual Review of Education*, 12. Retrieved 27 June 2005 from the Te Kete Ipurangi Online Learning Center Web site: [www.tki.org.nz/r.assessment/research/research20\\_1\\_e.php](http://www.tki.org.nz/r.assessment/research/research20_1_e.php)

Drummond, M. J. (2003). *Assessing children's learning*. London: David Fulton Publishers.

Duncan, C. R. & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *Alberta Journal of Educational Research*, 53 (1), 1-21.

Dunne, J. (1993). *Back to the rough ground: 'Phronesis' and 'techne' in modern philosophy and in Aristotle*. Notre Dame, IN: University of Notre Dame Press.

Durst, S., & Plourde, G. (2008, January 28). *Memorandum to Directors of Education*. [Subject: Growing Success. Assessment, Evaluation and Reporting: Improving Student Learning]. Retrieved on 24 April 2009 from <http://cal2.edu.gov.on.ca>

- Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press/Sage Publications.
- Earl, L. M. (2004). Collecting the evidence. *Curriculum Bulletin*, 2 (2), 41-43. Retrieved 28 April 2009 from <http://www.npbs.ca/06-elements/articles-clues.pdf>
- Ecclestone, K., & Pryor, J. (2003). 'Learning careers' or 'assessment careers'? The impact of assessment systems on learning. *British Educational Research Journal*, 29 (4), pp.471-487.
- Education Quality and Accountability Office. (2007). *Programme for International Student Assessment (PISA), 2006: Highlights of Ontario Results*. Retrieved on 4 December 2008 from [http://www.eqao.com/pdf\\_e/07/07P105e.pdf](http://www.eqao.com/pdf_e/07/07P105e.pdf)
- Education Quality and Accountability Office. (2008a). *School, Board and Provincial Results*. Retrieved December 10, 2008 from <http://www.eqao.com>
- Education Quality and Accountability Office. (2008b). *Ontario Student Achievement: EQAO's Provincial Report on the Results of the 2007-2008 Ontario Secondary School Literacy Test*. Retrieved December 10, 2008 from <http://www.eqao.com>
- Educational Testing Service. (2002). *Standards for quality and fairness*. Retrieved 29 July 2006 from the Educational Testing Service Web site at: [http://ets.org/media/About\\_ETs](http://ets.org/media/About_ETs)
- Eggen, A. B. (2004). *Alfa and omega in student assessment: Exploring identities of secondary school science teachers*. Oslo, Norway: Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo.
- Elbaz, F. L. (1980). The teacher's practical knowledge: A case study. *Proquest Digital Dissertations*, UMI # AAT NK47038
- Elbaz, F. L. (1992). Hope, attentiveness and caring for difference: The moral voice in teaching. *Teaching and Teacher Education*, 8 (5/6), 421-432.
- Elementary Teachers' Federation of Ontario. (2008). *Issues in Education: EQAO Testing*. Retrieved 10 December 2008 from <http://www.etfo.ca>
- Elwood, J. (2006). Formative assessment: possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy & Practice*, 13 (2), 215-232.
- Engemann, J. F., & Gallagher, T. (2006). The conundrum of classroom writing assessment. *Brock Education*, 15 (2), 33-44.

- Erzberger, C., & Kelle, U. (2003). Making inferences in mixed methods: The rules of integration. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research*, pp.457-488. Thousand Oaks, CA: Sage Publications.
- Feldt, L. S. (1962). The reliability of measures of handwriting quality. *Journal of Educational Psychology*, 53 (6), 288-292.
- Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. *Review of Research in Education*, 20, 3-56.
- Fenwick, T., & Parsons, J. (2000). *The art of evaluation: A handbook for educators and trainers*. Toronto: Thompson Educational Press.
- Field, J. C., & Latta, M. M. (2001). What constitutes becoming experienced in teaching and learning? *Teaching and Teacher Education*, 17, 885-895.
- Finklestein, I. E. (1913). *The marking system in theory and practice*. Baltimore, MD: Warwick & York, Inc.
- Flyvbjerg, B. (2004). Five misunderstandings about case-study research. In C. Seale, G. Gobo, J. F. Gubrium, & D. Silverman (Eds.), *Qualitative research practice*, pp.420-434. London: Sage Publications.
- Fox, J., & Cheng, L. (2007). Did we take the same test? Differing accounts of the Ontario Secondary School Literacy Test by first and second language test-takers. *Assessment in Education: Principles, Policy & Practice*, 14 (1), 9-26.
- Frankfort-Nachmias, C. & Leon-Guerrero, A. (2006). *Social statistics for a diverse society*. Thousand Oaks, CA: Pine Forge Press.
- Frey, N., Fischer, D. & Hernandez, T. (2003). "What's the gist?" Summary writing for struggling adolescent writers. *Voices from the Middle*, 11 (2), 43-49.
- Gardner, J. (2006). Assessment for learning: A compelling conceptualization. In J. Gardner (Ed.), *Assessment and learning* (pp.197-204). Thousand Oaks, CA: Sage Publications.
- Giddens, S. F. (2001). Factors that affect teachers' practices in classrooms with students with disabilities. *Proquest Digital Dissertations*, UMI No. AAT NQ63665
- Gilman, D. A., & Kiger, S. (2003). Should we try to keep class sizes small? *Educational Leadership*, 60 (7), 80-85.
- Giordano, G. (2005). *How testing came to dominate American schools: The history of*

- educational assessment*. New York: Peter Lang Publishing.
- Gipps, C. (1994). *Beyond testing: Toward a theory of educational assessment*. London: The Falmer Press.
- Gipps, C. (1999). Chapter 10: Socio-cultural aspects of assessment. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education*, 24, 355-392.
- Gipps, C. (2005). Commentary on "The idea of testing: Psychometric and sociocultural perspectives". *Measurement*, 3 (2), 103-107.
- Glaser, B. G. & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago: Aldine.
- Gorard, S. (2002). The role of secondary data in combining methodological approaches. *Educational Review*, 54 (3), 232-237.
- Government of Ontario (2008). *Education Facts: Ministry of Education*. Retrieved 01 December 2008 from <http://www.edu.gov.on.ca/eng/educationFacts.html>
- Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through preservice teacher education. *Teaching and Teacher Education*, 21, 607-621.
- Green, S. K., Johnson, R. L., Kim, D.-H., & Pope, N.S. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, 23, 999-1011.
- Gummer, E. S. (2000). Rhetoric and reality: Congruence between the knowledge and practice of assessment of the classroom science teacher and the reform initiative of the National Science Education Standards. *Proquest Digital Dissertations*, UMI No. AAT 3018203.
- Guskey, T. R. (2005, April). Formative classroom assessment and Benjamin S. Bloom: Theory, research, and implications. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
- Hall, K., & Harding, A. (2002). Level descriptions and teacher assessment in England: Towards a community of assessment practice. *Educational Research*, 44 (1), 1-15.
- Hand, B., & Prain, V. (2002). Teachers implementing writing-to-learn strategies in junior secondary science: A case study. *Science Education*, 86, 737-755.
- Hargreaves, A., Earl, L., & Schmidt, M. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal*, 39 (1), 69-95.

- Hargreaves, A. & Goodson, I. (2003). Forward. In E. Campbell, *The ethical teacher*, (pp.ix-xiii). Maidenhead, England: Open University Press.
- Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning*, pp.103-117. Thousand Oaks, CA: Sage Publications Ltd.
- Harlen, W. (2007). *Assessment of learning*. London, England: Sage Publications, Ltd.
- Harper, D. (November, 2001). *Online etymology dictionary*. Retrieved 12 July 2008 from <http://www.etymonline.com>
- Harris, T. L., & Hodges, R. E. (1995). *The literacy dictionary: The vocabulary of reading and writing*. Newark, DE: International Reading Association.
- Hartrick Doane, G. & Varcoe, C. (2009). Toward compassionate action: Pragmatism and the inseparability of theory/practice. In P. G. Reed & N. B. Crawford Shearer (Eds.), *Perspectives on nursing theory*, pp.111-121. Philadelphia, PA: Lippincott Williams & Wilkins.
- Hayward, L., Priestley, M., & Young, M. (2004). Ruffling the calm of the ocean floor: Merging practice, policy and research in assessment in Scotland. *Oxford Review of Education*, 30 (3), 397-415.
- Heldke, L. (1988). Recipes for theory making. *Hypatia*, 3 (2), 15-29.
- Helms, J. E. (2006). Fairness is not validity or cultural bias in racial-group assessment: A quantitative perspective. *American Psychologist*, 61 (8), 845-859.
- Hills, J. R. (1991). Apathy concerning grading and testing. *Phi Delta Kappan*, 72 (7), 540-545.
- Hodgen, J., & Marshall, B. (2005). Assessment for learning in English and mathematics: A comparison. *The Curriculum Journal*, 16 (2), 153-176.
- Hodgkinson, C. (1991). *Educational leadership: The moral art*. New York: State University of New York Press.
- Horn, R. A. (2002). *Understanding educational reform: A reference handbook*. Santa Barbara, CA: ABC-CLIO, Inc.
- Hughes, G. J. (2001). *Routledge philosophy guidebook to Aristotle on ethics*. London: Routledge.

