

Converting Student Achievements to Letter Grades:

Formative Assessment and the BC Performance Standards in Practice

Martin Bozic

July 2013

Simon Fraser University

Table of Contents

Abstract

Introduction: Statement of the Problem

Topic and Research Problem	1
Justification for the Research Problem	2
Deficiencies in the Evidence	3
The Audience	4
Research Question	5

Literature Review

Removing Subjectivity from Teacher Judgment	6
Rubrics in Standardized Testing	7
The Practice of Rubric Conversion for Mandated Reporting	8
Converting Rubric Outcomes for the Purpose of Grading	10

Methodology

Research Design	14
Participants	15
Data Collection and Instruments	16
<i>Phase One: Small Group Interviews</i>	16
<i>Phase One: Individual Interviews</i>	17
<i>Phase Two: Survey</i>	18
Limitations	19
Ethics	20

Data Analysis

Theme One: Teachers modify BCPS rubrics to better measure student performance	21
Theme Two: Different paths of assessment practice can be drawn from the depth and breadth in teachers' understanding and use of FA and the BCPS	23
Theme Three: Teachers correlate letter grades to BCPS achievement levels, but there is no regularity or consistency	25
Theme Four: Using the BCPS language to communicate academic performance to parents poses challenges	27
Theme Five: Assessment <i>for</i> learning occurs more often than assessment <i>as</i> learning	28
Theme Six: Moving away from numbers in assessment is desirable but personally challenging	30
Innovation in FA and Rubric strategies	31
Conclusion	33
References	35
Appendices	37
Appendix 1: Interview Protocol	38
Appendix 2: Survey Instrument	40
Appendix 3: Survey Data (Tables 2.0 thru 2.9)	43

ABSTRACT

Researchers touting the benefits of formative assessment and criterion-referenced assessment rubrics agree that there is no simple, single method for converting student achievements to letter grades. Rubric conversion methods vary considerably in purpose and design and present a challenge for teachers when reporting student progress via formal report cards. In British Columbia, the Ministry of Education's policy on reporting students' progress and the BC Performance Standards' (BCPS) suggested CRA rubrics and language provide direction but also create additional complexity for teachers' practices when needing to convert student achievement to letter grades. The central question in this research is "How do middle school teachers in School District 43 use the BCPS suggested CRA rubric language and matrix while needing to meet Ministry guidelines for student progress reporting?" The study also aims to discover how middle school teachers use rubrics toward formative and summative assessment *for, as* and *of* student learning while promoting a focus on learning rather than letter grades.

Keywords: rubric, BC Performance Standards, letter grades, formative and summative assessment.

INTRODUCTION: STATEMENT OF THE PROBLEM

Topic and Research Problem

In British Columbia, the Ministry of Education acknowledges and supports the importance of formative assessment (FA) and criterion-referenced assessment (CRA) by providing detailed strategies and rubrics in the BC Performance Standards guide (BCPS) for Grades 1 through 10 (BCPS, 2011). However, problems of ambiguity emerge when the language based descriptors used to assess the four levels of student performance outlined in the BCPS (Not Yet Within, Minimally Meets, Fully Meets, Exceeds Expectations) are transposed to fit the parameters of mandated letter grades for formal reporting. Even though research has demonstrated FA and CRA to be effective practice for improving student learning and promoting life-long learning (Clark, 2011), the standardization and certainty – as well as the currency – associated with numbers and quantified letter grades often prevails over the importance and philosophy of FA and CRA pedagogy. In my teaching experience, I and like-minded colleagues struggle each year to instill in students the value of deep learning above letter grades. Yet it seems that the extrinsic rewards and validity that come with quantitative assessment (number totals and percentages) outweigh – and in some cases, over-power – the intrinsic rewards of developmental learning and self-improvement. The purpose of this research is to share insight, innovations and common strategies teachers – including myself – have developed in using BCPS CRA rubrics and achievement language while fulfilling the provincially mandated policy and practice for reporting on student progress. In addition, the research aims to explore the philosophies and pedagogy of teachers who use FA and CRA rubrics as assessment *for, as and of* learning to improve and promote the importance of student learning above letter grades.

Justification for the Research Problem

Current research identifies mandated reporting practices as a contributor to the challenges faced in teaching and evaluating student learning. The *Washington State Diagnostic Assessment Guide* notes “weaknesses in current assessment practices that directly undermine learning and instructional effectiveness” (Stevens, 2009, p. 39). Black & Wiliam (1998) identify three popular but misguided assessment practices:

a) tests that emphasize superficial learning and recall, b) teachers who appear to be unaware of the assessment work of colleagues and do not trust or use other teachers’ assessment results, and c) an emphasis on quantity and presentation of work rather than on quality of work in relation to learning. (as cited in Stevens, 2009, p. 39-40)

Moreover, Cizek, Fitzgerald and Rachor (1995) note that “current assessment practices overemphasize grading functions and underemphasize feedback and advice for learning, focus on competition rather than personal improvement, and use comparative assessment interpretations” (as cited in Stevens, 2009, p. 41). Furthermore, when educators and parents think ‘*good grades equals success and opens opportunities*’ learning with depth and breadth is less likely to occur because attention is focussed on the grade rather than the learning.

In *Going Gradeless: Evaluation over Time Helps Students Learn to Write* (2006), Jan Loveless reminisces of her struggles with FA and promoting student learning and development above letter grades. Though Loveless (2006) had “successfully de-emphasized grades as an evaluation device” (p. 12) for her previous grade 7 class, the same success could not be as easily reproduced at higher levels. She became frustrated with her senior Advanced Composition students whom she identified as “egocentric, formulaic writers” (p, 12) valuing only the traditional letter grade attached to each assignment as opposed to reflecting and learning from the

comments she so painstakingly provided. As both Loveless and others have experienced, students accustomed to traditional grading usually think “so ‘furtherly’ in the future” that “they are not choosing and growing in the present” (Loveless, 2012, p. 12). Like Loveless, more and more teachers are experimenting with FA, CRA and removing student focus from traditional letter grades. Still the problem remains of how a teacher can assign and defend a letter grade to a student’s learning from formative and summative assessment that is primarily based on qualitative, informed data.

A problem and benefit when using rubrics for CRA is that there is no one way to convert the outcomes into letter grades (Mertler, 2001). Yet, teachers want solutions to problems and generally seek “to reconcile innovative student assessment designed to be standards-based and influence instruction with the need to grade” (Boston, 2002, p. 34). Boston notes four common approaches when converting rubric outcomes to letter grades: *Frequency of Scores Method* (logic based), *Total Points*, *Total Weighted Points*, and *Linear Conversion*. The last three methods of conversion –being based on number computations and percentages – seem more scientific than they should be (Boston, 2002). Though attaching numeric values to the BCPS suggested performance rubrics provides a simple solution, doing so contradicts the philosophy, nature and intention of the type of assessment itself. Ultimately, classroom teachers “must find a system of conversion that works for them and fits comfortably into their individual system of reporting student performance” (Mertler, 2001).

Deficiencies in the Evidence

Locating research and resources on how teachers develop and defend letter grades based on FA and CRA rubric results is difficult. My observations and experience in Middle Schools indicate that teachers are typically left to their own devices in ‘making it work’ but that the

capacity needed (e.g., training, resources, professional development, time, motivation) to support and use FA and CRA rubrics as they are intended falls short. As a result, very few teachers practice FA and CRA in the same way, while others find it too strange, difficult or time consuming to even attempt using these tools and prefer the traditional number-total approach.

Data collected from CRA rubrics often varies from teacher to teacher bringing into question rubric reliability (Lovorn & Rezai, 2011). For example, while one teacher may assess a student's work as 'Fully Meeting Expectations' another may assess it as 'Meeting Expectations'. This creates discrepancies in standardized test results where rubrics are used, especially when teachers alter the descriptive language to suit numeric values – a common practice I've experienced in Middle Schools that use standardized reading assessment packages such as the QCA or RADs. If discrepancies exist in standardized test assessment, then the reliability and validity of the outcomes must come into question.

The Audience

As researchers proclaim a paradigm shift in education and promote 21st Century Learning or Personalized Learning as its gateway, teachers are left to create a balance between using modern assessment rubrics designed to improve student learning and fulfilling policy. In exploring how teachers currently using FA and CRA rubrics as assessment *for, as* and *of* learning, we not only acknowledge difficulties that come with transposing qualitative assessment into the narrow parameters of letter grades, but also share innovative strategies to create a classroom environment that focusses on student learning and not merely academic achievement. With this knowledge, more dialogue for discussion can be created among teachers, administration, and the Ministry of Education on how to further support and evolve assessment

practices in the interest of creating learning that has depth, breadth and purpose while fulfilling mandated policy and practice for reporting student progress.

