



# Assessment for learning in the Hong Kong assessment reform: A case of policy borrowing

Zi Yan<sup>a,\*</sup>, Gavin T.L. Brown<sup>b,c</sup>

<sup>a</sup> The Education University of Hong Kong, Hong Kong

<sup>b</sup> University of Auckland, Auckland, New Zealand

<sup>c</sup> Umeå Universitet, Umeå, Sweden

## ARTICLE INFO

### Keywords:

Assessment for learning  
Assessment reform  
Policy borrowing  
Public examination

## ABSTRACT

This paper describes Hong Kong's borrowing, primarily from the UK, of the assessment for learning policy, in the context of prolonged use of formal summative public examinations. The narrative review and analysis are guided by a social positivist critique of the assessment for learning policy in the Hong Kong context. This paper concludes that Hong Kong's attempts to implement assessment for learning are ambitious but somewhat futile because of persistent use of public examinations for important decision-making. Change in Hong Kong may only happen through a gradual implementation of assessment for learning practices that takes into account cultural, societal, and historic norms. Critical issues are identified and recommendations are proposed for further implementation of assessment for learning in Hong Kong.

## 1. Introduction

Assessment has been conventionally used for accountability and selection purposes in East Asia for centuries (China Civilisation Centre, 2007). Although it was not a new idea of using assessment to support student learning (e.g., Bloom, 1969; Scriven, 1967), the wide recognition and acceptance of this idea happened late in the 20th century. Since the influential publication of Black and Wiliam (1998a, 1998b) on formative assessment and the work of the Assessment Reform Group (ARG) on assessment for learning (Assessment Reform Group, 1999, 2002), assessment reform with an emphasis on assessment for learning has become an international trend (Berry, 2011b). The challenge in adopting such a policy within a context that has prolonged and widespread acceptance of public examinations is that it requires addressing the concerns and power of that extant policy framework (Kennedy, Chan, & Fok, 2011). Hence, this paper describes the status quo of Hong Kong's borrowing of the assessment for learning policy in the context of prolonged use of formal summative public examinations, and identifies tensions in the policy implementation process that have still to be addressed.

### 1.1. Assessment of learning

Assessment of learning focuses on collecting learning evidence typically through large-scale written examinations. In this fashion, assessment happens at the end of a particular learning period and is regarded as separate from learning. The premise of assessment of learning is to hold students individually accountable for their learning, by assigning evaluative judgements (i.e., grades or scores), checking off student performance against criteria, placing students into classes or groups based on performance, and reporting grades to parents, future employers, and educators. It is also seen in the various qualifications examinations in which students participate for graduation or entry selection to higher levels of educational or employment opportunity (Guthrie, 2002).

Characteristics commonly associated with assessment of learning are formal, standardised methods of data collection from students, rigorous attention to accuracy in scoring of student responses (e.g., machine scoring of multiple-choice questions, strictly moderated human marking of written examination responses). The score reporting is usually a blend of criterion- and norm-referencing. For example, grades from the Hong Kong Diploma of Secondary Education (HKDSE) examinations are criterion-referenced aligned with five levels of performance (1–5), but performance in the top level 5 is supplemented with a norm-referenced

\* Corresponding author at: Department of Curriculum and Instruction, The Education University of Hong Kong, 10 Lo Ping Road, Tai Po, N.T., Hong Kong.  
E-mail address: [zyan@eduhk.hk](mailto:zyan@eduhk.hk) (Z. Yan).

system (5\*\* [5 star, star] for the highest 10 % and 5\* [5 star] for the next highest 30 % level 5 candidates). These processes generally lead to credible and reliable scores that can be used to rank students for grading, scholarship, entry, and so on. Students with extremely high levels of success on such examinations (e.g., multiple 5\*\* scores in Hong Kong) are publically identified and praised. The credibility of this method of identifying and rewarding talent depends on the impartiality and rigor of the system used to assess achievement or competence.