- Hughes, R., & Huby, M. (2004). The construction and interpretation of vignettes of social research. *Social Work & Social Sciences Review*, 11(1), 36-51.
- International Association for the Evaluation of Educational Achievement. (December 2007). *Progress in International Reading Literacy Study (2006): Ontario Report*. Retrieved on 04 December 2008 from the Education Quality and Accountability Office Web site at: <http://www.eqao.com/NIA/PIRLS/PIRLS.aspx?Lang=E>
- Jackson, P. W., Boostrom, R. E., & Hansen, D. T. (1993). *The moral life of schools*. San Francisco, CA: John Wiley & Sons.
- Jeranyama, L. N. (2001). Reshaping teacher thinking, planning and practice using embedded assessment: Case studies of three middle school science teachers. *Proquest Digital Dissertations*, UMI No. 3040263.
- Jimenez, R. T. (2004/2005). More equitable literacy assessments for Latino students. In S. J. Barrentine & S. M. Stokes (Eds.). *Reading assessment: Principles and practices for elementary teachers*, pp.49-55. Newark, DE: International Reading Association.
- Joint Advisory Committee. (1993). *Principles for fair student assessment practices for education in Canada*. Edmonton, Alberta, Canada: University of Alberta.
- Joint Committee on Standards for Teacher Competence in Educational Assessment of Students. (1990). *Standards for teacher competence in educational assessment of students*. American Federation of Teachers, National Council on Measurement in Education & National Education Association.
- Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Joint Committee on Standards for Educational Evaluation. (2003). *The student evaluation standards*. Thousand Oaks, CA: Corwin Press, Inc.
- Katz, S., Earl, L., & Olson, D. (2001). The paradox of classroom assessment: A challenge for the 21<sup>st</sup> century. *McGill Journal of Education*, 36(1), 13-26.
- Kearns, L. M. (2008). Equity, literacy testing and marginalized youth: The social construction of "illiterate" identities. *Proquest Digital Dissertations*, UMI No. AAT NR39807.
- Kern, A. L. (2007). Novice teacher action: Exploring secondary science teachers' actions

- through practical reasoning. *Proquest Digital Dissertations*, UMI No. 3287819.
- Kessels, J. P., & Korthagen, F. A. (1996). The relationship between theory and practice: Back to the classics. *Educational Researcher*, 17-23.
- Kitchen, J., Cherubini, L. Smith, D. Goldblatt, P., & Engemann, J. (2008). *Learning from experience: Supporting beginning teachers and mentors*. Retrieved 25 June 2009 from www.oct.ca
- Lam, T. C. M. (1995). Fairness in performance assessment. *ERIC Digest (No. ED391982)*.
- Lane, S., & Silver, E. A. (1999). Fairness and equity in measuring student learning using a mathematics performance assessment: Results from the Quasar Project. In A. L. Nettles & M. T. Nettles (Eds.), *Measuring up: Challenges minorities face in educational assessment* (pp.97-120). Norwell, MA: Kluwer Academic Publishers.
- Larabee, D. F. (2003). The peculiar problems of preparing educational researchers. *Educational Researcher*, 32 (4), 13-22.
- Laucius, J. (2009, April 22). Students aren't failing, but system is, teachers say. *The Ottawa Citizen*, pp. A1-A2.
- Laveault, D., & Miles, C. (2002). *The study of individual differences in the utility and validity of rubrics in the learning of writing ability*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1-5.
- Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning? *Assessing Writing*, 12, 180-198.
- Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment: Minute by minute, day by day. *Educational Leadership*, 63 (3), 18-24.
- Leibowitz, M. (2000). The work ethic and the habits of mind. In A. L. Costa & B. Kallick (Eds.), *Discovering & Exploring Habits of Mind*, (pp.62-78). Alexandria, VA: Association for Supervision and Curriculum Development.
- Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.). *Handbook of Qualitative Research* (2<sup>nd</sup> ed.), pp.191-215. Thousand Oaks, CA: Sage Publications.
- Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8<sup>th</sup> ed.). Upper Saddle River, NJ: Pearson Education/Prentice-Hall, Inc.
- Little, J. W., Gearhart, M., Curry, M., & Kafka, J. (2003). Looking at student work for

- teacher learning, teacher community, and school reform. *Phi Delta Kappan*, 85 (3), 184-192.
- Lock, C., & Munby, H. (2000). Changing assessment practices in the classroom: A study of one teacher's challenge. *Alberta Journal of Educational Research*, 46 (3), 267-79.
- Lovett, S., & Gilmore, A. (2003). Teachers' learning journeys: The quality learning circle as a model of professional development. *School Effectiveness and School Improvement*, 14 (2), 189-211.
- Luce-Kapler, R., & Klinger, D. (2005). Uneasy writing: The defining moments of high-stakes literacy testing. *Assessing Writing*, 10, 157-173.
- Mabry, L., Poole, J., Redmond, L., & Schultz, A. (2003). Local impact of state testing in Southwest Washington. *Education Policy Analysis Archives*, 11 (22). Retrieved July 28, 2004 from <http://epaa.asu.edu/epaa/v11n22>.
- Margolis, J. (2006). Introduction: Pragmatism, retrospective, and prospective. In J. R. Shook and J. Margolis (Eds.), *A companion to pragmatism*, pp.1-10. Oxford: Blackwell Publishing.
- Markus, J. (2002). Student assessment and evaluation in studio art. *Proquest Digital Dissertations*, UMI No. AAT NQ69131
- Marshall, C., & Rossman, G. B. (1999). *Designing qualitative research* (3<sup>rd</sup> ed.). Thousand Oaks, CA: Sage Publications.
- Maxcy, S. J. (2003). Pragmatic threads in mixed-method research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research*, pp.51-89. Thousand Oaks, CA: Sage Publications.
- Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage Publications.
- McDonald, M. (2002). The perceived role of diploma examinations in Alberta, Canada. *Journal of Educational Research*, 96 (1), 21-36.
- McDonald, B. & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education*, 10 (2), 209-220.
- McLaren, P. (2003a). Critical pedagogy: A look at the major concepts. In A. Darder, M. Baltodano, & R. R. D. Torres (Eds.), *The critical pedagogy reader*, pp. 69-96. New

York: RoutledgeFalmer/Taylor Francis Books, Inc

- McLaren, P. (2003b). Revolutionary pedagogy in post-revolutionary times: Rethinking the political economy of critical education. In A. Darder, M. Baltodano, & R. R. D. Torres (Eds.), *The critical pedagogy reader*, pp. 151-184. New York: RoutledgeFalmer/Taylor Francis Books, Inc.
- McLaughlin, T. H. (1999/2007). Beyond the reflective teacher. In R. Curren (Ed.), *Philosophy of education: An anthology*, pp.357-366. Oxford: Blackwell Publishing.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32.
- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, 22 (4), 34-43.
- McMillan, J. H. (2007). *Classroom assessment: principles and practice for effective teaching* (4<sup>th</sup> ed.). Boston, MA: Pearson Education, Inc.
- McMillan, J. H., & Nash, S. (2000). *Teacher classroom assessment and grading practices decision making*. Paper presented at the annual meeting of the National Council on Measurement in Education. New Orleans, LA: April 25-27.
- McNair, S., Bhargava, A., Adams, L., Edgerton, S., & Kypos, B. (2003). Teachers speak out on assessment practices, *Early Childhood Education Journal*, 31 (1), 23-31.
- Meijer, P. C., (1999). *Teachers' practical knowledge: Teaching reading comprehension in secondary education*. Leiden, Netherlands: Universiteit Leiden.
- Meijer, P. C., Verloop, N., & Beijaard, D. (1999). Exploring language teachers' practical knowledge about reading comprehension. *Teaching and Teacher Education*, 15, 59-84.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education* (Rev. ed.). San Francisco, CA: Jossey-Bass/John Wiley & Sons.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18 (2), 5-11.
- Messick, S. (2000). Consequences of test interpretation and use: The fusion of validity and values in psychological assessment. In R.D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment*, pp.3-20. Norwell, MA: Kluwer Academic Publishers.

- Monroe, W. S. (1923). *An introduction to the theory of educational measurements*. Cambridge, MA: Riverside Press/Houghton-Mifflin Company.
- Morgan, C., & Watson, A. (2002). The interpretive nature of teachers' assessment of students' mathematics: Issues for equity. *Journal for Research in Mathematics Education*, 33 (2), 78-110.
- Moss, P. (1998). The role of consequences in validity theory. *Educational Measurement: Issues and Practice*, 17 (2), 6-12.
- Moss, P. (2003). Reconceptualizing validity for classroom assessment. *Educational Measurement: Issues and Practice*, 22 (4), 13-25.
- Moss, P., Pullin, D., Gee, J. P., & Haertel, E. H. (2005). The idea of testing: Psychometric and sociocultural perspectives. *Measurement*, 3(2), 63-83.
- Nagy, P. (2000). Three roles of assessment: Gatekeeping, accountability, and instructional diagnosis. *Canadian Journal of Education*, 25 (4), 262-279.
- National Middle School Association. (2003). *This we believe: Successful schools for young adolescents*. Westerville, OH: National Middle School Association.
- Naylor, N. (2007). *Implementing primary class size caps in Ontario*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 9-13.
- Newkirk, T. (2004). A mania for rubrics. In A.S. Canestrari and B.A. Marlowe (Eds.), *Educational Foundations: An Anthology of Critical Readings* (pp.199-201). Thousand Oaks: Sage Publications.
- O'Donoghue, T. & Whitehead, C. (2008). A comparative analysis of teacher education across countries. In T. O'Donoghue, & C. Whitehead (Eds.), *Teacher education in the English-speaking world: Past, present, and future*, pp.191-199. Charlotte, NC: Information Age Publishing, Inc.
- Ontario College of Teachers. (1996). *Ontario College of Teachers Act, 1996 - Ontario Regulation 347/02*. Retrieved 08 December 2008 from <http://search.e-laws.gov.on.ca>
- Ontario College of Teachers. (2002, December). Defining special education. *Professionally Speaking*. Retrieved 08 December 2008 from <http://www.oct.ca/publications/>
- Ontario College of Teachers. (2003, September). *The state of the teaching profession in Ontario, 2003: A report to the Ontario College of Teachers based on a survey of the provinces' teachers*. Retrieved July 2006 from <http://www.oct.ca/publications>