Research Question

Central question. “How do middle school teachers in School District 43 use the BCPS suggested CRA rubric language and matrix while needing to meet Ministry guidelines for student progress reporting?”

Sub-questions. “How does teachers’ practise of the BCPS compare to its intent?”; How do teachers assign and defend a letter grade to a student’s learning based on FA and CRA rubric results?”; “What common strategies have teachers developed in using FA to better students’ learning?”

LITERATURE REVIEW

Removing Subjectivity from Teacher Judgement

In understanding how rubrics are used toward assessment and evaluative practice, it is important to know why rubrics came about at all. Answering a call in education circa 1912 to create a concrete standard measurement to judge writing compositions in the United States (US), Milo B. Hillegas, a Columbia University's Teachers College professor, created what is widely considered the original writing scale (Turley & Gallagher, 2008). The Hillegas scale, as it was commonly known, aimed to eliminate and replace subjectivity "for an objective and exact numerical measurement of student writing" (Turley & Gallagher, 2008, p. 88). Not without controversy, the scale was used as a scientific construct to quantify the quality of students' English compositions by providing artificially constructed samples with evaluative totals ranging from 0.0 to 9.0 (Trabue, 1922), as well as provide measures for teacher accountability (Trabue, 1922; Turley & Gallagher, 2008). More scientific assessment and measurement tools followed suit, (Turley & Gallagher, 2008); however the scales were designed to create standardization in assessment and evaluation, not to improve student performance (Hudelson, 1923; Wilson, 2006).

By the mid-twentieth century, US colleges were in need of standardized testing as part of their admissions criteria and process (Wilson, 2006). Changing from an open admissions policy for anyone who could afford it to one based on testable qualifications and standard ranking meant that public education needed to meet learning criteria established outside of schools (Wilson, 2006). In the early 1960s, the College Board spent a considerable sum of money to create a fairly reliable General Composition Test (Wilson, 2006). Paul Diederich, John French and Sydel Carlton of the Educational Testing Service "headed the study that would pave the

way for rubrics” (Wilson, 2006, p. 20). As standardized assessment tools continued to develop, the *rubric* (a theological term that began to emerge in the study of religion during the late 1980s) began to appear as a tool for grading curriculum-based tasks in education between 1990 and 1992 (Schneider, 2006). Today, holistic and analytical rubrics are widely used to assess and support student learning throughout North American schools.

Rubrics in Standardized Testing

The emphasis on standardized testing in the US with the No Child Left Behind (NCLB) act of 2001 is highly controversial. Opponents of the NCLB policy for testing and accountability in public education point out the weaknesses such as cultural biases, the sum of parts not being equal to the whole, difficulty in differentiating between matrix descriptors and variation in the methods of scoring in the evaluative rubrics used (Wilson, 2007; Kohn, 2006). The overwhelming negative effects of standardized testing on student learning in the US noted by many researchers created a need to analyze the effectiveness and use of rubrics. While some seek to re-evaluate or re-tool rubric intentions and designs (Andrade, 2005; Turley & Gallagher, 2008), others call for the elimination of rubrics as assessment tools because of the *lack of* subjectivity outlined in performance matrices, especially for student writing assessment (Wilson, 2006; Kohn, 2006). The controversy over standardized testing like in the US is seemingly not as prevalent in Canada. Standardized tests and performance criteria vary between provinces or territories, and though some may share the same test (e.g., The Yukon Territory and Alberta share the same standards), all have unique scoring rubrics – either holistic or analytical in type – used to evaluate student abilities in reading, writing and numeracy.

In British Columbia – the region of focus for this research – the provincial government requires the annual implementation of the Foundations Skills Assessment (FSA) for Grade 4 and

7 students in public and provincially funded independent schools. The FSA has been a contentious issue between the provincial government and the BC Teachers' Federation (BCTF) for a little more than a decade. Since the current government came to power in 2002 there has been what the British Columbia Public School Employers' Association (BCPSEA) characterizes as a "*philosophical divide*" (BCPSEA, 2010). With differences in philosophies aside, the BCPS CRA rubrics for reading, writing and numeracy hold value in what is expected for students to learn, and act as a teacher's resource by providing *suggested* assessment strategies for teachers using BCPS rubrics (www.bced.gov.bc.ca). However, when teachers follow the BC Ministry of Education's (BCME) policy, practice and suggestions for grading and reporting student progress, as well as use the rubric designs and strategies outlined in the BCPS, teachers are left to their own devices and professional judgement in needing to correlate rubric outcomes to letter grades.

The Practice of Rubric Conversion for Mandated Reporting

The BCME's Reporting Student Progress: Policy and Practice (RSPPP; 2009) document states that -

criterion-referenced letter grades in Grades 4 to 12 indicate students' level of performance in relation to the prescribed learning outcomes set out in provincial curriculum guides, ... Letter grades must be included on report cards in Grades 6 to 12.

(p. 7)

The RSPPP (2009) further states that "teachers must use the approved letter grades as set out in the Provincial Letter Grades Order. ... For students in the 2004 Graduation Program, percentages must accompany letter grades for courses numbered 10, 11 and 12" (p. 13; see Figure 1). An anecdotal descriptor for each letter grade is provided as well, and is intended to be used more for Grades 4 to 9 (RSPPP, 2009).

Percentages Associated with Letter Grades		
A	86	100
B	73	85
C+	67	72
C	60	66
C-	50	59
F	0	49

The successful completion of a course numbered 10, 11 or 12 requires a minimum of a C-.

Figure 1. Source: RSPPP, 2009, p. 13.

The BCPS were developed for voluntary use and “intended as a resource to support ongoing instruction and assessment” (RSPPP, 2009, p. 20). Frequent references are made to criterion referenced assessment and evaluation throughout the RSPPP, but only once does the document provide a rubric scoring example. A holistic rubric based on fictitious performance criteria established by a teacher and students for a “how-to” writing assignment in a secondary English course, the Ministry’s example provides little insight in how to develop a reliable and valid rubric or how to justify students’ marks according to rubric assessment or evaluation, especially if a grade of C- or C+ would be more appropriate (Figure 2).

A number of significant problems and questions emerge for teachers when analyzing and using a rubric such as this to assess or evaluate a students learning in relation to mandated letter grades. Beyond the fact that the descriptors for level *THREE: (B)* and *FOUR: (A)* provide little objective variance to help in distinguishing between either outcome, the four achievement level indicators are provided as numbers *and* equivalent letter grades, not as adjectival phrases such as those outlined the BCPS suggested rubric matrix (e.g., Fully Meets Expectations). Nowhere within the RSPPP or BCPS documents is a clear-cut method for converting rubric outcomes (especially adjectival performance level indicators) to legitimate letter grades described.

ONE: (I)	These are not really directions. It is impossible to do the activity using the directions. Only one or two steps are given. There is no introductory or concluding sentence. A maximum of two of the criteria are met.
TWO: (C)	Three or four steps are given in sequence. Of the eight criteria, up to four are present.
THREE: (B)	There are five to seven criteria of good instructions present, including numbers seven and eight. The rough copy shows evidence of revision.
FOUR: (A)	Six to eight criteria of good instructions are present including numbers six and seven. The rough copy shows evidence of revision.

Figure 2. Source: *RSPPP*, 2009, p. 27

Converting Rubric Outcomes for the Purpose of Grading

Boston (2002, p.vi) states that “When good rubrics are used well, teachers and students receive extensive feedback on the quality and quantity of student learning. The issue of then trying to feed this information into a more conventional grading system is complex and controversial”. In *Understanding Scoring Rubrics: A guide for teachers* (Boston, 2002), the Northwest Regional Education Laboratory (NREL) outlines suggested strategies for converting rubric scores to letter grades. Using a memo sent from Linda Elman, a district Testing Coordinator, to teachers of the Central Kitsap School District in Washington, the NREL not only addresses the issues surrounding methods for converting rubric outcomes to letter grades, but illustrates how such conversions can be accomplished. Elman explains that “there is no simple or single way to manipulate rubric scores so that they can be incorporated into end-of-quarter letter grades” (Boston, 2002, p. 35). She goes on to outline and describe four methods – all of which use number values and/or mathematical operations – for rubric score-to-letter grade conversions:

a) *Frequency of Scores*, b) *Total Points*, c) *Total Weighted Points*, and d) *Linear Conversion*. In her conclusion, Elman writes –

Once you, as teacher, arrive at a method for converting rubric scores to a scale that is comparable to other grades, the responsibility is on you to come up with a defensible system for weighting the pieces in the grade book to come up with a final grade for students. This part of the teaching process is part of the professional art of teaching.