Nonetheless, despite the power of examinations to overcome collusion, corruption, or deception (Cheung, 2008), it is clear that students from disadvantaged backgrounds tend not to perform as well as those with privileged homes (English, 2002). Furthermore, examination at the end of a cycle of learning, does not provide feedback to learners about what they could do to improve. Thus, while the unfair social consequences of student assessment may be problematic for critics of contemporary society, the obvious and powerful consequences of standardised, public examination contributes to shaping consensus as to its legitimacy.

### 1.2. Assessment for learning

In contrast, assessment for learning emphasises that assessment is an integral part of learning and should be used as a tool to support learning and teaching. ARG (2002) defines assessment for learning as “the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.” (pp. 2–3). Fundamentally, the goal is to overcome the challenges associated with summative or terminal assessment before any it’s too late to do anything about it effect.

National responses to this trend have varied, with some systems totally abolishing high-stakes examinations (e.g., Queensland) and others trying to balance classroom formative and external summative assessment (e.g., New Zealand) (Berry, 2011b). Although the ARG definition has been widely cited, the interpretation of the definition and the ways it is reflected in educational policy and practice are different and sometimes deviate from the original principles (Klenowski, 2009). For example,

*‘deciding where the learners are in their learning, where they need to go and how best to get there’, has sometimes been (mis)interpreted as an exhortation to teachers to (summatively) test their students frequently to assess the levels they attain on prescribed national/state scales in order to fix their failings and target the next level. In this scenario, scores, which are intended to be indicators of, or proxies for, learning, become the goals themselves. Real and sustained learning is sacrificed to performance on a test. (Klenowski, 2009, p. 263)*

Nonetheless, the fundamental purpose of assessment for learning is to use assessment processes (e.g., sharing intentions and criteria, engineering classroom activities to elicit learning, providing feedback, involving students in assessment; Leahy, Lyon, Thompson, & Wiliam, 2005) within teaching contexts, prior to terminal or summative decision making, to improve student learning. Teachers do this by monitoring closely the work students produce and the processes they use to complete that work and consequently adapting teaching activities to move students toward greater success. By treating assessment processes as pedagogical processes (Black & Wiliam, 2006; Stobart, 2006), assessment for learning seeks to change the way classrooms work. It could be argued that assessment for learning is good teaching and not actually assessment (Brown, 2013, 2019), but the general stance of assessment for learning is to prioritise pedagogical and learning-oriented practices (Mok, 2010) over summative evaluations.

Sometimes assessment for learning is used interchangeably with formative assessment (Wiliam, 2018), but some researchers see them as distinct concepts. Formative assessment, with its origins in Scriven (1967) and Bloom, Hastings, and Madaus (1971), positions assessment

(usually tests and quizzes) during teaching to guide adjustments by the teacher to curriculum and pedagogy so as to achieve intended learning outcomes. In contrast to this teacher- and test-centric approach, assessment for learning (Black & Wiliam, 1998b) reflects a pedagogical approach in which assessment-like techniques (i.e., clear learning intentions, questioning, feedback, peer assessment, and self-assessment) are marshalled to create a student-focused, non-test learning environment. Within assessment for learning, there has been resistance to this approach as being too clearly focused on pre-determined learning outcomes with Swaffield (2011) suggesting that pure assessment for learning focuses on learner-defined outcomes and processes.

The policy documents on assessment reform in Hong Kong use the term formative assessment as an umbrella term that covers assessment for learning and assessment as learning to contrast with the focus on summative examinations. Assessment for learning was promoted earlier (Curriculum Development Council, 2002, 2009) focusing on the teacher’s role in using assessment for formative purposes; while assessment as learning appears later (Curriculum Development Council, 2014, 2017) to highlight the importance of students’ active and reflective role in assessment processes. Obviously such a differentiation between assessment for learning and assessment as learning is oversimplified. Interested readers may refer to Yan & Boud (in press) for a more detailed discussion on this topic. This article focuses on assessment for learning for a practical reason: assessment for learning has been promoted for two decades in Hong Kong, while the concept of assessment as learning is relatively new.