- Ontario College of Teachers. (2004, September). *2004 annual state of education survey for the Ontario College of Teachers: A COMPAS special report on teacher and public perceptions*. Retrieved July 2006 from <http://www.oct.ca/publications>
- Ontario College of Teachers. (2005, September). *State-of-the-teaching-profession 2005 annual survey: A COMPAS report to the Ontario College of Teachers*. Retrieved July 2006 from <http://www.oct.ca/publications>
- Ontario College of Teachers. (2006, September). *State of the teaching profession 2006 annual survey: A COMPAS report to the Ontario College of Teachers*. Retrieved November 11, 2007 from <http://www.oct.ca/publications>
- Ontario College of Teachers. (2008a). *Foundations of professional practice*. Retrieved 1 July 2008 from [http://www.oct.ca/publications/PDF/foundation\\_e.pdf](http://www.oct.ca/publications/PDF/foundation_e.pdf)
- Ontario College of Teachers. (2008b). *Teacher Education*. Retrieved 08 December 2009 from [http://www.oct.ca/teacher\\_education/](http://www.oct.ca/teacher_education/)
- Ontario Ministry of Education. (1997). *The Ontario curriculum grades 1-8 Language*. Toronto, ON: author.
- Ontario Ministry of Education. (1999). *Ontario secondary schools, grades 9 to 12: Program and diploma requirements*. Toronto, ON: author. Retrieved 05 December 2008 from <http://www.edu.gov.on.ca>
- Ontario Ministry of Education. (2006). *The Ontario Curriculum Grades 1 - 8 Language*. Toronto, ON: author. Retrieved 22 November 2007 from <http://www.edu.gov.on.ca>
- Ontario Ministry of Education. (2007a). *The Ontario Curriculum Grades 9 and 10 English*. Toronto, ON: author. Retrieved 22 November 2007 from <http://www.edu.gov.on.ca>
- Ontario Ministry of Education. (2007b). *The Ontario Curriculum Grades 11 and 12 English*. Toronto, ON: author. Retrieved 22 November 2007 from <http://www.edu.gov.on.ca>
- Ontario Ministry of Education (2007c, May). *Report to the Partnership Table on teacher professional learning: Recommendations of the Working Table on teacher development*. Retrieved 03 May 2009 from <http://www.edu.gov.on.ca>
- Ontario Ministry of Education. (2007d, September). *Teacher moderation: Collaborative assessment of student work* (Literacy and Numeracy Secretariat Capacity Building Series). Retrieved April 28 2009 from <http://www.edu.gov.on.ca>
- Ontario Ministry of Education. (2008). *Growing success. Assessment, evaluation and reporting: Improving student learning*. Toronto, ON: author. Retrieved 21 November

2007 from <http://www.edu.gov.ca>

- Orton, R. E. (1997). Toward an Aristotlean model of teacher reasoning. *Journal of Curriculum Studies*, 29 (5), 569-583.
- Ozmon, H. A., & Craver, S. M. (2003). *Philosophical foundations of education* (7<sup>th</sup> ed.). Upper Saddle River, NJ: Merrill Prentice Hall.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage Publications.
- People for Education. (2007). *Class sizes in Ontario schools: The effects of provincial class size policy on Ontario's elementary and secondary schools*. Retrieved 30 April 2009 from <http://www.peopleforeducation.com>
- Peterson, S., Childs, R., & Kennedy, K. (2004). Written feedback and scoring of sixth-grade girls' and boys' narrative and persuasive writing. *Assessing Writing*, 9, 160-180.
- Peterson, S., & Kennedy, K. (2006). Sixth grade teachers' written comments on student writing: Gender and gender influences. *Written Communication*, 23 (1), 36-62.
- Phelan, A. M. (2005). A fall from (someone else's) certainty: Recovering practical wisdom in teacher education. *Canadian Journal of Education*, 28 (3), 339-358.
- Phillips, M. (2002). Is 'same' treatment 'fair' treatment? In J. Shulman, A.K. Whittaker & M. Lew (Eds). *Using assessments to teach for understanding*. New York: Teachers' College Press.
- Plake, B. S., & Impara, J. C. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, 12 (4), 10-12.
- Popham, W. J. (1993). Educational testing in America: What's right, what's wrong? A criterion- referenced perspective. *Educational Measurement: Issues and Practice*, 12 (1), 11-14.
- Popham, W. J. (2003). Seeking redemption for our psychometric sins. *Educational Measurement: Issues and Practice*, 22 (1), 45-48.
- Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, 34 (1), 1-20.
- Pryor, J. & Torrance, H. (1998). Formative assessment in the classroom: Where psychological theory meets social practice. *Social Psychology of Education*, 2, 151-176.

- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28 (1), 4-13.
- Raveaud, M. (2004). Assessment in French and English infant schools: Assessing the work, the child or the culture? *Assessment in Education*, 11 (2), 193-211.
- Rea-Dickens, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18 (4), 429-462.
- Rex, L. A. (2005). Testing as shaping: A response to "The idea of testing: Psychometric and sociocultural perspectives." *Measurement*, 3 (2), 103-107.
- Riessman, C. K. (2004). Narrative analysis. In M. S. Lewis-Beck, A. Bryman, & T. Futing Liao (Eds.), *The Sage encyclopedia of social science research methods*. Thousand Oaks, CA: Sage Publications.
- Rinsland, H. D. (1937). *Constructing tests and grading in elementary and high school subjects*. New York: Prentice-Hall, Inc.
- Rose, R. M., Kolb, F. A., & Barra-Zuman, N. (1991). Profile of an equitable classroom. As cited in Szpyrka (2001).
- Ross, M. E. (2005). Teacher education: Approaches for handling daily ethical dilemmas in the classroom. *Proquest Digital Dissertations*, UMI No.AAT 3205293.
- Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (1999). Effects of self-evaluation training on narrative writing. *Assessing Writing*, 6(1), 107-132.
- Ruddiman, J. (2004). Transactional theory in practice: Exploring eight-grade students' engagement with literature as reflected in reader response journals. *Proquest Digital Dissertations*, UMI No. AAT3135374.
- Ryan, T. G. (2001). An action research study of secondary science praxis. *Proquest Digital Dissertations*, UMI No. AAT NQ58601.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18 (2): 119-144.
- Salloum, S. L. (2006). Teaching as practice: Blending the intellectual and moral in pursuit of science teachers' practical knowledge. *Proquest Digital Dissertations*, UMI No. AAT 3223709
- Sandelowski, M. (2003). Tables or tableaux? The challenges of writing and reading mixed methods studies. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods*

*in social and behavioral research*, pp.321-350. Thousand Oaks, CA: Sage Publications.

Schendel, E., & O'Neill, P. (1999). Exploring the theories and consequences of self-assessment through ethical inquiry. *Assessing Writing*, 6 (2), 199-227.

Schmidt, M. J. (2001). Toward teacher inquiry when implementing classroom assessment. *Proquest Digital Dissertations*, UMI No. AAT NQ63809.

Scriven Carter, R. (1952). How invalid are marks assigned by teachers? *Journal of Educational Psychology*, 43 (4), 218-228.

Seidman, I. (1998). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (2<sup>nd</sup> ed.). New York: Teachers College Press.

Shepard, L. (1995). Using assessment to improve learning. *Educational Leadership*, 52, 38-43.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29 (7), 4-14.

Shepard, L. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, 63 (3), 66-71.

Shepard, L. (2006). Classroom assessment. In R.L. Brennan (Ed.), *Educational Measurement*, 4<sup>th</sup> ed., pp.623-646. Westport, CT: National Council on Measurement in Education, American Council on Education and Praeger Publishers.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-22.

Shulman, L. S. (1987/2004). The wisdom of practice: Managing complexity in medicine and teaching. In L. S. Shulman, *The Wisdom of practice: Essays on teaching, learning, and learning to teach* (S.M. Wilson, Ed.), pp.249-271. San Francisco, CA: Jossey-Bass.

Shulman, L. S. (1992/2004). Research on teaching: A historical and personal perspective. In L. S. Shulman, *The Wisdom of practice: Essays on teaching, learning, and learning to teach* (S.M. Wilson, Ed.), pp.364-381. San Francisco, CA: Jossey-Bass.

Shulman, L. S. (1996/2004). Just in case: Reflections on learning from experience. In L. S. Shulman, *The Wisdom of practice: Essays on teaching, learning, and learning to teach* (S.M. Wilson, Ed.), pp.249-271. San Francisco, CA: Jossey-Bass.

- Shulman, L. S. (1997/2004). Professional development: Learning from experience. In L. S. Shulman, *The Wisdom of practice: Essays on teaching, learning, and learning to teach* (S.M. Wilson, Ed.), pp.503-520. San Francisco, CA: Jossey-Bass.
- Smith, J. (2003). Reconsidering reliability in classroom assessment and grading. *Educational Measurement: Issues and Practice*, 22 (4), 26-33.
- Smith, E., & Gorard, S. (2006). Pupils' views on equity in schools. *Compare*, 36 (1), 41-56.
- Speck, B. W. (1998). Unveiling some of the mystery of professional judgment in classroom assessment. In R.S. Anderson & B.W. Speck (Eds). *Changing the way we grade student performance: Classroom assessment and the new learning paradigm*, pp.89-96. San Francisco, CA: Jossey-Bass, Inc.
- Spencer, B. L. (2006). The will to accountability: Reforming education through standardized literacy testing. *Proquest Digital Dissertations*, UMI No. AAT NR15947
- Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3<sup>rd</sup> ed.), pp.443-466. Thousand Oaks, CA: Sage Publications.
- Sternberg, R. J. (2000). Intelligence and wisdom. In R. J. Sternberg (Ed.), *Handbook of intelligence*, (pp.631-649). Cambridge: Cambridge University Press.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72, 534-539.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83, 758-765.
- Stiggins, R. J. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, 87, 324-328.
- Stiggins, R. J., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice*, 44 (1), 11-18.
- Strike, K., & Soltis, J. F. (1998). *The ethics of teaching*. New York: Teachers College Press.
- Stobart, G. (2005). Fairness in multicultural assessment systems. *Assessment in Education: Principles, Policy & Practice*, 12 (3), 275-287.
- Stokking, K., van der Schaaf, M., Jaspers, J., & Erkens, G. (2004). Teachers' assessment of students' research skills. *British Educational Research Journal*, 30 (1), 93-116.
- Sullivan, S. W. (2006). Feminism. In J. R. Shook and J. Margolis (Eds.), *A companion to*

*pragmatism*, pp.1-10. Oxford: Blackwell Publishing.