(as cited in Boston, 2002, p. 38)

Though three of the four conversion methods require a mathematical approach, the NREL describes the conversions being more scientific in nature, and found that “many teachers dislike the total points and weighting methods because they don’t keep as much to the spirit of rubric use in the classroom” (Boston, 2002, pp. 39-40). However, the vast majority of literature on rubric outcome-to-grade conversion, including web-based or other electronically generated rubrics and teacher mark books, follow one of the four methods outlined in Elman’s memo. Admittedly, levels of achievement described solely through adjectives, like those indicated in the BCPS, are the most challenging to convert (Stix, 2002, p. 87). For Kohn (2006) and Wilson (2006), using adjective descriptors would be ideal for assessing and evaluating student performance, but could only exist in a system where traditional letter grades are not mandate as part of student evaluation and reporting.

Intended to guide and support teachers’ FA practice, the BCPS outlines and defines the expected qualities for each of the four performance levels in the rubrics for Not Yet meeting expectations (NY), Meeting (Minimally) Expectations (MM), Fully Meeting Expectations (FM) and Exceeding Expectations (EX) (www.bced.gov.bc.ca). In using language, not numbers, to describe and illustrate the four levels of student achievement in key areas of learning in terms of

prescribed learning outcomes, teachers must use professional judgement if attempting to correlate a letter grade with each descriptor. Having a four-level scale further complicates the summative process. Though the descriptors easily and adequately describe what the qualifications for NY, MM, FM, or EX, they do not provide clarity into correlating each descriptor to letter grades. For instance, while EX (“the work exceeds grade-level expectations in significant ways”; “the student may benefit from extra challenge”) easily constitutes the parameters of an A in an exceptional way, FM (“the work meets grade-level expectations”; “there is evidence that relevant prescribed learning outcomes have been accomplished”) (www.bced.gov.bc.ca/perf) – can be translated by teachers to have an equivalent anywhere from C to A. Moreover, having only four performance levels can present a challenge to teachers when evaluating students’ achievements to determine a final letter grade.

In *Demystifying the Evaluation Process for Parents: Rubrics for Marking Student Research Projects (2004)*, Joan Shaw describes her experience as a BC teacher librarian collaborating with other teachers in creating an assessment and evaluation rubric. Shaw (2004) and her colleagues “began with the performance standards developed by” (p. 17) the BCME . The teachers adapted the BCPS four-scale rubric model and created a five-scale model that included adjectival levels of achievement along with associated letter grades (A, B, C+, C, C-). In doing so, Shaw found that moving to a five-level rating scale was necessary and challenging to create, but that it allowed for the desired flexibility in assessing student work (Shaw, 2004). In my experience, the initial instinct for a teacher to alter and adapt the BCPS suggested four-level matrix by adding a fifth descriptor – generally between Fully Meets and Meets (Minimally) – provides the teacher a desired ‘middle ground’ when marking student work. However, the solution of moving to a five-scale rubric using either letter grade or number values associated

with letter grades is one of many simple methods for fixing the rubric-to-grade dilemma. “Only when teachers compare classroom-based rubrics to published standards, share them with another teacher, or co-score student work with a colleague, will rubrics improve” (Andrade, 2005, p. 30), and allow for more innovative solutions to emerge.

METHODOLOGY

Research Design

This study followed exploratory sequential mixed methods designs in collecting both qualitative and quantitative data. The research aims to inform, improve, and provoke further discussion surrounding teachers' current student assessment and evaluative practices when using rubrics, as well as provide points of reflection about the problem of converting rubric outcomes to letter grades while following suggested strategies and mandated policy on reporting student progress set by the BC Ministry of Education. Information and Data collection occurred in two phases: qualitative data gathered in the first phase allowed for an "in-depth exploration of a few individuals" while quantitative data from the second phase provided "the opportunity to gather data and generalize results" (Creswell, 2012, p. 548). Personal interviews and rubric samples documented from phase one provided a means for developing a quantitative instrument, forming categorical information for later quantitative data collection, and developing generalizations from a few initial qualitative cases (Creswell, 2012). Through an exploratory design analysis, commonalities were found within the sample set of phase one, and a survey was created from the emerging common themes to provide additional information as a numeric score.

A major disadvantage in using mixed methods designs is the need to attend to issues that are reflected in both the qualitative and quantitative data. The researcher conveyed the purpose of the study, obtained permissions, was aware of potential power issues in data collection, did not disclose sensitive information, and ensured anonymity (Creswell, 2012). In addition, the researcher took the necessary precautions and made every effort not to use his close relationship with the participants and to coerce data collection (Creswell, 2012). Participants had the right to opt out of the study at any given time if they chose to do so. The ethical needs in collaborating

with participants are to clarify how the results will be used and the promise of anonymity (Creswell, 2012). As such, the researcher was committed to open and transparent participation, respect for people's knowledge, and democratic and non-hierarchical practices (Creswell, 2012).

Participants

Participants were chosen through the purposeful sampling of middle school teachers (Grades 6 to 8) currently practicing in School District 43, British Columbia. Years of teaching experience varied among participants, as well as the social and cultural demographics of the students and community for which each teacher's site serves. Teachers were contacted via email, phone or face-to-face and were informed of the study and its intentions. In addition, a letter stating pertinent study information and formally inviting participants in the study was provided. Individuals signed and submitted a portion of the formal letter indicating their willingness to participate in the study, as well as indicate a preference to participate in group or individual interviews.

Two strategies were used in purposefully sampling teachers: homogeneous sampling and snowball sampling. Since middle school philosophies are pedagogically different from those of elementary and high school teachers, homogeneous sampling is a necessary procedure. Typically, middle school teaching is defined in how the needs of young adolescents in the areas of their intellectual, emotional, social, moral and physical development are met. In support of the unique characteristics of middle schools, the *"This We Believe" for School District 43 Middle Schools* was created to provide guiding principles and essential belief statements for instruction, assessment and middle level pedagogy. Snowball sampling arose during interviews and informal conversations with individuals from various sites as the researcher asked participants to

recommend colleagues that demonstrate particular innovative strategies for rubric use (Creswell, 2012). To this end, the research studied twelve individuals from five different sites.

Data Collection and Instruments

Personal interviews are the foundation of this mixed methods study in collecting qualitative information. Data collection consisted of small group and/or individual interviews. The researcher noted participants' responses, audio taped interviews, and stored all the information collected onto secure computer drives with sole access to the researcher. All interviews followed semi-structured characteristics, allowing the necessary flexibility in moving interviews and discussions toward potentially enlightening information. Interview questions revolved around the professional experience, opinions, values, feelings and knowledge of each participant. Individual rubric examples along with personalized outcome-to-letter grade conversion practice were requested from each interviewee, but it will not be a requirement of participation in the study. Participants who chose to provide examples did so in written or electronic forms (e.g., email, Word documents, or fax). Each interview session was facilitated by the researcher and began by reminding participants of the purpose of the study (Creswell, 2012). After a thorough analysis of the qualitative data gathered through the first phase, the researcher created an appropriate quantitative measuring instrument to be completed by the participants in Spring.

Phase One: Small Group Interview

The purpose of collecting data through a small group is to provide an arena for professional discussion and reflection, as well as to obtain teacher opinions as a group of people within the same community (Creswell, 2012). Small group interview formats are an efficient way of gathering data from multiple participants and encourage in-depth discussions about the

topic, which can lead to some consensus or strong disagreement in opinions. Moreover, further questions may arise from the discussions that can be used in interviews (Creswell, 2012). The process began with a few ice-breaker questions to get participants comfortable and engaged in dialogue. After this activity, the participants were asked to give and discuss their opinions on the questions that the researcher gave to them. The session lasted thirty minutes.

Possible disadvantages of the focus group setting are that certain participants may not feel comfortable expressing and defending their opinions, especially if some of the other participants have dominating personalities, are more articulate with their arguments, or are seen as having a position of authority based on years of teaching experience (Creswell, 2012). Some may feel hesitant about admitting their own lack of knowledge and skills in using rubrics towards student assessment and grading practices in front of the interviewer or other seemingly more knowledgeable or experienced teachers. In addition, certain participants could monopolize the group discussions and less data will be gathered from the others (Creswell, 2012). This would then require more effort on the facilitator's part to maintain some equity in the group (Creswell, 2012). The facilitator's skills were also required to ensure that the discussions did not go too far off topic so that more time could be used to gather the pertinent information.

Participants were encouraged to participate and allow others to participate equally as well. They were also reminded of their right to abstain from any portion of the process.

Phase One: Individual Interviews

This form of data collection allowed the researcher to form an intimate, trusting relationship with the participant that led to deeper, more honest and personal responses and explanations (Creswell, 2012). The number of individual interviews provided multiple sources of data that were used to build on and confirm each other (Creswell, 2012).

