

August 2018

## Grading Policies in Canada and China: A Comparative Study

Liyong Cheng

Queen's University - Kingston, Ontario, liyong.cheng@queensu.ca

Christopher DeLuca

Queen's University - Kingston, Ontario, cdeluca@queensu.ca

Heather Braund

Queen's University - Kingston, Ontario, heather.braund@queensu.ca

Wei Yan

Queen's University - Kingston, Ontario, yan.wei@queensu.ca

Follow this and additional works at: <https://ir.lib.uwo.ca/cie-eci>

---

### Recommended Citation

Cheng, Liyong; DeLuca, Christopher; Braund, Heather; and Yan, Wei (2018) "Grading Policies in Canada and China: A Comparative Study," *Comparative and International Education / Éducation Comparée et Internationale*: Vol. 47 : Iss. 1 , Article 5.

Available at: <https://ir.lib.uwo.ca/cie-eci/vol47/iss1/5>

This Research paper/Rapport de recherche is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Comparative and International Education / Éducation Comparée et Internationale by an authorized administrator of Scholarship@Western. For more information, please contact [tadam@uwo.ca](mailto:tadam@uwo.ca), [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

# Grading Policies in Canada and China: A Comparative Study

## Politiques d'évaluation au Canada et en Chine : Une étude comparative

**Liying Cheng**, Queen's University, Canada  
**Christopher DeLuca**, Queen's University, Canada  
**Heather Braund**, Queen's University, Canada  
**Wei Yan**, Queen's University, Canada

### Abstract

The current trend towards globalization, immigration, and internationalization of schools and universities around the world has led to the increased use of grades across educational systems. Given the use of grades for student promotion, mobilization, and admission into educational programs internationally, there is an urgent need to understand how grades are constructed differently in diverse systems of education. This study specifically examines grading policies across two educational contexts—Canada and China—to gain an nuanced understanding of how grades are constructed in these two systems where we see a large fast increase of Chinese students studying at Canadian tertiary institutions. This comparative analysis of Ministry of Education documents *within* and *across* these two learning contexts indicates significant differences in policies that guide teacher constructed grades in Canada and China. In Canada, achievement is the primary consideration in the construction of classroom grades, whereas grades in China include considerations of both the learning (i.e., achievement) and the learner (i.e., learning skills and personal dispositions). The findings of the study have significant implications for understanding the validity of grade interpretations across educational systems.

### Résumé

La tendance actuelle envers la globalisation, l'immigration et l'internationalisation des écoles et universités à travers le monde a menée à augmenter l'usage des notes dans le système éducatif. Étant donnée l'utilisation des notes pour la promotion des étudiants, la mobilisation et l'admission dans des programmes éducatifs internationalement, il y a un urgent besoin de comprendre comment les notes sont conçues différemment dans divers systèmes éducatifs. Cette étude examine spécifiquement les politiques de classement dans deux contextes éducatifs—le Canada et la Chine— afin d'obtenir une compréhension nuancée de la construction des notes dans ces deux systèmes, lesquels sont marqués par une augmentation rapide du taux d'étudiants chinois inscrits dans des institutions supérieures au Canada. L'analyse comparée de documents du Ministère de l'Éducation, *interne et transversale* à ces deux contextes éducatifs, indique des différences importantes dans les politiques qui guident la construction des notes des enseignants au Canada et en Chine. Au Canada, la réussite est la considération principale dans la conception des notes en salle de classe, alors que les notes en Chine incluent des considérations sur l'apprentissage (réussite) ainsi que sur l'apprenant (aptitudes d'apprentissage et dispositions personnelles). Les résultats de cette étude ont des implications significatives pour comprendre la validité de l'interprétation des notes à travers les systèmes éducatifs.

**Keywords:** grading; government policy; educational systems; learning values; China; Canada  
**Mots clés :** classement, politique gouvernementale, système éducatif, valeurs d'apprentissage, Chine, Canada

### Introduction

The current trend towards globalization, immigration, and internationalization of schools and universities around the world has led to the increased use of grades across educational systems. Grades are the dominant currency that enables student migration patterns; in particular, the recent upsurge of Chinese students studying and settling in North America, Canada in particular. Given the use of grades for student promotion, mobilization, and admission into educational programs internationally, there is an urgent need to understand the function and construction of grades across learning contexts. However, despite their significant uses, very little is known

about the comparability of grades given differences in learning cultures (Brookhart et al., 2016; Brookhart, 2013; Guskey, 2011; Randall & Engelhard, 2009; Sun & Cheng, 2013).

This study examines grading policies that describe the purpose and composition of school grades within two distinct learning cultures: Canada and China. Since 1991, Chinese students have consistently been the largest international and immigrant population in Canadian schools and universities with Chinese students representing the largest foreign student population at 34.1% in 2013–14 (Statistics Canada, 2016); this statistic is on the rise. Grades are used for high-stakes decisions for these students, including admission, placement, and immigration to English-speaking countries and are the primary decision-making indicator of whether to accept Chinese students into Canadian universities. However, currently grades are interpreted and used at the face-value; there is limited information on the consistency and comparability of grades between China and Canada. Understanding differences in grading policies in these two countries is critical to enabling valid interpretations of student achievement and admission decisions. Ensuring the validity of grade interpretation is foremost essential given the direct impact of grades on students who come to Canada to study and settle, and the impact on Canada itself—socially, educationally, and economically. The short-term implications are evident in changing school and community dynamics across Canada; the long-term impact will affect the availability of knowledge workers and citizens supporting the Canadian economy.

Currently, the lack of comparative research on grading practices across learning cultures provides unprecedented challenges to grade interpretation and grade use between educational systems (Sun & Cheng, 2013). Thus, this study begins to address this gap by examining grading policies within Canada and China. As grading policies are intended to provide theoretical and pedagogical guidelines to educators about assessment and grading practices, our research presents an initial attempt to understanding the function and construction of grades across Canada and China. Specifically, our study analyzes contemporary government-based grading policies in the two countries in order to examine the degree of consistency, if any, between grading in these two educational cultures. In particular, we were interested in understanding the consistency in how grades were constructed, the purposes of grades, and the relationship between grades and other forms of assessment practices within the contexts of Canada and China.

## **Grading Research**

Grading involves the process of summing student learning using a numerical or ordinal scale and is a complex assessment and evaluation practice that requires teacher judgment on a body of evidence (Brookhart, 2013; Brookhart et al., 2016; Guskey, 2011). Grades are used, most notably, to make public statements to students, parents, and principals about student achievement. Grading is one of the most high-stakes classroom assessment practices, with significant consequences to student self-perception, motivation for learning, prioritization of curriculum expectations, parental expectations, and social relationships inside and outside of schools (Brookhart, 2013).

Research on grading has a long history in education (Brookhart et al., 2016). Early educators (Dobbin & Smith, 1960; Finkelstein, 1913; Teaf, 1964) were primarily concerned with the reliability and standardization of teachers' grading practices. Recent research has further explored factors that influence and shape teachers' grades related to both achievement and non-achievement (e.g., effort and behaviour) (Guskey, 2011; Randall & Engelhard, 2009; Yesbeck, 2011). Teachers try hard to be fair to their students (Brookhart, 2004; Sun & Cheng, 2013) as they juggle their dual roles of being a judge of student learning and a coach to student learning

(Bishop, 1992). However, these roles may be in direct conflict within grading practices, and thus may jeopardize the validity of grade interpretation and use (Brookhart, 1991, 1993; Cross & Frary, 1999).

McMillan (2008) pointed out that even when teachers use the same grading scale and the same grading guidelines, there remains little consistency in teachers' grades (Brookhart, 1994; Liu, 2013). Based on interview data with secondary and elementary classroom teachers in Virginia, McMillan and Nash (2000) proposed a model of teachers' classroom grading decision-making including both internal and external influencing factors. The most salient internal factor was the teachers' philosophy of teaching and learning. The major external factors were identified as district (school board) grading policies, mandated statewide learning standards, high stakes tests, and parents' expectations. This model is also supported by studies conducted in the Canadian and the Chinese contexts. For example, Cheng and colleagues (Cheng, Rogers, & Hu, 2004; Cheng, Rogers, & Wang, 2008; Cheng & Wang, 2007; Sun & Cheng, 2013) investigated teachers' assessment and grading practices in Canada, Hong Kong, and China. These studies show that teachers' grading preferences are influenced by their values about assessment, their teaching experiences and training, their instructional contexts, and the dominance of large-scale testing.