Suskie, L. (2002). Fair assessment practices: Giving students equitable opportunities to demonstrate learning, *Adventures in Assessment*, 14. Retrieved 21 April, 2004 from the System for Adult Basic Education Support/World Education Web site: <http://www.sabes.org/resources/adventures/vol14/14suskie.htm>

Szpyrka, D.A. (2001). Exploration of instruction, assessment, and equity in the middle school science classroom. *Digital Dissertations*, UMI No.AAT 3029061.

Taylor, B. M., Pearson, D. P., Peterson, D. S., & Rodriguez, M.C. (2005). The CIERA School Change Framework: An Evidence-based Approach to Professional Development and School Reading Improvement. *Reading Research Quarterly*, 40 (1), 40-69.

Threlfall, J. (2005). The formative use of assessment information in planning – the notion of contingent planning. *British Journal of Educational Studies*, 53 (1), 54–65.

Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education: Principles, Policy & Practice*, 13 (3), 239-264.

Tierney, R. D., & Charland, J. (2007, April). *Stocks and prospects: Research on formative assessment in secondary classrooms*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Tierney, R. D., Simon, M., & Charland, J. (2009, May). *Mathematics teachers' understanding of assessment principles in Ontario*. Paper presented at the annual meeting of the Canadian Society for Studies in Education, Ottawa, ON.

Tierney, R. J. (1998/2005). Reading assessment: Principles and practices for elementary teachers In S.J.Barrentine & S.M. Stokes (Eds.). *Reading assessment: Principles and practices for elementary teachers*, pp.29-40. Newark, DE: International Reading Association.

Tobin, G. A., & Begley, C. M. (2004). Methodological rigour within a qualitative framework. *Journal of Advanced Nursing*, 48 (4), 388-396.

Toronto District School Board. (2006). *Student Census, Grades 7-12: System Overview*. Toronto, ON: author.

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Buckingham, England: Open University Press.

Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using

- action research to explore and modify theory. *British Educational Research Journal*, 27 (5), 615-631.
- Treagust, D. F., Jacobowitz, R., Gallagher, J. L., & Parker, J. (2001). Using assessment as a guide in teaching for understanding: A case study of a middle school science class learning about sound. *Science Education*, 85(2), 137-157.
- Tunstall, P. (2001). Assessment discourse and constructions of social reality in infant classroom. *Journal of Education Policy*, 16 (3), 215-231.
- Uchiyama, M. K. (2004). Teachers' use of formative assessment in middle school-reform based mathematics classrooms. *Proquest Digital Dissertations*, UMI No. 3123299.
- Volante, L. (2006). Reducing bias in classroom assessment and evaluation. *Orbit Magazine*, 36 (2), 34-36.
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30 (3), 749-770.
- Waluchow, W. J. (2003). *The dimensions of ethics: An introduction to ethical theory*. Peterborough, ON: Broadview Press
- Warms, C. A., & Schroeder, C. A. (1999/2009). Bridging the gulf between science and action: The "new fuzzies" of neopragmatism. In P. G. Reed & N. B. Crawford Shearer (Eds.), *Perspectives on nursing theory*, pp.111-121. Philadelphia, PA: Lippincott Williams & Wilkins.
- Warne, C. (2006). *Aristotle's Nicomachean Ethics*. London: Continuum International Publishing Group.
- Watson, A. (1999). Paradigmatic conflicts in informal mathematics assessment as sources of social inequity. *Educational Review*, 51 (2), 105-115.
- Weeden, P., Winter, J., & Broadfoot, P. (2002). *Assessment: What's in it for schools?* New York: Routledge Falmer.
- Whittington, D. (1999). Making room for values and fairness: Teaching reliability and validity in the classroom context. *Educational Measurement: Issues and Practice*, 18 (1), 14-22.
- Wigfield, A., & Eccles, J.S. (2002). Students' motivation during the middle school years. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education*. San Diego, CA: Academic Press/Elsevier Science.

- Wiggins, G. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco, CA: Jossey-Bass, Inc.
- William, D. (1994). *Toward a philosophy for educational assessment*. Paper presented at the annual conference of the British Educational Research Association. Retrieved May, 2006 from [http://www.kcl.ac.uk/depsta/education/publications/BERA\\_94.pdf](http://www.kcl.ac.uk/depsta/education/publications/BERA_94.pdf)
- Wilks, T. (2004). The use of vignettes in qualitative research into social work values. *Qualitative Social Work*, 3(1), 78-87.
- Willie, C. V. (1985). The problem of standardized testing in a free and pluralistic society. *Phi Delta Kappan*, 66 (9), 626-627.
- Wilson, M. (2006). *Rethinking rubrics in writing assessment*. Portsmouth, NH: Heinemann, Reed Elsevier, Inc.
- Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research*, (2), 131-175.
- Wormeli, R. (2006). *Fair isn't always equal: Assessing and grading in the differentiated classroom*. Portland, MN: Stenhouse Publishers; Westerville, OH: National Middle School Association.
- Wright, R. L. (2002). Frictions in the machine: Teacher practical knowledge in a prison setting. *Proquest Digital Dissertations*, UMI No. AAT NQ77048.
- Yanchar, S. C., & Williams, D. D. (2006). Reconsidering the compatibility thesis and eclecticism: Five proposed guidelines for method use. *Educational Researcher*, 35 (9), 3-12.
- Yen, S.-H. (2006). Teaching Taiwanese indigenous students: Case studies of three Han Chinese teachers. *Proquest Digital Dissertations*, UMI No. 3241473.
- Yin, R. K. (2006). Case study methods. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research*, pp.3-31. Mahwah, NJ: Lawrence Erlbaum Associates; Washington, DC: American Educational Research Association.
- Yip, D. Y., & Cheung, D. (2005). Teachers' concerns on school-based assessment of practical work. *Journal of Biological Education*, 39 (4), 156-162.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45, 477-501.

- Young, T. R., & Arrigo, B. A. (1999). *The dictionary of critical social sciences*. Boulder, CO: Westview Press.
- Yung, B. H. W. (2001a). Examiners, policeman and students' companion: Teachers' perceptions of their role in an assessment reform. *Educational Review*, 53 (3), 251-260.
- Yung, B. H. W. (2001b). Three views of fairness in a school-based assessment scheme of practical work in biology. *International Journal of Science Education*, 23 (10), 985-1005.
- Zembylas, M. (2003) Emotion metaphors and emotional labor in science teaching. *Science Education*, 88 (3), 301-324.
- Zoeckler, L. G. (2005). Moral dimensions of grading in high school English. *Proquest Digital Dissertations*, UMI No. AAT3183500.

## APPENDICES

## **Appendix A: Learning English in Ontario's Educational System: An Overview**

In Ontario's English language schools, students learn English Language Arts from their entry in Junior Kindergarten (3 and 4-year olds) to their graduation at the end of Grade 12 (17 and 18-year olds). English is the only subject that is mandatory through all four years of secondary school, with one credit required for each grade (OME, 1999). In Grades 9 and 10, students choose academic or applied streams based on a recommendation from their previous English teacher. In Grades 11 and 12, English courses are stratified further with three options: university, college and workplace preparation. The curriculum for all of these English courses contains four strands. In Grades 7 and 8 the strands are Oral Communication, Reading, Writing, and Media Literacy (OME, 2006), and in Grades 9 to 12 they are Oral Communication, Reading and Literature Studies, Writing, and Media Studies. For students who are learning English as an additional language (in the English language system), there are five levels of courses for English as a Second Language and English Literacy Development (OME, 2007a, 2007b). In addition to the mandatory English curriculum, schools may offer different optional courses, depending on the interests and strengths of the teachers in their English departments. Optional courses include Business and Technological Communication, Canadian Literature, Contemporary Aboriginal Authors, Literacy Skills, Media Studies, Presentation and Speaking Skills, Studies in Literature, and Writer's Craft (OME, 2007b). In both mandatory and optional courses, writing is not only the subject of study, but also the vehicle for expressing learning in the other strands. For example, a teacher could assign a written essay to assess students' understanding of a dramatic text that had been studied as part of the reading strand.

## Appendix B: Ensuring Fairness in *The Student Evaluation Standards* (2003)

The 21 standards below are listed under “ensuring fairness” (p.x) in a thematic table of contents in *The Student Evaluation Standards* (JCSEE, 2003).

#	Standard
P1	<b>Service to Students</b> Evaluations of students should promote sound education principles, fulfillment of institutional missions, and effective student work, so that educational needs of students are served.
P2	<b>Appropriate Policies and Procedures</b> Written policies and procedures should be developed, implemented, and made available, so that evaluations are consistent, equitable, and fair.
P3	<b>Access to Evaluation Information</b> Access to a student’s evaluation information should be provided, but limited to the student and others with established legitimate permission to view the information, so that confidentiality is maintained and privacy protected.
P4	<b>Treatment of Students</b> Students should be treated with respect in all aspects of the evaluation process, so that their dignity and opportunities for educational development are enhanced.
P5	<b>Rights of Students</b> Evaluations of students should be consistent with applicable laws and basic principles of fairness and human rights, so that students’ rights and welfare are protected.
P6	<b>Balanced Evaluation</b> Evaluations of students should provide information that identifies both strengths and weaknesses, so that strengths can be built upon and problem areas addressed.
P7	<b>Conflict of Interest</b> Conflicts of interest should be avoided, but if present should be dealt with openly and honestly, so that they do not compromise evaluation processes and results.
U1	<b>Constructive Orientation</b> Student evaluations should be constructive, so that they result in educational decisions that are in the best interest of the student.
U3	<b>Information Scope</b> The information collected for student evaluations should be carefully focused and sufficiently comprehensive, so that evaluation questions can be fully answered and the needs of students addressed.
U5	<b>Explicit Values</b> In planning and conducting student evaluations, teachers and others who evaluate students should identify and justify the values used to judge student performance, so that the bases for the evaluations are clear and defensible.
U7	<b>Follow-Up</b> Student evaluations should include procedures for follow-up, so that students, parents/guardians, and other legitimate users can understand the information and take appropriate follow-up actions.
F2	<b>Practical Orientation</b> Student evaluation procedures should be practical, so that they produce the needed information in efficient, nondisruptive ways.
A1	<b>Validity Orientation</b> Student evaluations should be developed and implemented, so that interpretations made about the performance of a student are valid and not open to misinterpretation.