Personal interviews were an excellent method of acquiring qualitative data. All research participants were individually interviewed with the aid of an interview protocol. This was used to ensure that the interviews proceeded accordingly and that the interviewer could record notes during the process (Creswell, 2012). As with the small group interview, these sessions began by reminding participants of the study's purpose, along with some casual conversation to make the participant feel comfortable (Creswell, 2012). The format of the interviews was semi-structured, ensuring that the participants were given the opportunity to respond to all questions and yet have the freedom to elaborate on their responses as needed. Sessions lasted twenty to sixty minutes.

There are some disadvantages to the personal interview format. Participant honesty could be affected by personal pressures to respond a certain way (Creswell, 2012), feelings of being judged as an unknowledgeable or inadequate professional, or a lack of confidence toward the researcher keeping participant identity anonymous, especially if being quoted. Responsibility fell on the interviewer to ensure that these fears were minimized and that the participants could respond honestly and freely to all questions (Creswell, 2012). The questions were carefully phrased to avoid misinterpretation or misunderstanding, and the interviewer was also careful and tactful as not to influence the participants' responses. A final disadvantage to the individual interview process is that it required a lot of time to collect the data this way and each interview needed to be analyzed afterwards.

Phase Two: Survey

In developing the survey, the researcher identify questions connected to themes that emerged from the data collected during phase one. The researcher created and conducted a semi-structured, electronic survey generated through an online survey provider, and was comprised of 29 mostly closed-ended, Likert-type agreement questions. To remove the opportunity of

neutrality, individuals responded to each belief statement by choosing one of four agreement selections: Strongly Disagree, Disagree, Agree, Strongly Agree (see Appendix 1, Table 2). A formal invitation was emailed to all SD43 middle school teaching staff (N = 479) to elicit as many responses as possible. Anonymity was assured, and those who chose to complete the survey could do so only once. A total of 99 completed surveys were collected and tabulated through the online survey provide. Upon completion of the survey, participants were thanked, told of the next stages in the research and asked if they wished a copy of the study and its conclusions. Though the sample size may limit ability to “draw valid inference from the sample to the population” (Creswell, 2012, p. 382), the survey allowed analysis of the data in describing trends about responses to questions (Creswell, 2012). The trends can then be generalized to the population.

Limitations

The study’s findings are limited in its generalizability. This is because the sample size is relatively small, specifically comprised of middle school teacher and within the confines of School District 43. Teachers from different levels of schools (e.g., elementary, high school, independent) as well as other school districts will likely have different perspectives on pedagogy, rubric use and reporting practice on student progress and achievement. Also, findings are restricted to teachers of Grades 6 to 8, and therefore do not necessarily consider or reflect the opinions, thoughts and experiences of teachers of other grades, especially for Grades 10 to 12 for which a percent must accompany each letter grade assigned on end-of-term student reports. As such, purpose, methods and philosophy for rubric conversion in Grades 10 to 12 will be different than those of lower grades.

The reliability and validity of the study's design and implementation must be considered as a research limitation. Intrinsic to qualitative studies, reliability or possible reproduction of the research design is low (Creswell, 2012). Inconsistencies can be expected if different researchers were performing the interviews and/or if a different sample were interviewed. The validity of the study may also be questionable as participants' responses during interviews and the survey may not reflect their true practice due to the intimidating prospect of revealing inadequacies in pedagogical knowledge or ability, or by simply providing responses that lack care and attention.. The triangulation of the two interview methods and the member checking strategy should reduce these effects.

Ethics

Informed consent was obtained from the participants using a detailed Informed Consent Form outlining all pertinent information concerning the study, the role of participants as volunteers, and their right to withdraw or refrain from any portion of the interview process. It is also important to note that the researcher did not evaluate the participants and personal pedagogy – this is not the role of the researcher or the purpose of this study. This was explicitly stated in the informed consent form. All school and participant names were kept confidential. Pseudonyms were used in place of these. Pre- and post-interview reflections (as outlined in the Data Collection and Instruments section) were transposed to a Word document, and then shredded for confidentiality purposes.

DATA ANALYSIS

Information collected during Phase One (QUAL) of the research provided insight into the feelings, attitudes and opinions of middle school teachers have toward FA and the BCPS. Initial findings illuminated misunderstandings and differences in understanding teachers had of the BCPS intent and language. Interviews also highlighted how teachers try to make sense of FA and the BCPS achievements by correlating student performance with letter grades. Five over-arching themes were established from interviews and used to create the survey administered in Phase Two (quan). Those themes are used to outline and guide the data analysis.

Theme One: Teachers modify BCPS rubrics to better measure student performance.

All interviewees commented that BCPS CRA rubrics lack appropriate descriptors for students who do not completely fall into either FM or MM performance profile, but rather are somewhere in between. Nine of the 12 participants modified the BCPS four column rubric matrix by adding a fifth column between FM and MM to compensate for the achievement gap. The remaining three teachers acknowledged the gap but felt comfortable using the BCPS suggested matrix for formative assessment. When asked why he felt some teachers add a fifth column while he does not, teacher *K* responded, “Because they want to make it to make sense with letter grades. The M (meaning the added column) becomes the much needed C+.” Interviewee comments support *K*’s assertion.

Of the nine participants who preferred a five column rubric, one replaced the BCPS language with letter grades A to F where EX = A, FM = B, M = C+, MM = C/C-, and NY = I/F, three participants replace the BCPS language with number scores 5 to 1 where EX = 5, FM = 4, M = 3, MM = 2, and NY = 1, and five retained the overall BCPS language, but inserted a Meets Expectations achievement level between FM and MM. When the aforementioned five were

asked *How do you translate the language when it comes to report cards?*, all replied EX = A, FM = B, and M = C+, but MM and NY letter grade correlations differed slightly. Of the five participants, two associated MM with C and NY with C-, two associated MM with C/C- and NY as I (incomplete), and one associated MM as C/C- and NY as I/F. None the less, the interview information indicated that teachers agreed that the addition of a fifth column eased the challenges teachers face in converting performance outcomes to letter grades. For the purposes of fulfilling mandated reporting, such a method would be appropriate if the performance descriptors with each aspect were adjusted accordingly, but many did not do so (see Theme Two).

Data collected from the survey in phase two indicates that the majority of teachers in SD43 agree with phase one participants' sentiments that the BCPS suggested rubric matrix needs an additional column. Eighty-three of 99 respondents agree with the statement *I find that a fifth column needs to be added to the BCPS rubrics because the performance language used to differentiate Minimally Meeting from Fully Meeting does not account for student that perform somewhere in between*. Fifty-eight agree that they prefer using rubrics that have five or six levels of achievement, because it's easier to assign each column a letter grade equivalent. Of 83 teachers who agree that a fifth column is needed, 52, or 63%, prefer using a five or six column rubric. Of those who prefer using a five or six column rubric (N = 58), 40, or 69%, use the BCPS as a resource when creating rubrics, 41, or 71%, need to use numbers noted in their marks book to confirm letter grades at term's end, and 57, or 98%, agree with the statement *Parents want letter grades*.

Contrary to the common methods used for rubric outcome to letter grade conversion as outlined by Boston (2002), 65 teachers disagree with the statement *When converting rubric outcomes to letter grades, a mathematical approach is necessary*, and 41 of them do not feel the

need to use numbers to confirm letter grades a term end. Data also indicated that those who do not use numbers to confirm rubric outcomes or letter grades will generally record the majority of assessment as BCPS level of achievement in their mark books.

Theme Two: Different paths of assessment practice can be drawn from the depth and breadth in teachers’ understanding and use of FA and the BCPS.

Perhaps a factor in teachers modifying BCPS CRA rubrics is a lack of understanding of the performance standards intent and the language therein. Not knowing how the BCPS defines each level of achievement, specifically the Exceeds Expectations category, was a common factor among nine of the interviewees. Even though all participants claimed familiarity with the BCPS and use its language as part of their daily classroom assessment, only three of the 12 interviewed could paraphrase the BCPS definition of Exceeds Expectations. These three individuals – *K*, *L* and *Q* – share more commonalities than differences from the other participants.