McMillan (2008) argued that one of the most difficult issues in grading is how to deal with non-achievement factors such as effort, work habit, and motivation. He refers to these factors as academic enablers. Teachers consider these enabling factors in grading because they are traits that teachers cultivate and regard as important for student achievement. Zoeckler (2007) used a theoretical framework of truth, worthwhileness, trust, and intellectual and moral attentiveness to examine how US English language teachers attempted to assign fair grades while weighting both achievement and non-achievement factors. The results of this study indicated that grading was influenced by local grading systems, teachers' perceptions of student effort, and their concerns for moral development. Similarly, in a survey of 516 teachers in the USA, Randall and Engelhard (2010) found that under most circumstances, teachers were abided by the official grading policies of the participating school district, assigning grades based primarily on achievement. However, in borderline cases, teachers tended to value other characteristics such as ability, behaviour, and effort. In Ontario, Simon, Chitpin, and Yahya (2010) also found that pre-service teachers based grading decisions, in part, on non-achievement factors.

While a great majority of the studies above focus on grade construction, reliability, and validity within a given learning culture context, few studies have examined the comparability of grading policies and practices across contexts. Despite the fact that grading research has a long history, such research has been primarily driven by pedagogical concerns thus lacking explicit theoretical guidance. As currently regarded, grading is a social and contextual practice where "grades are acknowledged to have multiple meanings and multiple contexts, including social and legal contexts" (Brookhart, 2013, p. 265), understanding the similarities and differences in grading policies across educational systems has become increasingly important. This argument resides in a shift in our understanding of contemporary validity calling to a social and contextual understanding of the assessment context(s) where grading is constructed (Moss, 2003; Shepard, 2000). The current study begins to examine grading policies across two distinct learning cultures: Canada and China.

## **Education in Canada**

Formal education in Canada began during the 1850s with secondary education emerging in the

second half of the 19th century (Nagy, 2000). Initially, the emphasis was on educating the upper social class and focused on academic standards. Between 1846–1876, Dr. Egerton Ryerson, Superintendent of Education for Ontario, shaped education in Upper Canada towards public instruction drawing from Europe, the British Isles, and the United States (Ryerson, 1868). Currently, Canada’s education system is publically funded and accessible to all children, with compulsory elementary schooling commencing at either the Kindergarten (ages 4–5) or Grade 1 (age 6) levels. Education in Canada is decentralized with each of the ten provinces and three territories, with each responsible for the educational system within its jurisdiction (Klinger, DeLuca, & Miller, 2008; Volante & Jaafar, 2008). Each of the 13 systems has its own educational policies governing educational standards, curriculum, classroom practices, and assessment and grading; however, the nature of these policies is not prescriptive, leaving a range of possible interpretations for guiding assessment and grading in schools.

In contrast to other countries, there is no established national policy outlining a countrywide curriculum for students and teachers in Canada (e.g., National Curriculum in China; Common Core Standards in US). Instead, each provincial and territorial ministry of education assumes responsibility for overseeing compulsory elementary and secondary schooling through smaller agencies known as school boards or districts of education (Volante & Jaafar, 2008). As such, grading policies are established by ministries of education with interpretation and enactment occurring through local school boards and districts (Scott 1995).

In general, the education system in Canada is one that is predicated on standards and accountability (Klinger, DeLuca, & Miller, 2008). Standards-based grades are assigned according to a criterion-referenced framework where students’ performance is compared to established standards (Brookhart, 2013). Theoretically, standards-based grading is designed to be fairer for students because achievement is more accurately represented in relation to learning expectations or standards. Non-achievement factors such as learning skills (e.g., organization) are reported separately from grades (Tierney, Simon, & Charland, 2011). For accountability purposes, large-scale assessment programs are mainly used to indicate system effectiveness. While Canada does selectively participate in international standardized assessment programs (e.g., TIMSS, PISA), each province and territory has its own large-scale assessment system to monitor student achievement within an accountability framework of education (Klinger et al., 2008). In addition, this accountability has no intended negative consequences to schools or teachers, yet unintended negative consequences may happen and needs to be observed. Rather, it is expected that administrators and teachers will use the assessment results to inform and support their own ongoing school improvement efforts (Klinger & Rogers, 2011).

Concurrently with this emphasis on standards and accountability, more recently, educational assessment policies throughout Canada have also expanded to include explicit mandates towards *assessment for learning*, i.e., assessment practices designed to support learning as learning occurs. These mandates emphasize the use of formative assessments to provide ongoing feedback to support student learning as well as the use of summative assessments for standards-based grading purposes (e.g., OME, 2010) and large-scale assessments for accountability purposes.

## **Education in China**

The China known today as the People’s Republic of China is a country that represents “a long-standing civilization comprising successive states and cultures” (Cheng, 2008, p. 15). Its public education system started to develop around 4000 years ago in Xia Dynasty and matured around

Shang Dynasty (Fang, 1997). Because the education system is highly related to the political system, the education system of a politically centralized country is highly centralized accordingly (Xu, 2004). Today, China's education system remains centralized, though with aspects of provincial autonomy<sup>1</sup>, and is publically funded and controlled by its central government (Xu, 2004; Guan & Meng, 2007; Wang, 2012). It is a system accessible to all children, with nine years of compulsory education beginning from elementary (Grade 1–6) throughout to secondary (Grade 7–9) schooling (Wang, 2012). Education system in China is governed by its national Ministry of Education with provincial and municipal education ministries as level of administrations (Xu, 2004). The national Ministry of Education establishes the countrywide curriculum with grading policies embedded for all teachers and students in China. Provincial and municipal ministries and schools follow the overall guidance from the national ministry in their practice. China has a long history of public examinations. The testing and examination history in China can be traced back to the imperial period nearly 2000 years ago. Large-scale testing is evident at each key stage of the schooling (Grade 6 finishing elementary school, Grade 9 finishing junior high school and Grade 12 finishing senior high school).

## Methods

A policy review method (Arksey & O'Malley, 2005; Patton, 2015) was used to analyze contemporary grading policies from Canada and China. Data were obtained from two sources: (a) policy documents, and (b) relevant academic papers. Historical documents were only used to provide contextual information in relation to contemporary grading policies to identify and compare the influences of the psychometric tradition in Canada (Simon et al., 2010) and of the long-established Imperial Examinations in China (Cheng, 2010). In Canada, where education is under provincial jurisdiction, grading and assessment documents from all ten provinces and three territories were collected from Ministries of Education (MOE). In China, where education is governed at a national level, the country's MOE's documents were collected. In addition, curriculum documents and discussion papers<sup>2</sup> were also collected as these documents serve the function of disseminating central government documents to local provincial level of education.

In both contexts, we searched the MOE websites using the following key search terms (in either English or Chinese): grading policies and practices, assessment, summative assessment, and evaluation. Documents were included if they focused on grading policies in K–12 education. Our search resulted in a total of 65 documents (23 Canadian policy documents; and 13 Chinese policy documents, three curriculum documents with grading policy components, and 26 discussion papers). See Tables 1 and 2 in the appendix for complete list of documents.

Inductive content analysis (Arksey & O'Malley, 2005; Patton, 2015) was used for both the Canadian and Chinese contexts to identify themes related to the categories of (a) grade purposes, (b) methods for constructing grades, and (c) relationship between grading and formative assessment. *Grade purposes* included identifying the main purpose(s) of grading as explicitly

---

<sup>1</sup> China has 34 provincial-level administrative units: 23 provinces, 4 municipalities (Beijing, Tianjin, Shanghai, Chongqing), 5 autonomous regions (Guangxi, Inner Mongolia, Tibet, Ningxia, Xinjiang) and 2 special administrative regions (Hong Kong, Macau).