A3	<b>Context Analysis</b> Student and contextual variables that may influence performance should be identified and considered, so that a student's performance can be validly interpreted.
A4	<b>Documented Procedures</b> The procedures for evaluating students, both planned and actual, should be described, so that the procedures can be explained and justified.
A5	<b>Defensible Information</b> The adequacy of information gathered should be ensured, so that good decisions are possible and can be defended and justified.
A6	<b>Reliable Information</b> Evaluation procedures should be chosen or developed and implemented, so that they provide reliable information for decisions about the performance of a student.
A7	<b>Bias Identification and Management</b> Student evaluations should be free from bias, so that conclusions can be fair.
A8	<b>Handling Information and Quality Control</b> The information collected, processed, and reported about students should be systematically reviewed, corrected as appropriate, and kept secure, so that accurate judgments can be made.
A9	<b>Analysis of Information</b> Information collected for student evaluations should be systematically and accurately analyzed, so that the purposes of the evaluation are effectively achieved.
A10	<b>Justified Conclusions</b> The evaluative conclusions about student performance should be explicitly justified, so that students, their parents/guardians, and others can have confidence in them.

## Appendix C: Background Questionnaire

### Background Questionnaire

#### 1. Information and Instructions

##### Information About the Research Project

This research project is being conducted by Robin D. Tierney, a certified teacher and Ph. D. candidate, and supervised by Dr. Marjelle Simon at the Faculty of Education, University of Ottawa.

The purpose of this research is to understand experienced teachers' practical wisdom about formative in classroom assessment.

This is the first questionnaire for the research project. It asks you to provide information about your teaching experience and education, and to respond briefly to a set of items on the ethics of classroom assessment.

##### General Instructions

1. Please read, sign, and return the consent form before you begin this questionnaire.
2. To protect your confidentiality, please provide a pseudonym for the researcher to use.
3. As you complete the survey, follow the specific instructions given for different questions.
4. If you have any questions or concerns in completing the questionnaire, contact the researcher by email (rtier13@uottawa.ca) or by telephone 613-562-5900 x4510

#### 1. What is your pseudonym for this study?

\_\_\_\_\_

### Background Questionnaire

#### 2. Current Assignment

1. In which grades are you teaching English this year? Select all that apply for the current school year.

- Grade 7  
 Grade 8  
 Grade 9  
 Grade 10  
 Grade 11  
 Grade 12

Other (please specify): \_\_\_\_\_

2. What type of English classes are you teaching? Select all that apply for the current school year.

- Regular Classroom  
 English as a Second Language  
 English as a Second Language - Immersion  
 Applied English  
 Academic English  
 English for Specific Purposes (e.g., Writing a Term Paper, etc.)  
 Other (please specify): \_\_\_\_\_

3. What are your professional responsibilities? Select all that apply for the current school year.

- Classroom Teacher - English only  
 Classroom Teacher - Multiple Subjects  
 Practice Teacher (supervisor)  
 Guidance Counselor  
 Other (please specify): \_\_\_\_\_  
 Department Head  
 Administrator  
 Other (please specify): \_\_\_\_\_

### Background Questionnaire

#### 3. Length of Experience

1. Including this year, how many years of experience do you have TEACHING:

English as a subject	Years
The grade(s) you are currently teaching in the Ontario educational system	<input type="text"/>
In total, from the beginning of your first year of teaching	<input type="text"/>

2. Do you have any experience in education in a non-teaching role (e.g., curriculum leader, vice-principal)?

No (skip to next section)  
 Yes

3. Including this year, how many years of experience do you have in education in a non-teaching role?

### Background Questionnaire

#### 4. Previous Teaching Experience

1. What grade(s) did you teach in previous years?

Elementary (K to Grade 8)  Elementary (Grades 9 to 12)  
 Intermediate (Grades 7-8)  
Other (please specify): \_\_\_\_\_

2. Have you previously taught outside of Ontario?

No  
 Yes  
If yes, please specify where and when: \_\_\_\_\_

3. Have you previously taught any subjects other than English?

No  
 Yes  
If yes, please specify what(s): \_\_\_\_\_

**Background Questionnaire**

**5. Teaching Experience**

1. Is English the subject you prefer teaching the most?

- Yes  
 No

If yes, which methods do you prefer teaching most?

\_\_\_\_\_

2. Do you consider yourself a specialist in teaching English?

- No (not a great method)  
 Yes

3. On what basis do you consider yourself a specialist?

- Extensive background  
 Teaching experience  
 Professional development  
 A combination of education, and experience  
 Other (please specify): \_\_\_\_\_

**Background Questionnaire**

**6. Education**

1. Have you attended any educational programs as a student outside of Ontario?

- No  
 Yes

If yes, please specify where and what level.

\_\_\_\_\_

2. In which subject(s) did you major for your undergraduate degree(s)?

- English  
 Other

If Other, please specify the subject(s): \_\_\_\_\_

3. Did you complete a Bachelor of Education or teacher training program?

- No  
 Yes

If yes, please specify the length of the program in months: \_\_\_\_\_

4. Have you taken, or are you taking, any additional qualifications courses?

- No  
 Yes

If yes, please specify which: \_\_\_\_\_

\_\_\_\_\_

5. Have you taken, or are you taking, any graduate level university courses?

- No  
 Yes

If yes, please specify program: \_\_\_\_\_

\_\_\_\_\_

6. Have you attended any professional development sessions or workshops relating to teaching and/or assessing English Language Arts in the past five years?

- No  
 Yes

If yes, please specify: \_\_\_\_\_

\_\_\_\_\_

**Background Questionnaire**

7. Have you participated in any assessment activities outside of the classroom (e.g., EQAO)?

- No  
 Yes

If Yes, please describe:

8. Please briefly describe any other formal educational experiences that have not already been identified above:

**Background Questionnaire**

**7. Ethics of Classroom Assessment**

A recent survey by Green, Johnson, Kim & Pope (2007) showed high levels of disagreement among teachers about ethical issues in classroom assessment. The items in this section are based on those that the surveyed teachers agreed upon the LEAST.

1. For each of the items below, indicate whether you think the assessment practice is ethical or unethical, or whether it would depend on particular circumstances. Space is provided at the end for you to comment on any of these items.

	Ethical	Unethical	Depends on Circumstances
A teacher adds a few surprise items to a final exam about topics that were not included in the study guide she gave the students before the exam.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A teacher changes one student's final grade from C+ to A because his written work shows that he has mastered the learning expectations even though he did not complete all of the assigned work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A teacher weights homework heavily in his calculations for students' final grades.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A student receives a zero on a homework grade for not returning a form requiring a parent's signature.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To minimize guessing on a grammar test, a teacher announces that she will deduct more points for wrong answers than for leaving the answer blank.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A teacher creates learning activities with tasks that are similar to those included in the provincial assessment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
While administering a standardized test, a teacher notices that a student has skipped a problem and is now wandering off her answers out of boredom on the answer form. The teacher instructs the student to match the question number with the same number on the answer form, and shows the student where to record the answer she is working on.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A teacher always knows the identity of the students whose essay tests he is grading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A teacher includes only positive comments on students' strengths when writing report cards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Comments:

\_\_\_\_\_

**Background Questionnaire**

**8. Thank You!**

Thank you for completing this questionnaire.

## Appendix D: Blueprint for the Development of the Background Questionnaire

Section	Item	Item Description	Rationale
Instructions	1.1	Pseudonym	Confidential identification of participants.
Current Assignment	2.1	Grades	This information was used to describe individual cases, and provide evidence about the context in which the participant is currently operating, which will be especially helpful in understanding teachers' responses to the vignettes. Items in this section are similar to 2002 SAIP TQ items Q01A-H and Q02, but adapted specifically for the Ontario educational system.
	2.2	Type of Courses	
	2.3	Professional Responsibilities	
Length of Experience	3.1	Years of Experience; English as Subject; Current Grades; Ontario System; In Total	This information was used to confirm the participants' suitability for the study, and help establish the nature of the participant's teaching experiences. It was thought that differences in the nature of participant's experiences would be important in understanding their practical wisdom. For example, more teaching experience in the Ontario educational system might give a teacher a better understanding of particular circumstances, and more experience in a grade level could lead to better pedagogical items in this section are similar to 2002 SAIP TQ items Q28, Q29B, C, D.
	3.2	Experience in Non-Teaching Role	
	3.3	Years in Non-Teaching Role	
Previous Teaching Experience	4.1	Previous Grades (by panel)	This was used to describe the breadth and variety of educational contexts that the participant had encountered in each individual case, and to help understand participants' practical wisdom in analyzing vignette responses and interview transcripts. Items in this section are similar to 2002 SAIP TQ items Q01A-H and Q02C.
	4.2	Outside Ontario	
	4.3	Other Subjects	
Teaching Preference	5.1	English as Subject	This information was used to gain an idea of how participants viewed themselves, and indicated what they considered the basis of their specialization. Analysis of SAIP 2002 TQ data showed statistically significant associations between self-identification as specialist, years of experience and relevant degree.
	5.2	Self-Identification as Specialist	
	5.3	Basis of Specialization	