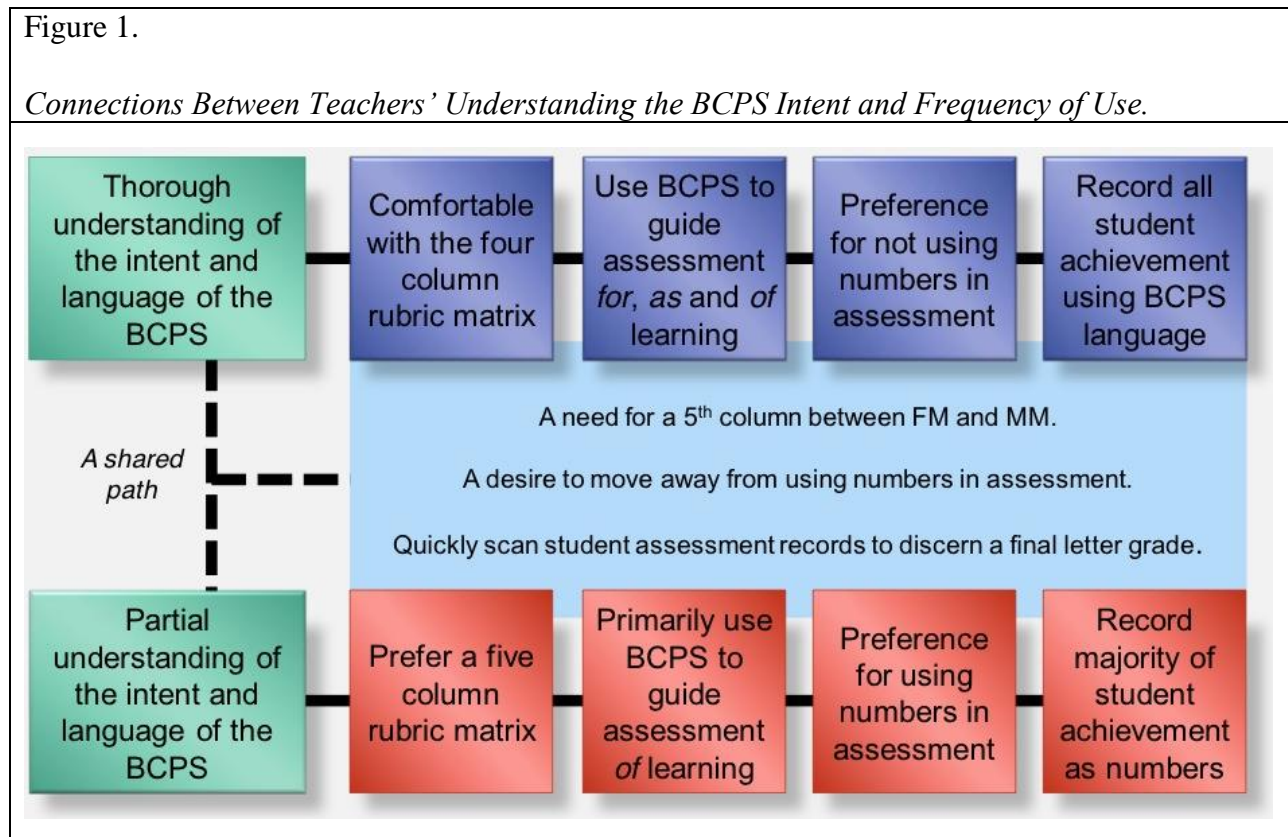
In comparison with other phase one participants, *K*, *L* and *Q* are unique in several ways. First, *K*, *L* and *Q* demonstrated a thorough understanding of the BCPS. They only use BCPS language when assessing student performance, have moved away from using numbers in their assessment practice altogether, and are completely comfortable using the BCPS four column rubric matrix and actually prefer it over a five column matrix. When asked *If you were to show me your marks book, what would I see?*, all stated that they record student achievement using the BCPS language only. Second, they have a passion and commitment to FA in the classroom. Though nine participants felt they did not do enough FA simply because it took too much time away from teaching curriculum, *K*, *L* and *Q* found the commitment of time far more rewarding because of increased student success and focus on learning, not letter grades. Third, they use the BCPS to guide their FA practice and assessment *for* and *as* learning, not for assessment *of*

learning. This is perhaps the greatest difference between these three interviewees and the others. All nine of the other participants mainly used the BCPS CRA rubric for assessment of learning.

A second theme emerged from noted differences between the group of *K, L* and *Q* and the rest of those interviewed (Figure 1). The grey area in the graphic contains common beliefs and practices shared between those who demonstrate a thorough understanding of the BCPS and regularly use the suggest CRA rubric matrix, and those who demonstrate a partial understanding of BCPS and occasional use of the CRA rubrics.

Figure 1.

Connections Between Teachers' Understanding the BCPS Intent and Frequency of Use.



Survey data indicates that 67% of those familiar with the BCPS as well as 95% of those that mainly use language descriptors record in their marks book want to move away from using numbers. Of those who agree that they do a lot of formative assessment (N = 70), 48, or 69%,

want to move away from using numbers and 43, or 62%, wish they did not have to give letter grades.

Theme Three: Teachers correlate letter grades to BCPS achievement levels, but there is no regularity or consistency.

Theme One, Two and Three are connected: all evolved from teachers trying to make sense of language-based achievement indicators and the BCPS rubrics. Theme Three, however, explores how misunderstandings, needs, and teaching variables lead to confusion and inconsistencies in converting BCPS language to letter grades. Metaphorically, teachers are trying to fit a square peg into a round hole.

In a personal communication with Don Gordon, Staff Development Coordinator for SD43, he explains the problem of correlating letter grades to the BCPS:

Based on my conversations with Sharon Jeroski who helped design them, there is no correlation between the Performance Standards and letter grades. There was never any intention to connect the two. The Performance Standards were designed to be used as formative assessment – assessment as and for learning, not so much assessment of learning in a summative evaluation sense. As I try to tell people, of the thousands of samples that were used to validate the scales and choose exemplars from, statistically only 2% would be exceeding and only 2% not yet meeting if these are used in late March/early April as intended. That is problematic for folks who want the Exceeding to be the A level. Realistically, the letter grade is based on a whole terms worth of evaluation, not just one score on a rubric. Fully meeting expectations could be an A, B, C+ or C, considering the reporting language of Excellent, Very Good, Good and Satisfactory.

It can be assumed that teachers who felt very familiar with the BCPS definitions of EX, FM, MM and NY would also agree with Gordon’s statement that FM “could be an A, B, C+ or C”. However, this is not the case. Further analysis suggests that teachers believe they know more about the BCPS that they really do.

Eight-eight percent of those who responded to the survey (N = 99) agreed with the statement *I am familiar with how the BCPS define Exceeding, Fully Meeting, Minimally Meeting and Not Yet Within Expectation*, of which 35% strongly agreed with the statement. When cross-tabulated with the open-ended response statement *When used to convert students’ performance assessments to letter grades, I translate the four levels indicated in the BCPS as having letter grade equivalents of*, surprising data emerged. As indicated in Table 1, the respondents’ conversions do not demonstrate a strong understanding of the BCPS definition but rather a lack of it. Not only does the modal correlation B (N = 11) support Gordon’s assertions, but connections can be made to those teachers identified in Theme One who prefer using five column rubrics, where FM = B.

Table 1	
<i>Letter Grade Equivalent of Those Very Familiar With BCPS Levels of Achievements Definitions</i>	
<u>Fully Meeting Expectations</u>	<u>Frequency</u>
A,B,C+,C	2
A,B,C+	5
A,B	9
B,C+	4
B	11
No Response	2
Total (N = 33)	

Tables 2.1 through 2.4 (see Appendix 3) indicate that letter grade correlations with BCPS language is very diverse. Assuming truthful responses, the following mode letter grade

equivalents emerge: EX = A (N = 79); FM = B (N = 33); MM = C (N = 27); NY = C- (N = 33).

This data further supports that Gordon is correct in his assertion that a teacher wanting “the Exceeding to be the A level” finds it difficult to correlate an FM achievement with an A level as well, let alone that it could also mean a B, C+ or C level. And it appears that parents share teachers’ confusion in understanding the BCPS language in needing to correlate each outcome to a letter grade.

Theme Four: Using the BCPS language to communicate academic performance to parents poses challenges.

Phase one participants were asked to provide insight in to parent reactions and preferences in communicating student achievement. All interviewees agreed that parents want and feel that they understand letter grades more than the BCPS language. *Q*, teaching in a low socio-economic middle school community, mentioned that the parents in his community were not so focussed on academic achievement, but that “They’re just happy with all the stuff we have going on for the kids outside the classroom. We take care of their kids.” Teacher *H*, serving a high socio-economic middle school community, shared a completely different experience to *Q*’s. “Everyone (teachers) around here uses numbers. When I first came to this school and used this rubric for Science (referring to his BCPS modified rubric language descriptors), I had a parent come in upset because he didn’t know what it meant by his kid getting an FM on the rubric. I tried explaining it, but he just didn’t get it. He needed me to translate it as a number or letter grade.” Teacher *S*, a French Immersion teacher working in a low socio-economy middle school community, regularly experience much the same a *H* saying, “They want letter grades. They want numbers. I tried for a while (using the BCPS language), but I just got tired of trying to explain what Fully Meets means. They wanted proof of why their kid got a B and not an A. They need numbers. Some time it got really bad.” In fact, *S* feels so much parental pressure to

communicate student achievement as a number that she no longer can be bothered using the BCPS language, saying, “The stress isn’t worth it.”