### 1.3. Educational assessment in Hong Kong

Historically, China used a progressive series of increasingly challenging examinations to select candidates from the general population for various levels of bureaucratic functions. Not only did the exam system permit social and personal life improvement, academic success was a signal of personal worth and virtue (China Civilisation Centre, 2007). This system is popularly described as ‘Confucian’ because its origins are associated with the philosopher Kong Fu Tse. However, the notion that Chinese or Hong Kong systems of assessment are somehow intrinsically Confucian reflects a ‘thin’ version of Confucianism, rather than one deeply embedded in Confucian philosophical principles (Kennedy, 2016). Nevertheless, ‘Chinese people have a tradition of changing their lives through examinations’ (Dorgan, 2000, p. 15). The imperial examination system continued until the early stages of the 20th century with the fall of the Qing dynasty. Nonetheless, even in modern China, there has been extensive use of high-stakes public examination at the end of middle school (*zhong kao*) and at the end of senior high school (*gao kao*) for selection into further opportunities. Hong Kong education is also characterised by its examination-dominated culture and reliance on high-stakes, summative public examinations (Biggs, 1998; Chan, 1986; Choi, 1999).

Hong Kong, drawing on imperial traditions and the British public school examination culture (e.g., 11+, ‘O’ levels, ‘A’ levels), has continued the use of public examinations for important decision making. The Hong Kong Certificate of Education Examination (HKCEE) conducted at Secondary 5 and the Advanced Level Examination (HKALE) in Secondary 7 have been replaced by the Hong Kong Diploma of Secondary Education (HKDSE) in 2012 in order to reduce the number of public examinations. In this background, one public assessment that existed solely for selection and accountability purpose, (i.e., Academic Aptitude Test; AAT) was abolished in 2001 for the sake of reducing the drilling that is likely to distort upper primary school education (Carless, 2005). Nevertheless, under the influence of two complementary traditions around examinations, Hong Kong has a widely accepted assessment of learning system that is used to make important decisions about students’ futures. The high-stakes examinations are regarded as a solid basis for selection into further schooling, as well as a legitimate means of moving up the social ladder (Berry, 2011a; Choi, 1999; Kennedy, 2005).

Although examinations provide surety against collusion, corruption, or nepotism (Cheung, 2008), they are often seen by students as oppressive burdens (Brown & Wang, 2013). The big problem of high-stakes examinations, as pointed out by many scholars (e.g., Biggs, 1996; Morris, 1985), is a negative test impact (also known as backwash or washback) effect on teaching and learning. That is, students and teachers focus their learning and teaching on the contents that are expected on the test. The curriculum is narrowed down and dominated by the examinations.

Some positive changes have been made to reduce this over-dependence on examination-oriented assessment culture and practices since the establishment of Hong Kong as a Special Administrative Region of China (Brown & Ngan, 2010). The Education Commission (2000), in the education reform proposal *Learning for Life – Learning through Life*, advocated less dictation exercises and mechanical drills, and recommended that schools should put more emphasis on assessment for learning. According to the curriculum structure that was revised in 2009, the 12-year schooling before university is broken into four 3-year Key Stages. Currently, all Hong Kong students receive 12 years of schooling, but just 20 % or so of students are selected, based on their HKDSE examination results, for government places in universities.

The major large-scale assessments in Hong Kong are summarised in Fig. 1 and readers are referred to Brown and Ngan (2010) for a detailed description of the assessment system in Hong Kong. The formative use of assessment has been promoted and practiced in various forms. For example, students with suspected special education needs, can be assessed by a registered educational psychologist with the Learning Achievement Measurement Kit (LAMK). Schools can monitor student attitudes and social experiences by using a battery of assessments that explore student attitudes, skills, and values in the Assessment Program for Affective and Social Outcomes (APASO) (Moore, Mok, Chan, & Lai, 2006). APASO is a component within the Education Bureau’s electronic School Development and Accountability system, which is designed to help schools conduct self-evaluations.