<sup>2</sup> Discussion papers are a unique feature of education documents in the Chinese context. These papers usually provide authors' perspectives, summaries, and reflections of grading and assessment practices, including MOE policies. In China these documents serve an intermediary role, offering interpretations of government policies on grading and creating a link between policies and practices for educators. Typically, authors of these discussion papers include provincial and local assessment officers, university professors specialised in educational assessment, and principals and teachers from primary and secondary schools (Sun & Cheng, 2013)

stated in policy documents (e.g., feedback, sorting students, and motivating learning). We also included the extent to which grading fairness, justice, or equity was discussed throughout the policies. Within the *methods for constructing grades* category, our analysis focused on the methods used to calculate grades, the evidence that contributes to grade, and the stakeholders involved in the grading process. Specifically, we examined how documents described the function of both achievement and non-achievement factors in grade construction, and the influence of large-scale assessments on grading policies and practices. In this category, we also considered how grades were reported (i.e., reporting periods, distribution/composition of grades, and scale systems). The final category, *grading in relation to formative assessment*, focused on (a) how policies articulated the distinction between formative and summative assessment, and (b) assessment terminology used to articulate the role of grades in relation to other assessment purposes. Two additional categories emerged from the Chinese data: Educational Philosophy (guiding values for education), and Coordination and Administration (the alignment of grading, and assessment training). Within all categories, relevant themes were identified within policy documents. All documents were coded by two researchers with an inter-rater reliability of 95%. In instances of code disagreement, raters discussed content until a consensus was reached.

## **Results**

Results are first reported with respect to Canadian and Chinese grading policies separately. We then compare the findings from both contexts to examine similarities and differences.

### ***Canadian Grading Policies***

The results of our systematic review of Canadian grading policies are presented in relation to three categories: (a) purposes of grades, (b) methods for constructing grades, and (c) grading in relation to formative assessment. Within each of these categories, specific themes were identified through the inductive analysis that presented general trends and described specific examples from provincial/territorial policy documents (see Table 3 in the appendix for summary of Canadian results). For the category of purposes of grades, the following five themes were identified: (a) support continuous learning, (b) planning, (c) feedback, (d) monitoring, and (e) reporting. For the category of method for constructing grades, the following two themes were identified: (a) curriculum-referenced, and (b) teacher-driven grading processes. For the category of grading in relation to formative assessment, the following two themes were identified: (a) explicitly not for grading purposes, and (b) assessment terminology.

It is evident that there are several consistent trends across provinces and territories in Canada about the purpose of grades. Namely, the primary functions of grades are to monitor and report on student achievement and to provide students with feedback about their learning. More specifically, nine provinces and three territories indicated that a primary purpose of grades was to monitor student achievement. The use of grades to support student learning was the least reported purpose of grades with only Alberta's policies explicitly stating this purpose.

Results suggested that the grading process is not necessarily always teacher-centred, as reported by only four provinces and one territory. Another consistent trend is that grades should primarily reflect students' learning of academic standards (i.e., criteria-referenced based on curriculum) rather than their learning skills or personal dispositions. In many provinces and territories, learning skills and work habits are represented separately through other reporting methods. Four provinces and three territories were explicit with the need for grades to be assigned in relation to student achievement of curriculum expectations. Two other provinces

implied this trend but it was not stated explicitly in the policy documents. Relatedly, the scales used for communicating grades are largely consistent across regions with secondary school grades typically reported using percentages and elementary grades typically using letter grades. The reporting of grades is typically standardized in each province with ministries of education providing guidelines and templates.

There still remains a large culture of large-scale assessments in Canada, which contributes to a significant amount of variability in how student grades are constructed in tested disciplines and grade levels. Large-scale provincial testing can contribute anywhere between 10% to 50% of a student's final grade depending on the course and province/territory. Despite consistencies in other areas of grading policies, significant variations in terminology was evident in grading policies across Canada related to the terms: formative, summative, assessment of learning (AoL), assessment for learning (AfL), and assessment as learning (AaL). While there is a consistent commitment to both formative and summative functions of assessment, with the latter reserved primarily for grading purposes, there are mixed trends across provincial/territorial systems about the use of formative assessment information in the construction of grades or the influence on teachers' judgments in grading decisions. Seven provinces and two territories explicitly stated that formative assessments should not contribute to the construction of grades. While AoL and AfL are used throughout documents to represent contemporary notions of formative and summative assessment, respectively, the integration of AaL in some documents suggests the prevailing importance of assessment used directly for metacognitive and self-regulation development. Typically considered as sub-concept of AfL, AaL focuses on supporting students' capacity for independent learning through processes of self-assessment, goal setting, and reflection.

### ***Chinese Grading Policies***

The findings from the Chinese government policies, curriculum documents, and discussion papers revealed how grading policies are directed from central government policy mandates through intermediary discussion papers to teachers' classroom practices within the Chinese context. In addition, curriculum documents serve as the bridge between the above policy documents and discussion papers. These curriculum documents incorporated policy guidelines with subject-specific goals and articulated specific assessment rules for classroom activities and teachers' grading practices. This bridging function is consistent with the expressed role of national curriculum documents as being "the foundation of classroom administration and assessment" (MOEPRC, 2001, p. 5).

As presented in Table 4, our analyses resulted in five overarching categories and related themes: (a) educational philosophy (two themes: comprehensive quality education vs. test-oriented education and learner-centered education); (b) grading purposes (two themes: promoting learning process and improving learning outcomes); (c) methods for constructing grades (four themes: assessing comprehensive quality, using multiple types of assessments, involving multiple stakeholders, e.g., teachers and parents as assessors, and reporting and using grades); (d) grading in relation to formative assessment (two themes: combining formative assessment with summative grading and focusing on formative assessment); and (e) coordination and administration (two themes: alignment with curriculum, large-scale tests, and admission system and assessment administration and training). The categories of educational philosophy, and coordination and administration, emerged as unique categories to the Chinese context, are closely related to methods for constructing grades and also impact on the other two categories

(i.e., grading purposes and grading in relation to formative assessment) within both Canadian and Chinese learning contexts. As the whole educational system in China is driven by comprehensive quality education, grading policies serve to promote comprehensive quality education. Hence the overall purpose of grading in China is to promote learning, in which comprehensive quality education, instead of academic performance measured by a final exam, is the primary focus.

**Table 4: Frequency counts of categories and themes per document source in Chinese context**

|   | Number of Documents       |                              |                            |
|---|---------------------------|------------------------------|----------------------------|
|   | Policy Documents (N = 13) | Curriculum Documents (N = 3) | Discussion Papers (N = 26) |
| <b>Educational Philosophy</b>   |                           |                              |                            |
| <i>Theme 1:</i> comprehensive quality education vs. test-oriented education   | 9                         | 1                            | 4                          |
| <i>Theme 2:</i> learner-centered education  | 4                         | 2                            | 3                          |
| <b>Grading Purposes</b>   |                           |                              |                            |
| <i>Theme 3:</i> promoting learning process (e.g., motivating students; improving teaching)  | 3                         | 2                            | 17                         |
| <i>Theme 4:</i> improving learning outcomes (e.g., improving students' development/performance; evaluating comprehensive quality defined in the curriculum) | 3                         | 3                            | 14                         |
| <b>Methods for Constructing Grades</b>  |                           |                              |                            |
| <i>Theme 5:</i> assessing comprehensive quality   | 8                         | 3                            | 18                         |
| <i>Theme 6:</i> using multiple types of assessments (e.g., not just tests)  | 8                         | 3                            | 16                         |
| <i>Theme 7:</i> involving multiple stakeholders as assessors  | 1                         | 3                            | 17                         |
| <i>Theme 8:</i> reporting and using grades  | 13                        | 3                            | 15                         |
| <b>Grading in Relation to Formative Assessment</b>  |                           |                              |                            |
| <i>Theme 9:</i> combining formative assessment with summative grading   | 2                         | 2                            | 13                         |
| <i>Theme 10:</i> focusing on formative assessment   | 6                         | 3                            | 5                          |
| <b>Coordination and Administration</b>  |                           |                              |                            |
| <i>Theme 11:</i> alignment with curriculum, large-scale tests, and admission system   | 7                         | 0                            | 1                          |
| <i>Theme 12:</i> assessment administration and training   | 8                         | 0                            | 0                          |

Across the categories observed in the Chinese results, three dominant trends emerged: (a) grading policies all endorse practices-oriented to comprehensive quality education; (b) grading

policies advocate practices that attempt to integrate Chinese assessment traditions with recent western educational assessment theories; and (c) areas of misalignment exist between government policy directives and curricular guidelines for grading in classrooms at the local provincial levels.