Data from the survey supports *H* and *S*’s assertions. 61 of 96 agree that *Parents find it difficult to understand the BCPS language of assessment (EX, FM, MM, NY) and how they indicate levels of academic performance*. Of those teachers who believe parents do not understand the BCPS language (N = 61), 55 feel a need to use numbers to confirm letter grades, and 30 of 32 who believe do not understand the BCPS language believe rubric conversion requires a mathematical approach. Regardless of years of teaching experience or demographics 89 of 98 agree that *Parent want letter grades*. Of teachers who believe parents want letter grades (N = 89), 54 need to use numbers to confirm letter grades and 34 mainly use language descriptors in their marks book. Of those who want to move away from using numbers (N = 63), 54 believe parent want letter grades. Of those who wish they didn’t have to use letter grades (N = 58), 49 agree that parents want letter grades. Of those that disagree with the statement *The majority of assessment data in my marks book is recorded as a language descriptor* (N = 58), 44 agree that parent don’t understand the BCPS. The data also indicates that the more affluent the school’s demographics, the more teachers believe that parents don’t understand the BCPS language.

Theme Five: Assessment *for* learning occurs more often than assessment *as* learning.

A final theme worth exploring evolved around interviewees’ practice of FA. All participants agreed that doing formative assessment well required time to build the necessary skills for both themselves and students but that time was in short supply. Though all participants believe that they do a lot of FA, the descriptions of strategies used by eight participants suggest that it is primarily assessment *for* learning –teacher self-reflection and guidance – and that they

used very few strategies for assessment *as learning* – student self-reflection and guidance). All participants believe that the purpose of a rubric is to provide students guidance and expectations for a given assignment, but the rubrics are primarily used for assessment *of learning*. Only two teachers, *K* and *Q*, regularly spend time building assessment criteria, direction and rubrics for assignments with students, while other participants avoid doing so, or did not do it as much as they would like, because this type of FA strategy takes too much time that they felt could be used more effectively elsewhere, such as teaching curriculum. A comment made by one teacher is a summation of how all others feel: “I find that building rubrics with kids takes too much time. I tried it once and spent two blocks trying to lead them through it. But it still wasn’t enough time. They just couldn’t do it. So now I just give them a rubric. I find that it’s just easier to give them the rubric.”

Survey data indicates that not building rubrics with students because of too much time spent on the activity is common practise. Seventy-one of 97 respondents believe that they do a lot of formative assessment in their classrooms, 62 of 99 agree with the statement *I frequently have my students self-ass their work before they submit it to me*, and 56 of 97 agree with the statement *I find that designing a rubric with students takes up too much valuable time. It’s quicker if I make the rubric on my own*. Of those who believe they do a lot of formative assessment in classroom (N = 71), 53 frequently have students self-assess work, 48 want to move away from using numbers, and 42 wish they did not have to use letter grades. Of those who frequently have students self-assess work (N= 60), 31 agree that it is quicker to make a rubric on their own. Of those that feel they do a lot of formative assessment (N = 69), 39 believe it is quicker to make the rubric on their own.

Theme Six: Moving away from numbers in assessment is desirable but personally challenging.

All but one of those who participated in phase one interviews indicated a desire not to use numbers in communicating assessment to students. While four teachers have moved to the point of not using numbers at all, seven are slowly flattening the field of quantitative assessment strategies that they have been so familiar with. Teacher *W* believes that the reason she feels more comfortable using number scores in assessment is because “that’s the way I was taught, and I guess I teach how I was taught.” *W* explained further that though she feels a desire to move away from numbers, she does not know how to and found that her past attempts left doubting the accuracy of her assessment. Numbers are infallible in her opinion. Of all participants, three teachers recorded mainly quantitative data in their mark books, two recorded a healthy mix of numbers, letter grade and BCPS language, two mainly used the BCPS language but included numbers as well, three used only the BCPS language, one used only letter grades, and one only use number scores. When asked *When it comes to report cards, how do you determine a letter grade?*, all participants replied that they quickly scan their data looking for a mode level of achievement. Those who use numbers would sometimes calculate the mean for a few assignments to confirm the observed mode.

Survey results indicate that 67% of teachers would like to move away from using numbers in their assessment, but that only 39% record the majority of assessment data as a language descriptor. Fifty-seven percent feel a need to use numbers in some way to confirm letter grades assigned on report cards. Of those who need to use numbers to confirm letter grades, 45% want to move away from using numbers and 96% feel parents want letter grades. Regardless of what method is used to record students’ achievement, 69% of teachers quickly

scan the data in their mark books and look for a norm level of achievement for each student to determine an overall letter grade for report card.

Innovation in FA and Rubric Strategies.

Only two teachers, *K* and *L*, demonstrated formative assessment and rubric practice that substantially differed from other interviewees'. *K* used an approach to increasing rubrics' informative accuracy by adding a plus or minus mark to a level of achievement when deemed necessary (e.g. FM-). The additional indicator communicates the degree to which a student has exceeded, fully met or minimally met CRA performance standards. Fifty-six of 98 survey responses agree that *Using a plus or minus to indicate the strength of how well a student performs on an assignment makes for more accurate rubric assessment*. The indicators also aid in converting a student's overall performance into a letter grade for report cards. *L* described a unique approach in using rubrics for formative assessment. Because *L*'s goal is to have all students fully meet expectations, and believing that some students are content at aiming for a personal performance that minimally meets expectations, *L* only has students use an FM column to guide them *as* they learn and to guide her assessment *for* their learning. When assessing students' work, *L* would provide oral and written anecdotes that included areas in need of further development and approaches for improvement. Both *K* and *L* feel very strongly about the importance of students responding to focus questions that require self-reflection and goal setting on the vast majority of assignments; both *K* and *L* are well versed in the BCPS intent and performance language and are highly focussed on FA practice in their classrooms. Neither *K* nor *L* uses numbers of any kind in their assessment practice. In fact, four of twelve teachers interviewed during phase one responded that they use rubrics for every assignment and that they have completely moved away from using numbers of any kind in their assessments of student

performance. Both *K* and *L* argue against using the language of exceeds expectation as part of rubrics used in the classroom. In their opinion, the term contradicts the ability to set such expectations, and doing so would hinder a student's own path and creativity to exceeding expectation. As *L* said, "How can I tell them (students) what exceeding my expectations looks like. It just seems odd to tell them what exceeding my expectations looks like by providing them those expectations. I don't even know what that looks like."

CONCLUSION

The majority of SD43 middle school teachers seem to lack understanding of the BCPS as a whole and are using it differently than intended in order to match their teaching practices and purposes with mandated reporting practice. Further inquiry is needed to determine why this is, and how it can be avoided in the future. Though the intent of the BCPS is for it to be used as a formative resource to aid and guide teachers to assess student performance in March or April, its suggested language and rubric design seem to be in conflict with how teachers perceive, need and use traditional grading practice to complete report cards. That being said, there are many courageous teachers within SD43 middle schools who are attempting to use the BCPS as intended and experimenting with FA strategies both in classrooms with students and in communication with parents.

Teachers' perceptions and personal experiences in communicating students' academic performance to parents can at times influence the type of assessment strategy a teacher uses. Many attitudinal variables, including those of parents, students, the general public, administration, and those of the teacher affect assessment and reporting practice.

Using FA and BCPS CRA rubrics continues to be an area of development for teachers. The findings in this research contradict literature that suggests the need for mathematical calculations when converting rubric outcomes to letter grades, but it confirms those that suggest teachers are left to their own devices to 'figure things out.' However, the data collected from teachers correlating letter grades to BCPS language suggests that a common system may be necessary.

As teachers explore and experiment with alternative forms of FA, the need, comfort and confidence found in using numbers to assess and communicate a student's level of achievement and overall letter grade diminishes, eventually leading to qualitative data being more valuable than quantitative. To support further development in FA and using the BCPS as intended, more capacity development is needed. As demonstrated with teachers' lack of understanding of the BCPS language and documents therein, simply making a resource available does not ensure that it will be used appropriately, or at all. Professional development focused on FA and the BCPS, along with the encouragement and support from administration and school districts for teachers to experiment, is necessary for growth and innovation in assessment to occur.