Section	Item	Item Description	Rationale
Education	5.1	Outside Ontario	This information was used to describe individual cases and determine the degree to which theoretical knowledge and technical could have been acquired. Items 5.2, 5.3, 5.5, 5.6 are similar to the 2002 SAIP TQ items Q30A-H and Q32. Other items aim to give an indication of the breadth or richness of the participant's educational experiences. An assumption here is that teachers who have been engaged in other educational systems, either as learners or teachers, may have been exposed to different classroom practices or beliefs about learning. Item 5.7 was asked to determine if participants had participated in assessment activities outside the classroom (e.g., marking for EQAO) because this would add to their understanding and experience in assessing student writing.
	5.2	Subject of Undergraduate Degree	
	5.3	Teacher Education	
	5.4	Additional Qualifications	
	5.5	Graduate Studies	
	5.6	Professional Development Sessions or Workshops	
	5.7	External Assessment Activities	
	5.8	Other Education	
Ethics of Classroom Assessment	7.1	Surprise	Green et al (2007): "For the final exam, a teacher always uses a few surprise items about topics that were not on the study guide" (Item 29, p.1006). Responses: 33.7% ethical / 66.3% unethical (degree of difference -32.6%).
		Change	Green et al (2007): "As a teacher finalizes grades, she changes one student's course grade from a B+ to an A because tests and papers showed that the students had mastered the course objectives even though he had not completed some of his homework assignments." (Item 13, p.1006). Responses: 37.3% ethical / 62.7% unethical (degree of difference -25.4%).
		Homework	Green et al (2007): "A teacher weighs homework heavily in determining report card grades." (Item 32, p.1006). Responses: 57.4% ethical / 42.6% unethical (degree of difference 14.8%).
		Zero	Green et al (2007): "A physical education teacher gives a student a zero as a homework grade for not returning a form requiring a parent's signature." (Item 8, p.1006). Responses: 42.6% ethical / 57.4% unethical (degree of difference -14.8%)

Section	Item	Item Description	Rationale
		Guessing	Green et al (2007): "To minimize guessing, a teacher announces she will deduct more points for a wrong answer than for leaving the answer blank." (Item 14, p. 1006). Responses: 30.8% ethical / 69.2% unethical (degree of difference -38.4%)
		Similar	Green et al (2007): "Based on his review of the district's mathematics frameworks, a teacher creates learning activities with specific math problems that are included in the annual achievement test" (Item 4, p. 1004). Responses: 53.3% ethical / 46.7% unethical (degree of difference 6.6%).
		Answer	Green et al (2007): "While administering a standardized test, a teacher notices that a child has skipped a problem and is now recording all his answers out of sequence on the answer form. The teacher stops at the child's desk and shows the student where to record the answer he is working on and instructs him to put the answer to each question with the same number on the answer sheet." (Item 17, p. 1005). Responses: 69.2% ethical / 30.8% unethical (degree of difference -38.4%).
		Identity	Green et al (2007): "A teacher always knows the identity of the students whose essay test she is grading." (Item 26, p. 1007). Responses: 48.5% ethical / 51.5% unethical (degree of difference -3.0%).
		Positive	Green et al (2007): "To enhance self-esteem, an elementary teacher addresses only students' strengths when writing narrative report cards" (Item 36, p. 1007). Responses: 40.8% ethical / 59.2% unethical (degree of difference -18.4%).

## Appendix E: Vignette Questionnaire

### Vignette Questionnaire

#### 1. Information

This research project is being conducted by Robin D. Tierney, a certified teacher and a Ph.D. candidate, and supervised by Dr. Marjelle Simon at the Faculty of Education, University of Ottawa. The project has been approved by the Social Sciences and Humanities Research Ethics Board (SSH REB) at the University of Ottawa.

The purpose of this research is to understand experienced teachers' practical wisdom about fairness in classroom assessment. Practical wisdom is defined as a responsive intellectual network that combines content knowledge, pedagogical expertise, moral beliefs, personal characteristics, experience, and understanding of circumstances in making decisions.

This is the second questionnaire for this research project. It asks you to respond to a set of eight vignettes that contain issues relating to fairness situated in the context of assessing student writing in English classes.

Both planned and interactive assessment events are included in these vignettes. Planned assessment events are those in which information about student learning is intentionally gathered at a predetermined point. Interactive assessment events are those which occur during learning activities when teachers recognize and respond to student learning.

#### 2. Instructions

The eight vignettes (A to H) in this questionnaire are all followed by the same three questions:

a. What decision would you recommend for the assessment to be fair?

For this question, include what you consider ideal for the assessment to be fair in supporting student learning. You may decide on one of the options that the teacher in the vignette is considering, or you may recommend any other course of action that you think would be fair.

b. How did you come to this decision?

For this question, please specify the knowledge, expertise, beliefs, experience, or understanding (elements of practical wisdom) that you used in answering the first question.

For example, in writing feedback on a student's essay, a teacher might draw on knowledge about writing, past experiences in teaching and assessing writing, beliefs about the importance of different aspects of writing, or an understanding of the particular student.

c. Which factors could make it easier or more difficult for a teacher to follow your recommendation?

For this question, include any factors in practice that you think would influence a teacher in following the decision you recommend.

Examples of the many factors that could help or hinder the use of fair assessment for learning are: administrative support, assessment policies, availability of teaching/learning materials, class size, colleague support, curriculum, grade level, parental expectations, preparation time, professional development, range of student abilities, report card requirements, subject matter, school climate, student behaviour, teacher characteristics, teaching traditions, team teaching, time constraints, and workload.

Please provide as much thought and detail as you can for each vignette. You are welcome to complete this questionnaire in more than one sitting.

If you have any questions or concerns as you complete the questionnaire, contact the researcher by telephone at 613-562-5800 x4510 or by email: [rdtier11@uottawa.ca](mailto:rdtier11@uottawa.ca)

#### 1. What is your pseudonym for this study?

### Vignette Questionnaire

#### 3. Vignette A

As the students began working on their article summaries, a teacher noticed a student who seemed unsure of what to do. The teacher realized that this student was away with the school band when the lesson on summarizing was taught. As the teacher set to work individually with the student, she noticed two other students who also seemed to be having trouble with the task. She wondered if she should work with all three of these students in a small group. But given that the other two were often inattentive in class, she wasn't sure if they should be given the same opportunity as the student who was away with the band.

1. What decision would you recommend for the assessment to be fair?

a)

b)

2. How did you come to this decision? Please specify the source(s) that you used in answering the first question (e.g., knowledge, expertise, beliefs, etc.).

a)

b)

3. Which factors would make it easier or more difficult for a teacher to follow your recommendation? You may refer to the list of possible factors in the instructions.

a)

b)

#### 4. Vignette B

A teacher picked up a student's essay with a sigh. He was thinking about the long, flowery sentences he'd been reading, and he was wishing that he'd drilled his students more on how to write short, clear sentences. Just then a colleague entered the stair room, and in honoring the name on the essay, she remained on the student's talent as a writer. She went on to say how much she'd enjoyed the student's writing during the previous year. Realizing that he and his colleague had different views, the teacher questioned his own assessment of the student's writing. He wondered whether he should consider his colleague's opinion, or talk to the student about his views, or perhaps discuss writing style with the whole class.

1. What decision would you recommend for the assessment to be fair?

a)

b)

2. How did you come to this decision? Please specify the source(s) that you used in answering the first question (e.g., knowledge, expertise, beliefs, etc.).

a)

b)

3. Which factors would make it easier or more difficult for a teacher to follow your recommendation? You may refer to the list of possible factors in the instructions.

a)

b)

#### 5. Vignette C

A group of English and History teachers collaborated in designing a unit on historical fiction. The teachers planned to have students write historical fiction, and they developed a rubric to use with the students from the beginning of the unit. After the meeting, though, one of the teachers was not sure about this approach. She wondered if giving

### Vignette Questionnaire

students the criteria beforehand would actually support their learning about writing, or if it would give them too much specific information, and maybe even discourage some students from developing as independent and creative writers.

1. What decision would you recommend for the assessment to be fair?  
 A  B
2. How did you come to this decision? Please specify the source(s) that you used in answering the first question (e.g., knowledge, expertise, beliefs, etc.).  
 A  B
3. Which factors would make it easier or more difficult for a teacher to follow your recommendation? You may refer to the list of possible factors in the instructions.  
 A  B

### 6. Vignette D

After the students finished the first draft of their persuasive essays, their teacher asked them to trade with a nearby classmate to edit each others' work. Two students who traded essays could not have been more different. The first had a neatly written, technically competent essay. The second student's work contained spelling mistakes, and it looked quite messy. However, the argument was logical and well developed. As the first student edited, she commented loudly on the second student's mistakes, and several classmates laughed. The teacher was concerned about this type of interaction, but he wasn't sure what to do about it. He wondered whether he should speak to this student, or address the whole class, or just avoid peer-editing in the future.

1. What decision would you recommend for the assessment to be fair?  
 A  B
2. How did you come to this decision? Please specify the source(s) that you used in answering the first question (e.g., knowledge, expertise, beliefs, etc.).  
 A  B
3. Which factors would make it easier or more difficult for a teacher to follow your recommendation? You may refer to the list of possible factors in the instructions.  
 A  B

### 7. Vignette E

A teacher read her students' plans for a research report, and she wrote feedback to help them organize their work. As she read the last research plan, the teacher realized that she had written increasingly more as she progressed through the class pile. The work that happened to be at the top of the pile received much less feedback than that at the bottom. The teacher considered taking the time to go back over all the students' research plans to make her feedback more consistent, but the class pile was large, and she still had to organize the next day's lessons. She wondered if she should add some feedback to the first few plans, or maybe she could just give some students oral feedback in class later.