The most salient feature of these grading policies was that they are all oriented to the concept of comprehensive quality education (Fang, 2006; MOEPRC, 2002). According to the documents we analyzed, the primary purpose of grading was to make a holistic evaluation of both the learner and his/her learning, instead of only assessing the learner’s academic performance. The embedded value of the comprehensive quality education framework in China has two important implications for grading policies and practices. First, grading decisions are rendered on both the process of learning and the outcomes of that learning process. Second, teachers are encouraged to move away from grading solely on achievements using summative tests towards formative assessments using diverse assessment methods.

In examining the Chinese documents, there appears to be areas of misalignment between the intended grading policies and practitioners’ interpretations. Formative assessment was emphasized across the three types of documents as an important concept that promotes a balanced approach towards classroom assessment that values both the learning process and learning outcomes. However, the implementation of this balanced approach in the Chinese education system remains largely undefined in policy and curriculum documents, which brings challenges to practitioners related to consistency of implementation across schools and classrooms. In order to assess and assign grades on students’ comprehensive quality education, multiple types of evidence are to be collected from students themselves, peers, and their parents in addition to teachers, coupled with various types of assessment other than tests. Formative assessments are emphasized in relation to summative forms of student work to promote learning processes, although how formative assessments are used within grading decisions is unclear and ambiguous across Chinese policy documents and discussion papers. As a result, teachers in this context are likely using multiple types of assessments—used for varying purposes (i.e., formative and summative)—to formulate grades that reflect a holistic assessment of the student and his or her learning.

### Comparative Analysis

After analyzing grading policies in both Canada and China, we are able to draw comparisons between the two contexts. Table 5 presents major areas of difference in the purpose and construction of grades.

**Table 5: Comparison of trends in grading policies in Canada and China**

| Canada  | China  |
|---|--|
| <ul style="list-style-type: none"> <li>• Provincially governed (variation among provinces)</li> <li>• Specific focus on practices with clear directives</li> <li>• Driven by measurement theories</li> <li>• Focus on separate grading of achievement and non-achievement (i.e., focus on learning)</li> <li>• Some integration (depending upon province) with large-scale assessments</li> </ul> | <ul style="list-style-type: none"> <li>• Centrally governed (country-wide)</li> <li>• General policy statements subject to interpretations</li> <li>• Driven by educational philosophy</li> <li>• Focus on combined grading of achievement and non-achievement (i.e., focus on learners)</li> <li>• Shaped by large-scale assessments but not directly integrated</li> </ul> |

Firstly, grading policies are governed differently across Canada and China, with grading policies governed at the provincial level in Canada and the national level in China. As a result, grading policies in Canada are more specific to local provincial contexts compared to more general guidelines presented in Chinese national documents. In Canada, there exist variations among provincial policies, specifically related to consistency of terminology and influence of large-scale assessments on grade calculations. In contrast, national policies in China provide high level information and guidance but require intermediary interpretation at more local provincial and municipal levels prior to reaching individual schools.

A major recurring difference involves the inclusion and exclusion of non-achievement factors in grade decisions. Canadian policies consistently support the practice of grading solely on achievement evidence, providing separate opportunities for reporting non-achievement outcomes. Chinese policies, on the other hand, encourage teachers to include both achievement and non-achievement factors in grading student. Such policies reflect the country's overall commitment to comprehensive quality education than the measurement guidance on grading accuracy taken by Canada. This substantive difference is in part derived from the driving forces behind the policies; measurement theories in the Canadian context and comprehensive educational philosophy in the Chinese context.

Another substantive difference between the two contexts is the influence of large-scale assessments on classroom grade calculations. While both contexts have a significant history of large-scale assessment, in certain Canadian provinces the grades from these assessments can comprise a small part of students' final grades (from 10–50% depending on level and subject). Hence there is no uniform approach across Canada to the inclusion of large-scale assessment results into grading decisions or the influence of these assessments on students' high-stakes decisions. In China, on the other hand, large-scale assessments are not included in grade calculations and reported separately from teacher constructed grades. Results from large-scale assessments (all external to the classrooms at the national, provincial, or municipal levels) are used for high-stakes decisions such as admission and graduation but are not used as part of teacher constructed grades. The analysis of grading policies in China indicates that the MOE is planning to use high school grades as a supplemental measure in the future for university admission decisions in addition to the results from National Matriculation Tests, which is one of the largest national examinations in China.

## **Discussion**

Grading is a socio-cultural activity and the interpretation of grades must account for their social construction. Hence, we have examined how grading policies differ between two countries, Canada and China, to achieve a deeper understanding of the learning values of grading shared by and embedded *within* and *across* these two national contexts.

Putting the findings from Canada and China together, an emergent pattern can be seen as summarized in Table 4 showing four distinct yet intertwined features. First, grading policies are governed at the provincial level for Canada and the national level for China. This reflects the nature of the educational system in the two countries. Due to this nature, there is great variation in policies across provinces in Canada. In China, national policies trickle down to more local levels, with policies interpreted at provincial and municipal levels prior to reaching schools. In addition, a large volume of discussion papers also mediates these policies and supports their interpretation. This indicates a centralized policy system with decentralization interpretations and possible implementations.

Second, the policies we examined in Canada are specific in nature detailing the purposes and processes of grading whereas policies in China are general and broad, leaving room for interpretation and re-interpretation at various levels in the school system. Hence, our findings were consistent with Liu and Yan's (2016) conclusion that grading policies in China are broadly defined and implicitly embedded across several government and non-government documents. This conclusion may account for practitioners' varied interpretations of grading policies and approaches to grade construction, as observed in the discussion papers analyzed in the Chinese portion of the study. In addition, inconsistencies in grading policies and their related practices may also be a result of low interest in formative classroom assessment across China compared to the high-stake, high-use, large-scale assessments that shape so much of the educational landscape across the country.

Third, across the Canadian and Chinese contexts, grading policies stem from and reflect different philosophies towards assessment and education. Grading policies in China derive from an educational philosophy rooted on a *Comprehensive Quality Education*, which leads to grades based on holistic student learning including achievement and non-achievement evidence. In contrast, Canadian policies focus specifically and solely on achievement factors in grade construction, as guided by contemporary measurement theories. In China, the embedded values within the comprehensive quality education framework stipulate two important implications for grading policies and practices. First, in alignment with a comprehensive view of education, the focus of grading is on both the learner (their effort and homework) and the learning (its processes and outcomes)—rather than solely on academic achievement as is in Canada. Second, teachers are encouraged to expand their forms of assessment and evidence for making grading decisions from strictly test-based assessment to other measures of student performance. The inclusion of formative assessment commitments across Chinese policy documents as well as recognizing the value of diverse assessment approaches (e.g., performance-based assessments, alternative assessments, authentic assessments) is relatively recent. Both values suggest that there remain commitments to time-honoured Chinese examination traditions but an attempt to recognize contemporary educational assessment theories originating from the western world.

Lastly, in China, public examinations have played an important role in shaping teachers' grading practices even though such examinations are not part of teachers' grades, yet their classroom assessment and grading practices are largely aligned with those examination methods and formats (Cheng, 2010). In Canada, this trend has not been as consistent or as influential, with large-scale assessment results variably impacting students' grades in certain subjects and levels of education. The move toward assessing the comprehensive quality of the learner in the Chinese context, however, is a deliberate effort to avoid the negative consequences of large-scale and classroom test use on teaching and learning (Qi, 2005, 2007). Furthermore, the inclusion of non-achievement factors such as effort and behaviour in grade decisions can be traced back to the earliest Confucian education text "On Learning," and has since remained highly valued in the contemporary Chinese learning culture (Carless, 2011; Cheng & Curtis, 2010; Huang, 2011; Yin, 2008). However, the inclusion of both achievement and non-achievement factors in grading is contradictory to measurement specialists' approach to grading. The variable practices associated with the inclusion of non-achievement factors in grade decisions and large-scale assessment results raises important validity concerns, especially when comparing and using grades across educational systems for high-stakes decisions: What does a grade include? What does the grade signify?