REFERENCES

- Alberta Education. (2011). *Examples of the standards for students' writing 2011: English language arts grade 3*. Retrieved on May 28, 2012 from: http://education.alberta.ca/admin/testing/achievement/examples_stand.aspx
- Andrade, H. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, 53(1), 27.
- Black, P. J. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139-144.
- Boston, C., & ERIC Clearinghouse on Assessment and Evaluation, C. D. (2002). *Understanding Scoring Rubrics: A Guide for Teachers*.
- British Columbia Public School Employers' Association.(2010). *Backgrounder: The foundation skills assessment*. Retrieved on June 2, 2012 from: <http://bcpsea.bc.ca/>
- British Columbia Ministry of Education. (2011a). *Reporting student progress: Policy and practice (2009)*. Retrieved on May 10, 2012 from: http://www.bced.gov.bc.ca/policy/policies/student_reporting.htm#
- British Columbia Ministry of Education. (2011b). *BC Performance Standards*. Retrieved on May 8, 2012 from: http://www.bced.gov.bc.ca/perf_stands/
- Clark, I. (2011). *Formative assessment: Policy, perspectives and practice*. *Florida Journal of Educational Administration & Policy*, 4(2), 158-180.
- Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th edition)*. Boston: Pearson.
- Hudelson, E. (1923). The development and comparative values of composition scales. *English Journal*, 12(3), 163-68.
- Jackson, C. W., & Larkin, M. J. (2002). *RUBRIC: Teaching students to use grading rubrics*. *Teaching Exceptional Children*, 35(1), 40-45.
- Kohn, A. (2006). *Speaking my mind: The trouble with rubrics*. *English Journal*, 95(4), 12-15.
- Loveless, J. (2006). *Going gradeless: Evaluation over time helps students learn to write*. *Language Arts Journal of Michigan*. 22(2), 12-17.
- Lovorn, M. G., & Rezaei, A. (2011). *Assessing the assessment: Rubrics training for pre-service and new in-service teachers*. *Practical Assessment, Research & Evaluation*, 16(16). Retrieved on May 12, 2012 from: <http://pareonline.net/pdf/v16n16.pdf>

- Mertler, C. A. (2001). *Designing Scoring Rubrics for Your Classroom*. Practical Assessment, Research & Evaluation, 7(25). Retrieved on May 13, 2012 from: <http://pareonline.net/getvn.asp?v=7&n=25>
- Northwest Regional Educational Laboratory (Oregon). (2002). *Converting rubric scores to letter grades*. In C. Boston & ERIC Clearinghouse on Assessment and Evaluation (Eds.), *Understanding Scoring Rubrics: A Guide for Teachers* (pp. 34-40). Portland, OR.
- Schneider, F. J. (2006). *Rubrics for teacher education in community college*. Community College Enterprise, 12(1), 39-55.
- Shaw, J. (2004). *Demystifying the evaluation process for parents: Rubrics for marking student research projects*, Teacher Librarian, 32(2), 16.
- Stevens, J. (2009). Washington State diagnostic assessment guide. MESA University of Oregon. Retrieved on May 10, 2012 from: <http://www.k12.wa.us/CurriculumInstruct/DiagnosticAssessmentGuide.aspx>
- Stix, A. (2002). *Creating rubrics through negotiable contracting*. In C. Boston & ERIC Clearinghouse on Assessment and Evaluation (Eds.), *Understanding Scoring Rubrics: A Guide for Teachers* (pp. 66-71). New Rochelle, NY: The Interactive Classroom.
- Trabue, M. R. (1922). *Supplementing the Hillegas Scale: A description of the derivation and use of the Nassau County supplement to the Hillegas composition scale*. Teacher College Bulletin, 13(12). New York: Teachers College, Columbia University.
- Turley, E. D., & Gallagher, C. W. (2008). *On the "Uses" of rubrics: Reframing the great rubric debate*. English Journal, 97(4), 87-92.
- Volante, L. & Beckett, D. (2011). *Formative assessment and the contemporary classroom: Synergies and tensions between research and practice*. Canadian Journal of Education, 34(2), 239-255.
- Wilson, M. (2006). *Rethinking rubrics in writing assessment*. Portsmouth, NH: Heinemann.
- Wilson, M. (2007). *Why I won't be using rubrics to respond to students' writing*. English Journal, 96(4), 62-66.

APPENDICES

Appendix 1: Interview Protocol.....38

Appendix 2: Survey Instrument40

Appendix 3: Survey Data (Tables 2.0 thru 2.9).....43

Appendix 1: Interview Protocol

Time of Interview:

Date:

Place:

Interviewer:

Interviewee:

Position of Interviewee:

The purpose of this study is to share insight, innovations and common strategies teachers have developed to use CRA rubrics for reporting on student progress. In addition, the research aims to explore the philosophy and pedagogy of teachers to identify ways that improve and promote student learning.

Data in the form of personal responses to set questions, as well as rubric samples voluntarily provided by the interviewee, will be collected from 10 to 15 middle school teachers (grades 6 to 8). Interviews will be taped and transposed to a computer document at a later date. All information collected, including electronic files, will be saved to a personal computer drive and folder only accessible by the researcher. The anonymity of the interviewee is insured in that real-life names and locations will be replaced with aliases that will not identify or link the interviewee to his/her current teaching assignment. The interviewee's personal responses are confidential and will not be shared with other interviewees. Information and data collected through the interview process will be used to create a survey that will be completed by a sample of middle school teachers within School District 43.

The interview time is expected to last fifteen to thirty minutes, depending on the depth of responses to each question.

Interview Format:

Introduction: First off, I'm not here to make any judgements on your teaching or how you prefer to assess and report on students' progress. I'm simply interested in how you use rubrics to for the purposes of assessing student progress and determining letter grades for report cards.

Questions:

Teacher Background & Rubric Design

- How long have you been teaching for? How many years at the Middle School level?
- What does a typical rubric matrix look like in your classroom? (Describe and draw; explain why that style is preferred)
- Do your students ever participate in designing the rubrics? In what situations do you find this useful or not useful?
- Since you first started using rubrics, how has your approach changed?

Formative and Summative Assessment

- Do you use rubrics more for formative assessment or summative assessment?

- Run me through how you would typically use your rubric to evaluate student performance?
- How do you come to a conclusion for a student's overall performance according to the rubric?
- Do you also use the rubric results as a guide for your own teaching?
- What does your current formative assessment practice involve?

Student and Parent reactions

- How do you use the rubric results to provide feedback for students?
- What do you find students or parents prefer: qualitative or quantitative assessment?
- What are students' typical reactions to rubrics and their use in assessing performance?
- Do the demographics of your schools community influence the method (numbers, letter grades, or BCPS language) you use for student assessments?

Mark Book & Reporting

- How do you record the rubric results in your mark book?
- If I were to look at your mark book, what would I see?
- When it comes time to completing report cards, how do you convert your rubric results to a letter grade?

Philosophy toward using letter grades/rubrics to improve student learning

- What approaches to assessment do you prefer and why?
- What do you see as the pros and cons of letter grades?
- Do you know of any middle school teachers who you recommend I contact regarding their use of rubrics?

If the teacher's rubric differs from the BCPS matrix, ask the following

- Have you ever used the BCPS suggest four levels of student performance language (NY, MM, FM, EX) for the rubrics in your class?
- For the purpose of evaluating and reporting on individual student performance, how do you/would you translate the four levels (NY, MM, FM, Ex) provided by the BCPS as letter grades?
- Are you familiar with how the BCPS defines each of the performance levels, particularly Exceeds Expectations.
- What are your thoughts on using the four levels of achievement?

Appendix 2: Survey Instrument

<u>Agreement Statement</u>	<u>Strongly Disagree</u>	<u>Disagree</u>	<u>Agree</u>	<u>Strongly Agree</u>
1. I am familiar with how the BC Performance Standards (BCPS) define Exceeding, Fully Meeting, Minimally Meeting and Not Yet Within Expectations.	○	○	○	○
2. I find that a fifth column needs to be added to BCPS rubrics because the performance language used to differentiate Minimally Meeting from Fully Meeting does not account for students that perform somewhere in between.	○	○	○	○
3. The language of EX, FM, MM and NY is too 'fluffy' for students. Students take their learning more seriously when the achievement level indicators are letter grades or number scores.	○	○	○	○
4. My students need to be evaluated using percentages and letter grades because that is what they will encounter in high school.	○	○	○	○
5. I'd like to move away from using numbers in my assessment practice.	○	○	○	○
6. I wish I didn't have to give letter grades.	○	○	○	○
7. At my school, teachers often discuss and attempt non-traditional assessment strategies.	○	○	○	○
8. When I create rubrics, I use the BCPS as a resource and guide.	○	○	○	○
9. Students take an assignment more seriously if the rubric uses letter grades or number scores, not EX, FM, MM, or NY.	○	○	○	○
10. Rubrics are best used for evaluating a student's performance on a writing assignment, not test.	○	○	○	○
11. Using a plus or minus to indicate the strength of how well a student performs on an assignment (e.g. FM+, FM-, A+, A-) makes for more accurate rubric assessment.	○	○	○	○
12. I do a lot of formative assessment in my classroom.	○	○	○	○
13. I frequently have my students self-assess their work before they submit it to me.	○	○	○	○