### Vignette Questionnaire

1. What decision would you recommend for the assessment to be fair?  
 A  B
2. How did you come to this decision? Please specify the source(s) that you used in answering the first question (e.g., knowledge, expertise, beliefs, etc.).  
 A  B
3. Which factors would make it easier or more difficult for a teacher to follow your recommendation? You may refer to the list of possible factors in the instructions.  
 A  B

### 8. Vignette F

At the end of the first week of school, a teacher wrote comments in his students' writing folders. To get an idea of their writing level, he had asked the students to respond to a few quick prompts. Overall, he was quite pleased. However, the quality of the writing in one folder did surprise him. The writer had been formally identified as a gifted student, but the writing was just average. He had assumed that this student's writing would be more advanced, and he wondered how to respond. He wasn't sure if his feedback should reflect the usual expectations for this grade, or focus just on what the student had written, or perhaps be more demanding given the student's identification.

1. What decision would you recommend for the assessment to be fair?  
 A  B
2. How did you come to this decision? Please specify the source(s) that you used in answering the first question (e.g., knowledge, expertise, beliefs, etc.).  
 A  B
3. Which factors would make it easier or more difficult for a teacher to follow your recommendation? You may refer to the list of possible factors in the instructions.  
 A  B

### 9. Vignette G

As the students wrote in their journals, the teacher circulated and offered encouragement. But she began to notice many capitalization and punctuation errors, problems with subject-verb agreement, and incomplete sentences in the students' entries. She expected students at this level to have already mastered these skills, and she believed that it was important to develop good writing habits. She considered saying this immediately, but she was concerned about interrupting with negative feedback while the task was in progress. She wondered if it would be better to discuss her expectations at a later point, or maybe just ignore these types of errors in the students' journals.

1. What decision would you recommend for the assessment to be fair?  
 A  B

### Vignette Questionnaire

2. How did you come to this decision? Please specify the source(s) that you used in answering the first question (e.g., knowledge, expertise, beliefs, etc.).

4 |

9 |

3. Which factors would make it easier or more difficult for a teacher to follow your recommendation? You may refer to the list of possible factors in the instructions.

4 |

9 |

### 10. Vignette H

The student's portfolio was missing several assignments, but the writing that was included was clearly expressed and well organized with very few mechanical errors. The teacher realized the excellent quality of the student's work, but he wasn't sure how to react to an incomplete portfolio. Given that they were only midway through the course, he wanted to encourage the student to develop further. But he wondered if it would be best for the student to finish all the work that had already been assigned, or if he should plan some additional writing tasks that would be more challenging for the student. Or maybe he should just talk to the student about the missing assignments.

1. What decision would you recommend for the assessment to be fair?

4 |

9 |

2. How did you come to this decision? Please specify the source(s) that you used in answering the first question (e.g., knowledge, expertise, beliefs, etc.).

4 |

9 |

3. Which factors would make it easier or more difficult for a teacher to follow your recommendation? You may refer to the list of possible factors in the instructions.

4 |

9 |

### 11. Thank You

Thank you for your responses. Your participation is appreciated.

## Appendix F: Blueprint for the Development of the Vignette Questionnaire

Vignette	Fairness	Assessment for Learning		Writing Activity		
Key Aspect (McMillan, 2007)	Principles (JCSEE, 2003)	Core Practice (Tierney & Charland, 2007)	Nature (Cowie & Bell, 1999)	Common (SAIP, 2002)		
A	Opportunity to learn	U1, A1, A3, A7	Adjustment - Using assessment information to adjust teaching	Interactive - Recognizing and considering how to respond	Individual instruction (Q06G, Q14S); Writing summaries (Q14K, Q17H)	Summary as form of writing (Glossaries); Writing summaries (SE in G8 to G12)
B	Avoiding bias (teacher)	U5, A1, A5, A7, A9	Feedback - Giving oral feedback to students; Adjustment - Using assessment information to adjust teaching	Planned - Interpreting and considering how to act	Regular feedback (Q14N, Q14Q, Q24D); Difficulty judging writing (Q13F)	Varying sentence types (SE in G7 to G12)
C	Student knowledge of targets & criteria	P1, U1, U5, A1, A2, A4, A7, A9, A10	Criteria - Clearly communicating learning expectations and assessment criteria;  Students - Actively involving students in learning and assessment	Planned - Planning how to eliciting and interpret	Team planning time (Q06A, Q06D, Q11); Writing narrative (Q14A, Q14H, Q17D); Assessment criteria (Q12H, Q14U)	Rubrics (Achievement chart, G7 to G12; SE in G11 & G12); Communication about assessment (Assessment & Evaluation, G7 to G12); Writing fiction (SE in G7, G10-G12); Historical fiction (SE in G7, G10, G12)

<b>Vignette</b>	<b>Fairness</b>	<b>Assessment for Learning</b>	<b>Writing Activity</b>
	Key Aspect (McMillan, 2007)	Principles (JCSEE, 2003)	Core Practice (Tierney & Charland, 2007)
		Nature (Cowie & Bell, 1999)	Common (SAIP, 2002)
			Mandated (OME, 2006, 2007a,b)
<b>D</b>	Avoiding bias (students); and Student knowledge of targets & criteria	P1, P3, P4, P5, U1, A5, A7	Students - Actively involving students in learning and assessment with focus on importance of constructive learning environment
		Interactive - Noticing, beginning to recognize and considering how to respond	Persuasive essays (Q14A, Q20C); Peer assessment (Q14M, Q24F)
			Peer editing (SE in G7 to G12); Need for supportive environment (Writing Strand, G9 to G12)
<b>E</b>	Equal treatment of students*	U1, A1, A5, A6, A9	Feedback - Giving written feedback to students
		Planned - Considering how to act	Hours marking (Q06B); Feedback to students (Q14Q); Longer writing projects (Q14A, Q14C); Report writing (Q17I)
			Research report as form of writing (Glossaries); Organizing ideas (OE in G7 to G12); Planning research (SE in G7 to G12)
<b>F</b>	Avoiding teacher stereotypes; and Accommodating special needs	A1, A2, A6, A7	Feedback - Giving written feedback to students
		Planned - Interpreting and considering how to act	Hours marking (Q06B); Feedback to students (Q14Q); Samples of students' writing in weighting assessment (Q25E)
			Samples of student work as evidence; frame of reference and use of professional judgements; modifying expectations (Assessment & Evaluation, G7 to G12)

<b>Vignette</b>	<b>Fairness</b>	<b>Assessment for Learning</b>	<b>Writing Activity</b>
	<p><b>Key Aspect</b> (McMillan, 2007)</p> <p><b>Principles</b> (JCSEE, 2003)</p> <p><b>Core Practice</b> (Tierney &amp; Charland, 2007)</p> <p><b>Nature</b> (Cowie &amp; Bell, 1999)</p> <p><b>Common</b> (SAIP, 2002)</p> <p><b>Mandated</b> (OME, 2006, 2007a,b)</p>		
<b>G</b>	<p>Student knowledge of targets &amp; criteria; and Prerequisite knowledge and skills</p> <p>P1, P4, P6, U5, A2, A3, A5, A7, A9</p>	<p>Criteria - Clearly communicating learning expectations and assessment criteria; Feedback - Giving oral feedback to students</p> <p>Interactive - Noticing and recognizing in relation to teacher expectations</p>	<p>Use of language conventions (Achievement Chart, G7 to G12); Journal writing (SE in G7 to G12)</p> <p>Knowledge of grammar (Q13C, Q13H, Q14F); Use of language conventions (Q20M, Q20N, Q20O); Journal writing (Q17E)</p>
<b>H</b>	<p>Accommodating special needs; Prerequisite knowledge and skills</p> <p>U5, A2, A6, A7, A9</p>	<p>Approaches - Using varied methods to elicit information about learning</p> <p>Planned - Interpreting and considering how to act</p>	<p>Use of portfolios (SE in G7 to G12)</p> <p>Hours marking (Q06B); Completing work (Q24A); Samples of students' writing and portfolios in weighting assessment (Q25E, Q25F)</p>

## Appendix G: Template for Individualized Interview Protocol

*This template was used to develop an individualized protocol for each participants' follow-up interview. All of the participants were asked questions in each of the sections below, but the actual questions varied in response to their questionnaire answers. Questions F1 and G1 were asked of all participants.*

### Questions for \_\_\_\_\_

Interview Time and Date: \_\_\_\_\_

Introductions. Glad you could make time for this. These are copies of your responses to the questionnaires for us to go over together as we talk.

#### **A. Teacher's Profile (Background Questionnaire)**

Let's start with the background questionnaire. I'll be using it to write up a profile of each teacher in the study. These are your responses, and these are the bits I've added in from your certificate.

- A1. It looks to me like you've been [ask for any clarifications needed]
- A2. So you have been, or can you tell me more about [ask for any elaboration that might be helpful]
- A3. I just want to check here, or is that accurate [ask for confirmation]
- A4. Is there anything you want to add about yourself as a teacher for me to include in your profile?

#### **B. General Ethics (Background Questionnaire)**

Okay, let's move on to the general ethics. These were the items drawn from the survey of American teachers that showed a lack of consensus about the ethics of classroom assessment.

- B1. For this item you've indicated (ethical or unethical) and I'm wondering if there are any circumstances where you'd think it was/wasn't?
- B2. For this one you said that you do x. Do you think that makes an assessment fairer?
- B3. I'd like to clarify what you mean by x in your response here?
- B4. You've indicated that this item would depend on the circumstances. Can you give me an example of the circumstances that would change the ethics?

### **C. Recommendations for Fair Assessment (Vignette Questionnaire)**

For the vignettes, what I've done is grouped your responses so that we have all of them for the first question, all for the second, and all for the third, and I'd like to talk about them that way.

When I reviewed the research on fair assessment, I found that people interpret fairness in a lot of different ways. Here's a list of some of the aspects or interpretations that I found.

C1. One aspect that stands out to me in your recommendations is x. For example, you say in Vignette X that . . . . So would you say that an assessment is fairer if it is x?