These four distinct features represent what we have learned about grading policies in these two distinct learning cultures (i.e., the learning values embedded in these policies documents). Most importantly, these features demonstrate what we can learn from each other thus improving our understanding of what is in a grade. We hope this study can begin to contribute theoretically to our understanding of the learning values embedded in grading.

Grading is one of the most ubiquitous yet high-stakes practices in education. For hundreds of years, grades have been used as the key metric in decisions about student promotion, admissions, scholarship, and work placements (Atkinson & Geiser, 2009; Bowers, 2010; Pattison, Grodsky, & Muller, 2013; Thorsen & Cliffordson, 2012). Further, the function of grades has become even more important given the increased rates of student mobility across education systems throughout the world. Grades have become the primary currency for selection, placement, and admissions processes. As classroom grades are beginning to factor into such significant decisions for students (e.g., admission, selection, graduation), using a “test score + classroom grade” model, grading policies need to be clearly defined in order to guide teachers’ grading practices and ensure the validity of their assigned grades. Moreover, an organized assessment system needs to be developed (Brookhart, 2013) and aligned with other educational systems—such as curricula, large-scale tests, and admissions—so that a robust education macrosystem that meets 21st century educational needs and challenges can be established internationally.

## References

- Arksey, H., & O’Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. doi:10.1080/1364557032000119616
- Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher*, 38(9), 665–676.
- Bishop, J. H. (1992). Why U.S. students need incentives to learn. *Educational Leadership*, 49(6), 15–18.
- Bowers, A. J. (2010). Analyzing the longitudinal K–12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis. *Practical Assessment Research and Evaluation*, 15(7), 1–18.
- Brookhart, S. M. (1991). Letter: Grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1), 35–36.
- Brookhart, S. M. (1993). Teachers’ grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), 123–142.
- Brookhart, S. M. (1994). Teachers’ grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279–301.
- Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, 106(3), 429–458.
- Brookhart, S. M. (2013). Grading. In J. H. McMillan (Ed.), *Research on classroom assessment* (pp. 257–272). Los Angeles, CA: Sage.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T. & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803–848.
- Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. New York, NY: Routledge.
- Cheng, L. (2008). The key to success: English language testing in China. *Language Testing*, 25(1), 15–37. doi:10.1177/0265532207083743
- Cheng, L. (2010). The history of examinations: Why, how, what and whom to select? In L. Cheng & A. Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 13–26). New York, NY: Routledge.
- Cheng, L., & Curtis, A. (Eds.). (2010). *English language assessment and the Chinese learner*. New York, NY: Routledge.
- Cheng, L., & Wang, X. (2007). Grading, feedback, and reporting in ESL/EFL classrooms. *Language Assessment Quarterly*, 4(1), 85–107.

- Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods and procedures. *Language Testing*, 21(3), 360–389.
- Cheng, L., Rogers, T., & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment & Evaluation in Higher Education*, 33(1), 9–32.
- Cross, L., & Frary, R. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53–72.
- Dobbin, J. E., & Smith, A. Z. (1960). Marks and marking systems. In C. W. Harris (Ed.), *Encyclopedia of Educational Research* (3rd ed.) (pp. 783–791). New York, NY: Macmillan.
- Fang, W. (1997). 孔子教育的素质教育因素 [Comprehensive quality education factors in Confucius education]. *池州师专学报 [Journal of Chizhou Teachers College]*, 4, 89–95.
- Fang, X. (2006). 素质教育的提出与发展 [The proposal and development of comprehensive quality education]. *教育研究史 [History of Educational Research]*, (4), 397–401.
- Finkelstein, I. E. (1913). *The marking system in theory and practice. Educational psychology monographs No. 10.* Baltimore, MD: Warwick & York.
- Guan, Q., & Meng, W. (2007). China's new national curriculum reform: Innovation, challenges and strategies. *Frontiers of Education in China*, 2(4), 579–604. doi:10.1007/s11516-007-0043-6
- Guskey, T. (2011). Five obstacles to grading reform. *Educational Leadership*, 69(3), 17–21.
- Huang, G. (2011). 大学公共课程平时成绩评定的新思考 [New thinking on usual performance assessment of college common required courses]. *高教论坛 [Higher Education Forum]*, 4, 80–82.
- Klinger, D. A., & Rogers, W. T. (2011). Teachers' perceptions of large-scale assessment programs within low-stakes accountability frameworks. *International Journal of Testing*, 11(2), 122–143. doi:10.1080/15305058.2011.552748
- Klinger, D., DeLuca, C., & Miller, T. (2008). The evolving culture of large-scale assessments in Canadian education. *Canadian Journal of Educational Administration and Policy*, 76, 1–34.
- Liu, X. (2013). *Investigating factors influencing grading decisions among teachers of Chinese to speakers of other languages.* Unpublished M.Ed. thesis. Queen's University, Kingston, ON, Canada.
- Liu, X., & Yan, W. (2016). Grading in East Asia: An examination of government policies. *Graduate Student Symposium Selected Papers*, 10, 73-96.
- McMillan, J. H. (2008). *Assessment essentials for standards-based education* (2nd ed.). Thousand Oaks, CA: Sage.
- McMillan, J. H., & Nash, S. (2000). *Teachers' classroom assessment and grading decision making.* Paper presented at the Annual Meeting of the National Council of Measurement in Education, New Orleans, LA.
- (MOEPRC) Ministry of Education of the People's Republic of China. (2001). *基础教育课程改革纲要(试行) [Compendium of basic education curriculum reform (trial implementation)]*. Retrieved from [http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe\\_309/200412/4672.html](http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_309/200412/4672.html)
- (MOEPRC) Ministry of Education of the People's Republic of China. (2002). *教育部关于积极推进中小学评价与考试制度改革的通知 [Notice on promoting the reform of the assessment and examination systems of elementary and secondary schools]*. Retrieved from [http://www.moe.gov.cn/srcsite/A26/s7054/200212/t20021227\\_166074.html](http://www.moe.gov.cn/srcsite/A26/s7054/200212/t20021227_166074.html)
- Moss, P.A. (2003). Reconceptualizing validity for classroom assessment. *Educational Measurement: Issues and Practice*, 22(4), 13–25.
- Nagy, P. (2000). The three roles of assessment: Gatekeeping, accountability, and instructional diagnosis. *Canadian Journal of Education*, 25, 262–79.
- Pattison, E., Grodsky, E., & Muller, C. (2013). Is the sky falling? Grade inflation and the signaling power of grades. *Educational Researcher*, 42(5), 259–265.
- Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods.* Thousand Oaks, California: Sage Publications.
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing*, 22, 142–173.
- Qi, L. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education: Principles, Policy & Practice*, 14, 51–74.
- Randall, J., & Engelhard, G. (2009). Differences between teachers' grading practices in elementary and middle schools. *Journal of Educational Research*, 102(3), 175–85.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372–1380.