14. I find that designing a rubric with student takes up too much valuable time. It's quicker if I make the rubric on my own.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Formative assessment and summative assessment are basically the same thing.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Parents find it difficult to understand the BCPS language of assessment (EX, FM, MM, NY) and how they indicate levels of performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. The vast majority of my students are indifferent as to whether or not I use BCPS language, letter grades, or number scores to assess their work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. Parents want letter grades.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. I prefer using rubrics that have five or six levels of achievement because it's easier to assign each column a letter grade equivalent (e.g. EX = A, FM = B, M = C+...; 5 = A, 4 = B, 3 = C+...).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. I feel a need to use numbers in some way to confirm letter grades I assign for report cards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. The data I collect through formative assessment influence the letter grade I assign a student on his or her report card.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. When converting rubric outcomes to letter grades, a mathematical approach is necessary.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. The majority of assessment data in my marks book is recorded as a language descriptor (e.g. EX, FM, FM-, MM+, M...).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. I quickly scan the data in my marks book and look for a norm level of achievement for each student to determine and overall letter grade fro report cards. Outliers are generally omitted in this process.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. When used to convert students' performance assessments to letter grades, I translate the four levels indicated in the BCPS as having letter grade equivalents of: (open ended responses)	<i>Exceeds =</i> <i>Fully Meeting=</i> <i>Minimally Meeting=</i> <i>Not Yet Within=</i> <i>I'm not sure=</i>			
26. Years of teaching experience.	(0-5) <input type="radio"/> (6-10) <input type="radio"/> (11-20) <input type="radio"/> (21-30) <input type="radio"/> (31+) <input type="radio"/>			
27. Years of teaching experience in Middle Schools.	(0-5) <input type="radio"/> (6-10) <input type="radio"/> (11-20) <input type="radio"/> (21-30) <input type="radio"/> (31+) <input type="radio"/>			

28. My current teaching position is:	<i>Grade 6</i> <i>Grade 6/7</i> <i>Grade 7</i> <i>Grade 7/8</i> <i>Grade 8</i> <i>Grade 6/7/8</i> <i>Explorations</i> <i>Student Services</i> <i>A Mix of Responsibilities</i>
29. The families in our school's community are mainly of (you may choose more than one):	<i>Low socio-economic status</i> ○ <i>Mid socio-economic status</i> ○ <i>High socio-economic status</i> ○ <i>A good mix of all three</i> ○
Total completed surveys (N = 99)	

Appendix 3 – Survey Results (Tables 2.0 thru 2.9)

Table 2.0				
Converting Student Achievement To Letter Grades – Survey Data				
<u>Agreement Statement</u>	<u>Strongly Disagree</u>	<u>Disagree</u>	<u>Agree</u>	<u>Strongly Agree</u>
1. I am familiar with how the BC Performance Standards (BCPS) define Exceeding, Fully Meeting, Minimally Meeting and Not Yet Within Expectations.	8	4	52	35
2. I find that a fifth column needs to be added to BCPS rubrics because the performance language used to differentiate Minimally Meeting from Fully Meeting does not account for students that perform somewhere in between.	3	13	40	43
3. The language of EX, FM, MM and NY is too 'fluffy' for students. Students take their learning more seriously when the achievement level indicators are letter grades or number scores.	13	45	29	12
4. My students need to be evaluated using percentages and letter grades because that is what they will encounter in high school.	29	36	28	6
5. I'd like to move away from using numbers in my assessment practice.	3	29	43	21
6. I wish I didn't have to give letter grades.	10	30	29	29
7. At my school, teachers often discuss and attempt non-traditional assessment strategies.	1	32	59	6
8. When I create rubrics, I use the BCPS as a resource and guide.	4	24	57	14
9. Students take an assignment more seriously if the rubric uses letter grades or number scores, not EX, FM, MM, or NY.	13	41	33	10
10. Rubrics are best used for evaluating a student's performance on a writing assignment, not test.	10	19	52	18
11. Using a plus or minus to indicate the strength of how well a student performs on an assignment (e.g. FM+, FM-, A+, A-) makes for more accurate rubric assessment.	6	36	46	10
12. I do a lot of formative assessment in my classroom.	3	23	52	19
13. I frequently have my students self-assess their work before they submit it to me.	1	36	43	19
14. I find that designing a rubric with student takes up too much valuable time. It's quicker if I make the rubric on my own.	5	36	50	6
15. Formative assessment and summative assessment are basically the same thing.	41	45	13	0
16. Parents find it difficult to understand the BCPS language of assessment (EX, FM, MM, NY) and how they indicate levels of performance.	6	29	46	15
17. The vast majority of my students are indifferent as to whether or not I use BCPS language, letter grades, or number scores to assess their work.	6	35	50	7
18. Parents want letter grades.	2	7	54	35
19. I prefer using rubrics that have five or six levels of achievement because it's easier to assign each column a letter grade equivalent (e.g. EX = A, FM = B, M = C+...; 5 = A, 4 = B, 3 = C+...).	10	31	43	15
20. I feel a need to use numbers in some way to confirm letter grades I assign for report cards.	15	28	40	16
21. The data I collect through formative assessment influence the letter grade I assign a student on his or her report card.	1	5	66	26
22. When converting rubric outcomes to letter grades, a mathematical approach is necessary.	17	48	28	5
23. The majority of assessment data in my marks book is recorded as a language descriptor (e.g. EX, FM, FM-, MM+, M...).	15	45	27	11
24. I quickly scan the data in my marks book and look for a norm level of achievement for each student to determine and overall letter grade fro report cards. Outliers are generally omitted in this process.	5	24	57	13
25. When used to convert students' performance assessments to letter grades, I translate the four levels indicated in the BCPS as having letter grade equivalents of:	<i>*see Tables 2.1 through 2.4</i>			
26. Years of teaching experience.	<i>*see Table 2.5</i>			

27. Years of teaching experience in Middle Schools.	*see Table 2.6
28. My current teaching position is:	*see Table 2.7
29. The families in our school's community are mainly of (you may choose more than one)	*see Table 2.8
Total completed surveys (N = 99)	

Table 2.1	
<i>Survey Question 25 – Letter Grade Equivalent</i>	
<u>Exceeds Expectation</u>	<u>Frequency</u>
A+	3
A	79
No response	15
Total (N = 97)	
<i>Note: 2 responses indicating an A,B value were removed as outliers.</i>	

Table 2.2	
<i>Survey Question 25 – Letter Grade Equivalent</i>	
<u>Fully Meets Expectations</u>	<u>Frequency</u>
A,B+	1
A,B	23
A-,B+	1
B	33
A,B,C+	10
B,C+	13
A,B,C+,C	2
C+	2
C+,C	1
No Response	13
Total (N = 99)	

Table 2.3	
<i>Survey Question 25 – Letter Grade Equivalent</i>	
<u>Meeting Expectations (Minimally)</u>	<u>Frequency</u>
B,C+	1
B-,C+	1
C+	6
C+,C	16
C+,C,C-	6
C+,C,C-,I	1
C	27
C,C-	20
C-	8
No Response	13
Total (N = 99)	

Table 2.4	
<i>Survey Question 25 – Letter Grade Equivalent</i>	
<u>Not Yet Meeting</u>	<u>Respondents</u>
C	1
C,C-	6
C-	33
I	16
C-,I	9
F	9
C-,F	4
C-,I,F	1
D	1
Total (N = 97)	

Table 2.5.	
<i>Survey Question 26 – Years of teaching experience</i>	
<u>Current Teaching Position</u>	<u>Respondents</u>
0-5	15
6-10	31
11-20	42
21-30	10
31+	1
Total (N = 99)	

Table 2.6.	
<i>Survey Question 27 – Years of teaching experience in Middle School</i>	
<u>Current Teaching Position</u>	<u>Respondents</u>
0-5	25
6-10	32
11-20	42
21-30	0
31+	0
Total (N = 99)	

Table 2.7	
<i>Survey Question 28 – Current Teaching Position</i>	
<u>Current Teaching Position</u>	<u>Respondents</u>
Grade 6 Core	3
Grade 6/7 Core	44
Grade 7 Core	0
Grade 7/8 Core	0
Grade 8 Core	18
Grade 6/7/8 Core	4
Explorations	7
Student Services	8
A Mix of Responsibilities	15
Total (N = 99)	

Table 2.8	
<i>Survey Question 29 – School Demographics</i>	
<u>Socio-economic Status</u>	<u>Respondents</u>
Low	12
Low/Mid	13
Mid	21
Low/High	2
Med/High	10
High	9
Primarily low but mixed	1
A good mix of all three	30
Total (N = 99)	