The Hong Kong Examinations and Assessment Authority (HKEAA) developed the Basic Competency Assessment (BCA) system from 2004. BCA contains a formative Student Assessment (SA) and a Territory-wide System Assessment (TSA). The SA provides teacher resources for assessing Chinese, English, and Mathematics from P1 to S3. The SA consists of a web-based central assessment item bank, online assessments, computer marking of student performance on online assessments, and instant reports as to students’ performance. Teachers are expected to flexibly use the SA, according to student needs and learning progress, to enhance student learning. Teachers design a test by selecting criteria that they wish to evaluate (i.e., they align the assessment to the curriculum they are implementing in their own classroom) and the system indicates task difficulty in terms of Key Stages. The system provides individual and class reports about items answered correctly or incorrectly, suggests possible causes for inaccuracy, and links each question to its basic competency criterion.

In contrast, the TSA is a system monitoring assessment conducted in all schools with samples of pupils at Primary 3, 6 (implemented every second year starting from 2011), and Secondary 3. The goal of the TSA is to 1) provide an overall picture of Hong Kong student performance so as to facilitate government’s policy-making; and 2) help schools to adjust

learning and teaching strategies according their students’ performances. The TSA are administered and scored against Key Stage competence descriptors by the HKEEA at set times in the school year. While the focus is not on students, public reports are made about levels of competence across the territory. Schools which have disproportionate levels of below basic competence are offered targeted support to raise student performance. The identity of schools needing support is kept confidential and all schools are told the performance of the school relative to the territory-wide averages. It is noteworthy that this monitoring system uses formal examination protocols to provide system diagnostic information.

There are two major summative public assessments. The Pre-S1 Hong Kong Attainment Test (Pre-S1 HKAT) is administered to all Primary 6 students. The results are used to moderate school-based assessment scores of the subsequent cohort in the same school and determine in part the type of secondary school students attend. This moderation of teacher judgements by examination is expected to raise the credibility of decisions attached to the assignment of students to Band 1–3 secondary schools.

To better account for aspects of secondary school competencies that require performances (e.g., oral language, laboratory skills, etc.), school-based assessments (SBA) contribute toward certification in the HKDSE. The SBA is worth 15 %–50 % of the total mark in a subject, with the balance obtained through the end-of-year written examination. While SBA is administered by teachers, quality assurance is obtained by the HKEEA by having common formats for SBA tasks across schools, quite specific rules guide the conduct of the SBA events, common scoring rubrics are used, and extensive moderation procedures are used. The HKDSE, supplemented by SBA, is the most significant public examination system that students encounter. Pupils are required to study four compulsory ‘core subjects’ (i.e., Chinese and English languages, mathematics, and liberal studies) and choose one to four elective subjects from a pool of 20 options. The performance in the HKDSE is reported in five levels (level 1–5; 5\*\* for the highest 10 % and 5\* for the next highest 30 % level 5 candidates). Given the level of government funding for places at university, candidates consistently scoring below Grade 5 are unlikely to be awarded a government place in universities.

#### 1.4. Purpose of this paper

The success of the assessment for learning reform depends on many factors (Lam, 2018). Consistent with Smith (1973), a number of components in the policy implementation process have to successfully collaborate for policy to be implemented. The policy making process has to create a policy that actively supports, rather than benignly tolerates or resists, the objective. Implementation organisations (e.g., examination authority, teacher professional education and development, subject professional associations), the target group (i.e., teachers, school leaders, parents, and students), and environmental factors (cultural, social, political, and economic conditions) all have to align to ensure success of a policy innovation.