- Ryerson, E. (1868). *A special report on the systems and state of popular education on the continent of Europe, in the British Isles, and the United States of America, with practical suggestions for the improvement of public instruction in the province of Ontario*. Toronto, ON: Leader Steam Press.
- Scott, E. L. (1995). 'Mokita'. The truth that everyone knows but nobody talks about: bias in grading. *Teaching English in the Two Year College*, 22(3), 211–216.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Simon, M., Chitpin, S., & Yahya, R. (2010). Pre-service teachers' thinking about student assessment issue. *The International Journal of Education*, 2(2), 1–22.
- Statistics Canada. (2016). *International students in Canadian universities, 2004/2005 to 2013/2014*. Retrieved: <http://www.statcan.gc.ca/pub/81-599-x/81-599-x2016011-eng.pdf>
- Sun, Y., & Cheng, L. (2013). Teachers' grading practices: Meanings and values assigned. *Assessment in Education*. doi:10.1080/0969594X.2013.768207
- Teaf, H. M. (1964). Grades: Their dominion is challenged. *The Journal of Higher Education*, 35(2), 87–88.
- Thorsen, C., & Cliffordson, C. (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades. *Educational Research and Evaluation*, 18(2), 153–172.
- Tierney, R. D., Simon, M., & Charland, J. (2011). Being fair: Teachers' interpretations of principles for standards-based grading. *The Educational Forum*, 75(3), 210–227. doi:10.1080/00131725.2011.577669
- Volante, L., & Jaafar, B. (2008). Educational assessment in Canada. *Assessment in Education: Principles, Policy & Practice*, 15(2), 201–210.
- Wang, J. (2012). Curriculum reform in mainland china, 1978–2008: Change, maintenance, and conflicts. *Chinese Education & Society*, 45(1), 59–68. doi:10.2753/CED1061-1932450107
- Xu, W. (2004). 社会形态, 政治权利和教育体制: 当代中国体育体制改革的逻辑 [Society modality, politics power and education system: The logic of modern Chinese education system reform]. *复旦教育论坛 [Fudan Education Forum]*, 2(4), 21–25.
- Yesbeck, D. M. (2011). *Grading practices: Teachers' considerations of academic and non-academic factors* (Unpublished doctoral dissertation). Virginia Commonwealth University, Richmond, Virginia.
- Yin, L. (2008). 学习态度与学习成绩的相关研究 [A Research on the Relationship Between Study Attitude and Study Achievement]. *心理科学 [Psychological Science]*, 31(6), 1471–1473.
- Zoeckler, L. (2007). Moral aspects of grading: A study of high school English teachers' perceptions. *American Secondary Education*, 35(2), 83–102.

**Liying Cheng**, PhD is Professor and Director of the Assessment and Evaluation Group (AEG) at the Faculty of Education, Queen's University. Her primary research interests are the impact of large-scale testing on instruction, the relationships between assessment and instruction, and the academic and professional acculturation of international and new immigrant students, workers, and professionals to Canada.

**Christopher DeLuca**, PhD is an Associate Professor and Graduate Faculty member in Classroom Assessment at the Faculty of Education, Queen's University. He leads the Classroom Assessment Research Team and is a member of the Queen's Assessment and Evaluation Group. Dr. DeLuca's research examines the complex intersection of curriculum, pedagogy, and assessment as operating within the current context of school accountability and standards-based education. In particular, Dr. DeLuca's research centers on how pre-service and in-service teachers learn to engage the complexities of assessing student learning in relation to the evolving accountability culture in today's classrooms.

**Heather Braund** is a doctoral student at the Faculty of Education in the Assessment and Evaluation Group. Her doctoral research is focused on bridging elementary teachers' classroom assessment practices with metacognitive development and self-regulated learning behaviours in elementary students.

**Wei Yan** is a PhD student in the Assessment and Evaluation Group at the Faculty of Education, Queen's University. He is an experienced language instructor and curriculum developer. His research focuses on grading practices and validity.

## Appendix 1

**Table 1: List of Policy Documents Reviewed and Province/Territory Abbreviations**

| <b>Province/<br/>Territory &amp;<br/>Abbreviation</b> | <b>Document</b>   |
|---|---|
| <b>Alberta (AB)</b>                                   | Alberta Education. (2015). <i>Guide to Education. ECS to Grade 12</i> . Retrieved from <a href="https://education.alberta.ca/guide-to-education/">https://education.alberta.ca/guide-to-education/</a>  |
| <b>British<br/>Columbia (BC)</b>                      | British Columbia Ministry of Education. (2009). <i>Reporting student progress: Policy and practice</i> . Retrieved from <a href="https://www.bced.gov.bc.ca/classroom_assessment/09_report_student_prog.pdf">https://www.bced.gov.bc.ca/classroom_assessment/09_report_student_prog.pdf</a>     |
| <b>Manitoba (MB)</b>                                  | Department of Manitoba Education, Citizenship and Youth. (2008). <i>Communicating Student Learning</i> . Retrieved from <a href="http://www.edu.gov.mb.ca/k12/assess/docs/csl/csl_doc.pdf">http://www.edu.gov.mb.ca/k12/assess/docs/csl/csl_doc.pdf</a>   |
|   | Manitoba Education and Advanced Learning. (2015). <i>Provincial Assessment and Policy. Kindergarten to Grade 12</i> . Retrieved from <a href="http://www.edu.gov.mb.ca/k12/assess/docs/policy_k12/">http://www.edu.gov.mb.ca/k12/assess/docs/policy_k12/</a>                                    |
|   | Department of Manitoba Education, Citizenship and Youth. (2006). <i>Rethinking Classroom Assessment with Purpose in Mind</i> . Retrieved from <a href="http://www.edu.gov.mb.ca/k12/assess/wncp/">http://www.edu.gov.mb.ca/k12/assess/wncp/</a>   |
| <b>New Brunswick<br/>(NB)</b>                         | New Brunswick Department of Education and Early Childhood Development. (2013). <i>Framework for provincial assessments</i> . Retrieved from <a href="http://www.gnb.ca/0000/results/pdf/AssessmentFrameworkDocument.pdf">http://www.gnb.ca/0000/results/pdf/AssessmentFrameworkDocument.pdf</a> |
|   | New Brunswick Department of Education and Early Childhood Development (2013). Assessment policy statement 331. Retrieved from <a href="http://web1.nbed.nb.ca/sites/ASD-E/policies/District%20Policies/331n.pdf">http://web1.nbed.nb.ca/sites/ASD-E/policies/District%20Policies/331n.pdf</a>   |
| <b>Newfoundland<br/>and Labrador<br/>(NFL)</b>        | Newfoundland and Labrador Department of Education. (2013). <i>K –12 student assessment, evaluation and grading policy</i> . Retrieved from <a href="https://www.nlesd.ca/about/doc/policies/archive/labrador/112.pdf">https://www.nlesd.ca/about/doc/policies/archive/labrador/112.pdf</a>      |
| <b>North West<br/>Territories<br/>(NWT)</b>           | Department of Education, Culture & Employment. (2011–2012). <i>Educating All Our Children: Procedures, Roles and Responsibilities for Student Assessment, Evaluation and Reporting</i> .  |
|   | Department of Education, Culture & Employment. (2010). <i>Educating All Our Children: Departmental Directive on Student Assessment, Evaluation and Reporting</i> .  |