C2. A second point you made seems to relate to x. That could be interpreted as several of these aspects of fairness, so I'm wondering if you can tell me a bit more.

C3. Something that I think I'm hearing in your responses is x. Would you agree with that interpretation? Can you tell me more?

### **D. Sources of Practical Wisdom (Vignette Questionnaire)**

In this section you identified the sources for your recommendations, and you seem to draw on a combination of things, so I'd like to try to tease that apart a little further.

This is the conceptual model I'm using to describe practical wisdom. It's based on research and theories about our practical knowledge.

D1. What I'd like to understand better is when you say x, do you mean x?

D2. In your response for Vignette X, you wrote x. Do think you've always felt that way as a teacher, or is that something that has developed over time?

D3. The other thing I'd like to discuss, especially for this Vignette, is what's behind that. Is that a [moral belief, policy - query]?

### **E. External Factors (Vignette Questionnaire)**

In this section you identified quite a few external factors. Here's the list of possibilities that was on the questionnaire.

E1. Looking at over your responses, you identified x, x, and x most often. Would you say they are the factors that would be most likely to make it harder or easier for a teacher to assess fairly?

E2. Looking at the list again, are there any that you'd like to add or discuss further?

E3. So for this one (identified on questionnaire), how would you say that impacts a teachers' ability to assess fairly?

### **F. Teachers' Experience (Open)**

With the vignettes in the questionnaire I tried to illustrate fairness issues that come up in the classroom, but of course they aren't the only one. I could have included others.

F1. So I'm wondering if there's any fairness issues that you've encountered in your practice, either like the ones in the vignettes, or different. [If needed: Can you give me an example, and tell me about what you decided to do?]

### **G. Sharing Knowledge**

G1. I have one very last question for you. If you could tell new teachers or teacher educators, or researchers like me who are trying to understand fairness in classroom assessment what is *most important* for fair assessment, what would you say?

**Thank you very much for your time!**

## **Appendix H: Email Recruitment Text**

Dear [Name],

The information in your teaching certificate indicates that your knowledge and experience could be of great value for novice teachers, educational researchers, and teacher educators.

I am a certified teacher and a doctoral student, and I am currently looking for participants for a research project that draws on the practical wisdom of experienced teachers to understand fairness in assessing student writing in English Language classes from grades 7 to 12. This research project is being supervised by Dr. Marielle Simon, Faculty of Education, University of Ottawa.

Participating in this research would involve you in completing a background questionnaire, responding to a set of vignettes about the classroom assessment, and engaging in a short interview. This will take no more than three hours of your time in total.

If you are interested in sharing your knowledge and experience, or you would like to know more about this research project, please contact me, either through the Teachers' Network, or by emailing

Robin D. Tierney  
PhD Candidate  
Faculty of Education  
University of Ottawa

## Appendix I: Information Sheet

University of Ottawa and Faculty of Education Logos

### Invitation to Participate

You are invited to participate in a research project on Teachers' Practical Wisdom about Fairness in Classroom Assessment. This research is being conducted by Robin D. Tierney, a certified teacher and a Ph.D. candidate, and supervised by Dr. Marielle Simon, the Director of Graduate Studies (Francophone sector) at the Faculty of Education, University of Ottawa. The project has been approved by the Social Sciences and Humanities Research Ethics Board (SSH REB).

### Purpose of the Research

Teachers' assessment decisions can have a significant and long-lasting impact on students, yet novice teachers often receive very little guidance in this area. The purpose of this research project is to draw on the practical wisdom of experienced teachers to understand fairness in assessing student writing in English.

Practical wisdom is defined as a responsive intellectual network that combines content knowledge, pedagogical expertise, moral beliefs, personal characteristics, experience, and understanding particular circumstances in making decisions.

### Participation

Participants in this research project will be experienced teachers who are willing to share their practical wisdom about assessing writing in English Language classes, from Grades 7 to 12 in Ontario. Participants should have a Bachelor of Arts and a minimum of ten years of teaching experience.

Participation in this study will involve completing a background questionnaire (10 to 15 minutes), responding to a set of vignettes in a questionnaire (50 to 65 minutes), and engaging in a follow-up interview with the researcher (30 minutes). The interview will be scheduled at a convenient time and location for participants. Following the interview, participants will be asked to review the interview transcript and provide the researcher with feedback (10 to 15 minutes).

### Confidentiality

The information you share as a participant will remain strictly confidential and will be used solely for the purpose of this research. Access to the research data will be limited to the researcher and her supervisor. You will be asked to provide a pseudonym, and this will be used by the researcher to ensure that you will not be identifiable in any of the research documents, or publications and presentations that result from this research.

## **Risks**

Your participation in this study will involve reflecting on fairness in classroom assessment. If this causes you any discomfort, you may communicate this to the researcher, and every effort will be made to minimize this risk. As classroom assessment is an area of practice in which more than one option is often defensible, the dialogue between yourself and the research is intended to be of a respectful and collegial nature.

## **Conservation of Data**

All data collected during this research project, including completed questionnaires, interview transcripts, and the researcher's notes, will be kept in a locked cabinet in the supervisor's office at the University of Ottawa for a period of five years, at which time it will be destroyed.

## **Benefits**

By sharing your knowledge and experience, you will be providing important and useful information for novice teachers, researchers, and teacher educators.

Although there is no financial compensation for participating in this research project, you will receive a book store gift certificate from the researcher as a token of appreciation for your time. You will also receive a complete electronic copy of the final report, and if you wish, copies of any subsequent publications or presentations that result from the research.

## **Voluntary Participation**

You are under no obligation to participate. If you would like to participate, please read and sign the consent form that accompanies this information sheet.

### **Researcher**

Robin D. Tierney  
Faculty of Education  
University of Ottawa

### **Supervisor**

Marielle Simon, Ph.D  
Faculty of Education  
University of Ottawa

## Appendix J: Consent Form

### Consent Form



Title of the study: **Teachers' Practical Wisdom about Fairness in Classroom Assessment**

Researcher: Robin Tierney, Ph.D. Candidate  
Faculty of Education, University of Ottawa

Supervisor: Marielle Simon, Ph.D.  
Faculty of Education, University of Ottawa

I am invited to participate in the abovementioned research study conducted by Robin Tierney, a doctoral candidate supervised by Dr. Marielle Simon, Faculty of Education at the University of Ottawa.

This study has been approved by the Social Sciences and Humanities Research Ethics Board (SSH REB) at the University of Ottawa.

The purpose of this study is to understand the practical wisdom of experienced teachers about fairness in classroom assessment. Practical wisdom is defined as a responsive intellectual network that combines content knowledge, pedagogical expertise, moral beliefs, personal characteristics, experience, and understanding particular circumstances in making decisions.

The purpose of this study is to understand the practical wisdom of experienced teachers about fairness in classroom assessment. Practical wisdom is defined as a responsive intellectual network that combines content knowledge, pedagogical expertise, moral beliefs, personal characteristics, experience, and understanding particular circumstances in making decisions.

As a participant in this study, I will complete a background questionnaire (10 to 15 minutes), respond to a set of vignettes in a written questionnaire (50 to 65 minutes), and engage in a follow-up interview with the researcher (30 minutes). The interview will be scheduled at my convenience and will take place at a location of my choice. I will also be asked to review the interview process and transcript (10 to 15 minutes), and I will have up to two weeks following receipt of the documents to contact the researcher for any clarifications or modifications. My participation in this study will involve reflecting on fairness in classroom assessment. If this causes me any discomfort, I may communicate this with the researcher, and/or refuse to answer any questions. I have received assurance from the researcher that every effort will be made to minimize this risk. As classroom assessment is

1967 | 1968 - 2007 | 2008

UNIVERSITY OF OTTAWA

an area of practice in which more than one option is often defensible, the dialogue between myself and the researcher is intended to be of a respectful and collegial nature. If I become interested in knowing more about teachers' practical wisdom or fair classroom assessment through this study, the researcher will provide me with a list of resources, or an electronic copy of selected articles.

By participating in this study I will have the opportunity to share my knowledge and experience, and contribute to the advancement of research and practice of fair classroom assessment. If I indicate an interest in reading the research report, or any publications or presentations that result from the study, I will be given a complete electronic copy.

I understand that the results will be used for the purposes of this research only. The information I share as part of this study will remain strictly confidential. I will provide the researcher with a pseudonym to protect my confidentiality, and my name not be revealed in any step of the research process or resulting publication. All electronic communication will be printed and then deleted immediately.

Printed communication and all data collected, including written questionnaire responses, interview transcripts, and the researcher's notes, will be kept in a secure cabinet in the thesis supervisor's office at the University of Ottawa, and will be destroyed five years after the study is completed.

I am under no obligation to participate. If I choose to participate, I may withdraw from the study at any time. I may refuse to answer any questions without suffering any negative consequences. If I choose to withdraw, any data incomplete at the time of withdrawal will not be considered in the research.

-----  
**Acceptance:** I, \_\_\_\_\_, agree to participate in this research project conducted by Robin Tierney, Ph.D. Candidate at the Faculty of Education, University of Ottawa.

I, \_\_\_\_\_, consent to participate in an audio-recorded interview.

If I have any questions about the study, I may contact the researcher or her thesis supervisor at the telephone numbers or e-mail addresses given above.

If I have any ethical concerns regarding my participation in this study, I may contact the Protocol Officer for Ethics at the University of Ottawa, 550 Cumberland Street, Room 159, 613-562-5841 or [ethics@uottawa.ca](mailto:ethics@uottawa.ca).

There are two copies of the consent form, one of which is mine to keep.

Participant's signature:

Date:

Researcher's signature:

Date:

**PLEASE RETURN A COPY OF THIS CONSENT FORM USING THE FOLLOWING FAX NUMBER FOR THE FACULTY OF EDUCATION, UNIVERSITY OF OTTAWA: 613-562-5987**