In a systematic review, Heitink, Van der Kleij, Veldkamp, Schildkamp, and Kippers (2016) identified prerequisites for the successful implementation of assessment for learning from four aspects: (1) teachers with sufficient capacities and beliefs; (2) students with

Function	Key Stage 1			Key Stage 2			Key Stage 3			Key Stage 4		
	P1	P2	P3	P4	P5	P6	S1	S2	S3	S4	S5	S6
Formative	Basic Competency Assessments (BCA)											
			TSA			TSA			TSA			
Summative							Pre-S1 HKAT					
											HKDSE SBA	

Fig. 1. Hong Kong’s formative and summative education assessments by year level, updated from Brown and Ngan (2010).

sufficient knowledge, skills, and beliefs; (3) a school context that is supportive to assessment for learning and provides professional development; and (4) appropriate assessment design and implementation. Furthermore, culture and the local education context are also crucial factors influencing such an educational reform because assessment is not simply a technical issue (Brown, Kennedy, Fok, Chan, & Yu, 2009; Yan, 2014b). Any attempt to borrow assessment policy and practices from the international research literature should take into account local cultures and context factors (Adamson & Davison, 2003). When there are competing or contradictory tensions within or between various components in the implementation process (e.g., resistance by parents, continued use of tests to evaluate schools, or continued use of examinations to ration further educational opportunities), then discrepancies arise that can undermine a policy initiative (Smith, 1973). Hence, the goal of this paper is to examine the degree to which the assessment for learning policy has been successfully implemented in Hong Kong and identify tensions in the implementation process that impact its widespread replacement of the assessment of learning culture.

## 2. Method

This is a narrative review (Paré & Kitsiou, 2017) of Hong Kong's policy borrowing in assessment over the last two decades. Narrative reviews usually "do not involve a systematic and comprehensive search of all of the relevant literature. Instead, narrative reviews are often opportunistic in that they survey only that literature and evidence that are readily available to the researchers." (Paré & Kitsiou, 2017, p. 185, *italics in original*). In that spirit, the authors take advantage of their own privileged position as teaching and research academics in Hong Kong.

Intermediary organisations that connect research to policy and practice are necessary for effective policy reform (Corcoran, Rowling, & Wise, 2015). In that sense, both authors work or worked in the largest teacher education body in Hong Kong, putting them in the central intermediary organisation for developing teacher knowledge and skill in assessment for learning. The first author teaches future Hong Kong school teachers about assessment and has a special interest in self-assessment practices, having developed the Self-assessment Practices Scale with Hong Kong primary to university-level students (Yan, 2018b, 2020). The second author was also a teacher educator of curriculum and assessment in Hong Kong for three years, during which time he was the lead author of a Hong Kong teaching text on educational assessment. He continues his involvement through Honorary Professor status in the same department as the first author. His research in Hong Kong has focused on understanding teacher and student conceptions of assessment in collaboration with both mainland and Hong Kong Chinese researchers. The joint experience of the authors allows them to formulate a qualitative interpretation of curriculum policy documents and empirical research on assessment for learning in Hong Kong. While this approach is somewhat subjective and opportunistic, this type of review is useful in gathering together a volume of literature in a specific subject area and synthesising it (Paré & Kitsiou, 2017).

Our analysis is guided by a social positivist approach to educational transfer or borrowing (Perry & Tor, 2008) as applied to the assessment for learning policy in the Hong Kong context. The social positivist perspective takes the view that "larger global practices shape educational practices within a given national context" (Perry & Tor, 2008, p. 513). As Lingard and Lewis (2016) make clear, global priorities and practices can influence local situations and local situations can exploit and reinterpret global practices resulting in both globalised localisms and localised globalisms. This means that, while external forces (e.g., the U.K. assessment for learning reform) can be viewed as a transformative force on Hong Kong's educational policy and practice, it is also likely that local pressures to reform schooling (e.g., extending free education to 12 years of schooling) might exploit perceived advantages in the external policy to modify local conditions.