---

|                                   |  |
|-----------------------------------|--|
| <b>Nova Scotia (NS)</b>           | <p>Nova Scotia Department of Education (NSDE). (2002). <i>Reporting Policy Framework</i>.</p> <p>Annapolis Valley Regional School Board (AVRSB). (2005). <i>Student assessment, evaluation, and reporting practices and procedures</i>. Retrieved from <a href="https://www.avrsb.ca/sites/default/files/402.6%20Appendix%20A%20Evaluation%20of%20Student%20Progress.pdf">https://www.avrsb.ca/sites/default/files/402.6%20Appendix%20A%20Evaluation%20of%20Student%20Progress.pdf</a></p> <p>South Shore Regional School Board (SSRSB). (2015). <i>Student assessment, evaluation, and communication of student learning. Governance policy 213</i>. Retrieved from <a href="http://www.ssrbs.ca/wp-content/uploads/2015/11/213-Student-Assessment-Evaluation-and-Communication-of-Student-Learning-Appr-09-23-15.pdf">http://www.ssrbs.ca/wp-content/uploads/2015/11/213-Student-Assessment-Evaluation-and-Communication-of-Student-Learning-Appr-09-23-15.pdf</a></p> |
| <b>Nunavut (NU)</b>               | <p>Nunavut Department of Education. (2008). <i>Foundation for Dynamic Assessment as Learning in Nunavut Schools</i>, 1-60. Retrieved from <a href="http://www.gov.nu.ca/education/information/assessment-and-evaluation">http://www.gov.nu.ca/education/information/assessment-and-evaluation</a></p> <p>Nunavut Department of Education, Curriculum and School Services Division. (2007). <i>Education Framework Inuit Qaujimagajatuqangit for Nunavut Curriculum</i>, 1-67.</p>  |
| <b>Ontario (ON)</b>               | <p>Ontario Ministry of Education. (2010). <i>Growing Success: Assessment, Evaluation, and Reporting in Ontario Schools</i>. Toronto, ON: Queen's Printer for Ontario.</p>  |
| <b>Prince Edward Island (PEI)</b> | <p>English Language School Board. (2015). <i>Assessing, evaluating, monitoring and reporting student achievement. Policy 407.1</i>. Retrieved from <a href="http://www.gov.pe.ca/edu/elsb/files/2015/10/407.1_Assessing_Evaluating_Monitoring_and_Reporting_Student_Achievement.pdf">http://www.gov.pe.ca/edu/elsb/files/2015/10/407.1_Assessing_Evaluating_Monitoring_and_Reporting_Student_Achievement.pdf</a></p> <p>English Language School Board. (2015). <i>Assessing, evaluating, monitoring and reporting student achievement. Policy 407</i>. Retrieved from <a href="http://www.gov.pe.ca/edu/elsb/files/2015/10/407_Assessing_Evaluating_Monitoring_and_Reporting_Student_Achievement.pdf">http://www.gov.pe.ca/edu/elsb/files/2015/10/407_Assessing_Evaluating_Monitoring_and_Reporting_Student_Achievement.pdf</a></p>  |
| <b>Quebec (QC)</b>                | <p>MEQ. (2003). <i>Policy on the evaluation of learning</i>. Retrieved from <a href="http://www.education.gouv.qc.ca/references/publications/resultats-de-la-recherche/detail/article/politique-devaluation-des-apprentissages/pubLang/1/">http://www.education.gouv.qc.ca/references/publications/resultats-de-la-recherche/detail/article/politique-devaluation-des-apprentissages/pubLang/1/</a></p> <p>MEQ. (2011). <i>Framework for the evaluation of learning: English as a second language</i>. Retrieved from <a href="http://www1.education.gouv.qc.ca/sections/programmeFormation/secondaire2/index_en.asp">http://www1.education.gouv.qc.ca/sections/programmeFormation/secondaire2/index_en.asp</a></p>  |

---

---

MEQ. (2015). *Administrative Guide–2015 Edition. Certification of Studies and Management of Ministerial Examinations*. Retrieved from [www.education.gouv.qc.ca/fileadmin/site\\_web/.../Guide-sanction-2015\\_ang.pdf](http://www.education.gouv.qc.ca/fileadmin/site_web/.../Guide-sanction-2015_ang.pdf)

**Saskatchewan (SK)**

Saskatchewan Education. (1991). *Student Evaluation: A Teacher Handbook*. Retrieved from <http://www.publications.gov.sk.ca/details.cfm?p=10168>

**Yukon (YT)**

Yukon Education. (2011). *Reporting on Student Progress in Yukon Schools*. Retrieved from [http://www.yesnet.yk.ca/pdf/11-12/r\\_student\\_progress\\_dec\\_11.pdf](http://www.yesnet.yk.ca/pdf/11-12/r_student_progress_dec_11.pdf)

---

## Appendix 2

Table 2: List of documents reviewed in Chinese context

| Document Types               | Document   |
|------------------------------|--|
| Policy Documents<br>(n = 13) | Ministry of Education of the People's Republic of China. (2001a). <i>国务院关于基础教育改革与发展的决定</i> [The State Council decisions on the basic education reform and development]. Retrieved from <a href="http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_719/200409/3843.html">http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_719/200409/3843.html</a>   |
|                              | Ministry of Education of the People's Republic of China. (2001b). <i>基础教育课程改革纲要(试行)</i> [Compendium of basic education curriculum reform (trial implementation)]. Retrieved from <a href="http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_309/200412/4672.html">http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_309/200412/4672.html</a>   |
|                              | Ministry of Education of the People's Republic of China. (2002). <i>教育部关于积极推进中小学评价与考试制度改革的通知</i> [Notice on promoting the reform of the assessment and examination systems of elementary and secondary schools]. Retrieved from <a href="http://www.moe.gov.cn/srcsite/A26/s7054/200212/t20021227_166074.html">http://www.moe.gov.cn/srcsite/A26/s7054/200212/t20021227_166074.html</a>  |
|                              | Ministry of Education of the People's Republic of China. (2005). <i>教育部关于基础教育课程改革试验区初中毕业考试与普通高中招生制度改革的指导意见</i> [Guidelines on the reform of junior high school graduation examination and high school enrollment system in the basic education curriculum reform pilot area]. Retrieved from <a href="http://www.moe.gov.cn/srcsite/A26/s8001/200501/t20050112_167346.html">http://www.moe.gov.cn/srcsite/A26/s8001/200501/t20050112_167346.html</a> |
|                              | Ministry of Education of the People's Republic of China. (2006). <i>教育部关于贯彻《义务教育法》进一步规范义务教育办学行为的若干意见</i> [Guidelines on the implementation of the compulsory education law to further regulate the compulsory education schooling]. Retrieved from <a href="http://www.moe.gov.cn/srcsite/A06/s3321/200608/t20060824_81811.html">http://www.moe.gov.cn/srcsite/A06/s3321/200608/t20060824_81811.html</a>   |
|                              | Ministry of Education of the People's Republic of China. (2010). <i>关于在小学减轻学生过重负担的紧急通知</i> [Notice of reducing elementary school students' study load]. Retrieved from <a href="http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s3321/201001/81821.html">http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s3321/201001/81821.html</a>   |
|                              | Ministry of Education of the People's Republic of China. (2011a). <i>教育部关于印</i>  |

发义务教育语文等学科课程标准(2011年版)的通知 [Notice on issuing the compulsory education subject curriculum standards (2011 Edition) such as the Chinese Literacy etc.]. Retrieved from [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk\\_167340.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk_167340.html)

Ministry of Education of the People's Republic of China. (2011b). *小学教师专业标准(试行)* [Elementary school teachers' professional standards (trial implementation)]. Retrieved from <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s6127/201112/127836.html>

Ministry of Education of the People's Republic of China. (2011c). *中学教师专业标准(试行)* [Secondary school teachers' professional standards (trial implementation)]. Retrieved from <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s6127/201112/127830.html>

Ministry of Education of the People's Republic of China. (2013). *教育部关于推进中小学教育质量综合评价改革的意见* [Guidelines on promoting the comprehensive assessment of educational quality in elementary and secondary schools]. Retrieved from [http://www.moe.gov.cn/srcsite/A26/s7054/201306/t20130608\\_153185.html](http://www.moe.gov.cn/srcsite/A26/s7054/201306/t20130608_153185.html)

Ministry of Education of the People's Republic of China. (2014a). *国务院关于深化考试招生制度改革的实施意见* [The State Council guidelines on deepening the reform of examination and enrollment systems]. Retrieved from [http://www.moe.gov.cn/jyb\\_xxgk/moe\\_1777/moe\\_1778/201409/t20140904\\_174543.html](http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/201409/t20140904_174543.html)

Ministry of Education of the People's Republic of China. (2014b). *教育部办公厅关于开展义务教育阶段学校“减负万里行·第2季”活动的通知* [Notice of the “study load reduction Stage 2” campaign in the compulsory education schools]. Retrieved from [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s3321/201404/xxgk\\_166578.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s3321/201404/xxgk_166578.html)

Ministry of Education of the People's Republic of China. (2014c). *教育部关于加强和改进普通高中学生综合素质评价的意见* [Guidelines on strengthening and improving the comprehensive quality assessment of high school students]. Retrieved from <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s4559/201412/181667.html>

---

**Curriculum Documents**  
(n = 3)

- Ministry of Education of the People's Republic of China. (2011d). 义务教育语文课程标准 [Compulsory education Chinese Literacy curriculum standards]. Retrieved from [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk\\_167340.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk_167340.html)
- Ministry of Education of the People's Republic of China. (2011e). 义务教育数学课程标准 [Compulsory education Mathematics curriculum standards]. Retrieved from [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk\\_167340.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk_167340.html)
- Ministry of Education of the People's Republic of China. (2011f). 义务教育英语课程标准 [Compulsory education English curriculum standards]. Retrieved from [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk\\_167340.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk_167340.html)

---

**Discussion Papers**  
(n = 26)

- Bei, X. (2009). 中学语文教学评价改革的实践与思考 [Practice and thinking on the assessment reform of Chinese Literacy classes in junior high and high schools]. *江苏教育学院学报(社会科学版)* [Journal of Jiangsu Institute of Education (Social Sciences)], 25(5), 118-121.
- Bian, Y., & Shan, H. (2006). 基于标准的教育:利用评分规则促进学生学习 [Standards-based education: Promoting students' learning by applying the marking rule]. *教育理论与实践* [Theory and Practice of Education], 26(7), 30-33.
- Chen, S. (2010). 小学语文、数学、英语成绩等级制存在的问题及解决策略 [Problems and solutions in elementary school Chinese Literacy, Mathematics, English grading system]. *江苏教育(小学教学版)* [Jiangsu Education (Elementary Education Edition)], (3), 49-50.
- Chen, W. (2009). 浅谈学生评估在中学教育中的应用 [The application of student assessment in junior high and high school education]. *中国科教创新导刊* [China Education Innovation Herald], (6), 96.
- Gong, M., & Gao L. (2006). 新课程背景下高中语文必修课程评价方案构想 [The evaluation scheme of the required high school Chinese Literacy classes under the new curriculum]. *淮南师范学院学报* [Journal of Huainan Teachers College], 38(8), 95-98.
- Li, H. (2011). 谈小学数学学业成绩评价 [Elementary school mathematics

- achievement assessment]. *学科建设 [Discipline Construction]*, (12), 178.
- Li, Y., & Zhang, X. (2000). 教育评价发展的历史轨迹及其规律 [The history of education assessment development and its law]. *江苏高教 [Jiangsu Higher Education]*, (3), 62-65.
- Liang, Y., & Lv, L. (2010). 学生学业评价与考试改革研究及探索 [Research and exploration on the reform of students' achievement assessment and examination]. *教育实践与研究 [Educational Practice and Research]*, (10), 10-12.
- Ma, N., Li, W., & Xue, Y. (2009). 基础教育学业成绩元评价的探析 [A study on the meta-evaluation of academic achievements in basic education]. *中国电力教育 [China Electric Power Education]*, 135, 29-30.
- Ma, Y. (2014). 高中音乐教学中学生成绩评定标准浅析 [An analysis on the grading standards of students' performance in high school music teaching]. *学周刊(教育科学) [Learning Weekly (Educational Science)]*, (6), 213.
- Quan, S., & Fang, F. (2007). 学业评估改革与创新型人才培养 [The reform of achievement assessment and the cultivation of innovative talents]. *科教文汇 [The Science Education Article Collects]*, (12), 3 & 9.
- Sun, N. (2013). 等第制与百分制孰优孰劣 [Letter grading vs percentage grading: strengths and weaknesses]. *上海教育 [Shanghai Education]*, (15), 73.
- Sun, S., & Zhang, Q. (2004). 关于基础教育中对学生评价问题的探讨 [A tentative study of the assessment of school students]. *玉溪师范学院学报 [Journal of Yuxi Teachers College]*, 19(6), 62-64.
- Sun, W. (2011). 新课程初中物理教学评价体系的建构探索 [An exploration on the construction of the evaluation system of junior high school Physics classes under the new curriculum]. *湖南中学物理 [Hunan Middle School Physics]*, (10), 47-51.
- Wang, C. (2008). 基于新课程理念下化学教学评价体系的建构 [The construction of the evaluation system of Chemistry teaching based on the new curriculum concept]. *中学化学教学参考 [Teaching Reference of Middle School Chemistry]*, (9), 7-9.
- Wang, D. (2001). 深化“等级+评语”改革构建学生评价体系 [Deepening the reform of "letter grading + comment" to construct the students' evaluation system]. [ *辽宁教育 Liaoning Education*, (9), 17-18.

- Wang, L. (2008). 学业评估的多重参照性特征及其教育意蕴 [Multiple references of academic evaluation and its educational implications]. *全球教育展望* [Global Education], 37(8), 20–24 & 16.
- Wu, X. (2001). 论教育评定: 《素质教育施教策略研究》研究报告之一 [On the education assessment: A research report of comprehensive quality teaching strategy research]. *中学教育* [Middle School Education], (3), 5–7 & 41.
- Xia, Z. (2007). 分数功能的异化与回归 [On the deviation and correction of the function of testing scores]. *课程教材教法* [Curriculum, Teaching Material and Method], 27(6), 18–23.
- Xue, Y. (2015). 英语评价不妨用“等级制度” [Why not using “letter grading” for English class evaluation]. *小学教学参考* [Reference for Elementary Education], (2), 65.
- Yang, C. (2007). 基础性与预备性: 基础教育学业评估的内在张力 [Basic and preparatory: the intrinsic tension of the academic evaluation in basic education]. *现代中小学教育* [Modern Primary and Secondary Education], 158(4), 66–68.
- Zhang, J. (2007). 创新时代的教育目的与我国基础教育学业评估 [The educational purpose of the era of innovation and academic evaluation of basic education in China]. *天津市教科院学报* [Journal of Tianjin Academy of Educational Science], (2), 18–19 & 55.
- Zhang, S. (2011). 提高学生多元等级评价效度的六“性” [Six “ways” to improve the validity of students’ multiple scale evaluation]. *考试周* [Testing Weekly], (55), 219–220.
- Zhang, X-J. (2008). 浅谈高中英语课程教学评价的设计原则 [The designing principles of the evaluation of English classes in high school]. *内江科技* [Neijiang Science], (8), 177.
- Zhang, X-L. (2011). 学生学业成绩评价改革研究 [Reform of secondary school academic evaluation of students]. *太原大学教育学院学报* [Journal of Education Institute of Taiyuan University], 29(3), 21–24.
- Zhu, X. (2006). 高中地理学业成绩评价方案构想 [The conception of the performance assessment programs on Geography in senior high schools]. *课程教材教法* [Curriculum, Teaching Material and Method], 26(5), 53–58.
-

### Appendix 3

**Table 3: Analysis of Canadian grading policies (provinces and territories)**

| Province                       | Purpose of Grading             |          |          |            |           | Method for Constructing Grades |                                | Grading in Relation to Formative Assessment |                         |
|--------------------------------|--------------------------------|----------|----------|------------|-----------|--------------------------------|--------------------------------|---|-------------------------|
|                                | To Support Continuous Learning | Planning | Feedback | Monitoring | Reporting | Curriculum-referenced          | Teacher driven grading process | Explicitly not for grading purposes         | Assessment terminology* |
| Alberta (AB)                   | X                              |          |          |            |           | X                              |                                |   |                         |
| British Columbia (BC)          |                                | X        | X        | X          |           | X                              | X                              | X   |                         |
| Manitoba (MB)                  |                                |          | X        | X          |           | X                              | X                              | X   | AfL/AaL/AoL             |
| New Brunswick (NB)             |                                |          | X        | X          |           |                                |                                |   | AfL/AoL                 |
| Newfoundland and Labrador (NL) |                                |          | X        | X          |           |                                |                                | X   | AfL/AaL/AoL             |
| Northwest Territories (NWT)    |                                | X        |          | X          | X         | X                              |                                | X   | AfL/AoL                 |
| Nova Scotia (NS)               |                                |          |          | X          | X         | X                              | X                              |   | AfL/AoL                 |
| Nunavut (NU)                   |                                | X        | X        | X          |           | X                              |                                | X   | AfL/AaL/AoL             |
| Ontario (ON)                   |                                | X        | X        | X          |           |                                | X                              | X   | AfL/AaL/AoL             |
| Prince Edward Island (PE)      |                                | X        | X        | X          | X         |                                |                                | X   | AfL/AoL                 |
| Québec (QC)                    |                                | X        | X        | X          |           |                                | X                              | X   |                         |
| Saskatchewan (SK)              |                                |          |          | X          |           |                                |                                | X   |                         |
| Yukon (YT)                     |                                |          | X        | X          |           | X                              | X                              |   |                         |

\*Assessment for Learning=AfL, Assessment of Learning = AoL, Assessment as Learning = AaL