The methodological priority of social positivism is to examine

primary source materials (e.g., textbook and curriculum frameworks, policy documents) and survey studies involving large samples of the population (Perry & Tor, 2008). Hence, we examine official documents and empirical studies. The official documents analysed were policy documents from the Hong Kong government, such as the Education Commission and the Hong Kong Examination Authority, and the curriculum guide for primary and secondary education from the Curriculum Development Council, the curriculum agency of the government. Empirical studies included were research studies of teachers, school leaders, and students around the uses, purposes, and effects of assessment for learning in Hong Kong. Selection and interpretation of documents and studies was governed by the authors' insider experience of educational assessment practices and policies in Hong Kong.

## 3. Results

### 3.1. Development of the assessment for learning policy in HK

Despite the examination-dominated norms, the Hong Kong government made some initial attempts at better alignment between curriculum and assessment from as far back as the 1970s. In 1978, the Teacher Assessment Scheme (TAS), a form of school-based assessment (SBA), was incorporated into the Hong Kong Advanced Level Examination (HKALE) for the Chemistry subject and it accounted for 20 % of the final grade. Since then, the TAS was gradually extended to other subjects as well as the Hong Kong Certificate of Education Examination (HKCEE). Although the main purpose of TAS for HKALE and HKCEE was to cover learning areas (e.g., practical skills) that were difficult to be assessed by public examinations, it did provide assessment for learning opportunities embedded with the process of continuous assessment for qualifications over the whole course period (Yan, 2014b). Unfortunately, the implementation of TAS failed mainly because of heavy workloads for teachers, and a lack of confidence among teachers as to the accuracy and consistency of teacher scoring within and across schools (Berry, 2008; Cheung & Yip, 2004; Yip & Cheung, 2005).

The Target-Oriented Curriculum (TOC) in the 1990s was another large-scale initiative to involve teachers in monitoring and responding to student learning as part of regular coursework teaching. TOC was a form of outcome-based education that required teachers to record students' learning progress and provide formative feedback to support learning (Berry & Adamson, 2011; Lam, 2018). However, TOC did not flourish for a number of reasons including the discrepancy between TOC principles and the prevailing system of formal examinations in schools, as well as unfavourable attitudes and beliefs held by teachers, the heavy workload involved, lack of school support, and relatively low assessment literacy (Carless, 2011; Lam, 2019; Morris, Lo, & Adamson, 2000).

In response to Black and Wiliam's (1998a, 1998b) reviews on formative assessment and ARG's research on assessment for learning (1999, 2002), Hong Kong teacher education (Berry, 2008) and curriculum agencies (Curriculum Development Council, 2001) have systematically sought to introduce assessment for learning. In 2001, a large-scale curriculum reform was initiated emphasising students' "learning to learn" capacity (Curriculum Development Council, 2001). Specific calls to introduce assessment for learning were enunciated in 2002:

All schools should review their current assessment practices and *put more emphasis on assessment-for-learning*. The latter is a process in which teachers seek to identify and diagnose student learning problems, and provide quality feedback for students on how to improve their work. Different modes of assessment are to be used whenever appropriate for a more comprehensive understanding of student learning in various aspects. (CDC, 2002, Chapter 5, p. 1; *italics added*)

Assessment for learning was advocated in part because it was consistent with a curricular shift in educational goals from purely

examination or mastery of subject content knowledge to acquisition of more generic life-long skills. Subsequent policy documents have placed assessment as an integral part of the curriculum, learning and teaching and feedback cycle.

Assessment is an *integral part of the curriculum, pedagogy and assessment cycle*. It involves collecting evidence about student learning, interpreting information and making judgements about students' performance with a view to providing feedback to students, teachers, schools, parents, other stakeholders and to the education system. (CDC, 2009, Booklet 4, p. 1; italics added)

Assessment is the practice of collecting evidence of student learning in various aspects (including the learning process and learning outcomes); interpreting data, assessing students' performance for the purpose of providing feedback to students, teachers, schools, parents and other stakeholders as well as the education system, which are fundamental to improving learning and teaching. Therefore, *assessment is an integral part of the curriculum, learning and teaching and feedback cycle*. (Curriculum Development Council, 2014, Chapter 5, p. 2; italics added)

In the most updated curriculum and assessment guide for primary (Curriculum Development Council, 2014) and secondary education (Curriculum Development Council, 2017), the Curriculum Development Council reinforces the idea of assessment for learning and made a further step by encouraging schools to integrate assessment as learning into the learning and teaching circle:

Under "Assessment as Learning", *students should understand their learning targets, monitor their learning progress, reflect on what to learn and the learning strategies to adopt based on feedback, adjust their learning methods and future learning targets, or even plan for their future direction of learning*. In the long run, *students should become their own best assessor and provide feedback for their own learning*. (CDC, 2014, Chapter 5, p. 8; italics added)

*Assessment-as-learning engages students in reflecting on and monitoring their progress of learning through establishing their roles and responsibilities in relation to their learning and assessment*. Students use feedback from reflection and monitoring to make adaptations and adjustments to the learning objectives and strategies. (CDC, 2017, Chapter 4, p. 6; italics added).

Thus, it is clear that the Curriculum Development Council has fully adopted the notion of assessment for learning coherent with the UK Assessment Reform Group, with a strong emphasis on assessment as something that students do in learning, rather than something done at the end of learning or by teachers. It is also clear that the adoption of an assessment for learning policy built upon earlier efforts to move assessment from purely summative formal examinations. Nonetheless, the curriculum position of assessment enunciated in these extracts does not remove the important role of formal assessment mechanisms, such as testing and examination. Indeed, it appears that assessment for learning is an adjunct teaching strategy that has to co-exist with the public examination system.

### 3.2. Stakeholders' perceptions and practice of assessment for learning in Hong Kong

After a decade of promotion, CDC (2014) reported that Hong Kong schools have made various attempts to adopt assessment for learning, including formulating assessment policies that emphasise formative assessments, apply diversified modes of assessment, emphasise formative feedback, making full use of assessment data to support teaching and learning, and so on. A number of empirical studies focusing on assessment for learning tends to affirm its potential positive impact on student learning in Hong Kong. For example, assessment for learning practices in Secondary 1 classrooms showed a positive impact on

students' motivation in writing (Lee, 2011a) and their positive learning experience (Lee, 2011b). Lee and Coniam (2013) found that implementing assessment for learning resulted in better performance of Secondary 1 students in writing tasks, although no explicit evidence on enhancing student motivation was reported. They attributed the learning gain to students' better understanding of the requirements of writing tasks, which was a positive outcome of how teachers implemented assessment for learning. In a study on student teachers, Ng (2014) found that assessment for learning from the teacher and peers demonstrated positive impact on the quality of wiki projects. However, it should be noted in these studies that assessment for learning from the teacher was preferred because students appreciated the teacher's authority and experience. This means the student-oriented focus of assessment for learning in CDC (2014) may not yet be prevalent in practice.

Students' perception about assessment is crucial for investigating the potential contribution of assessment for learning (Conlon, 2006), especially if a view of assessment for learning is taken that prioritises the students' role. In a large-scale survey of Hong Kong primary school students ( $N = 3019$ ), Guo and Yan (2019) found that generally students had positive instrumental yet negative affective attitudes towards formative assessment. In other words, students perceived formative assessment as useful for their learning, but they did not like it. Gao (2009) reported that secondary students' perceptions of SBA for English ranged from SBA being just like exams to being indistinguishable from learning, indicating that much depends on the specifics of the teacher and the student. Brown and Wang (2013, 2016) reported that university students acknowledged assessment as a useful tool for supporting their learning, although they emotionally disliked assessment. Hue, Leung, and Kennedy (2015) found that Hong Kong students held a generally positive attitude toward the pedagogical role of assessment. In particular, primary students considered teacher-student interactive assessment as the most important practice because primary students preferred interacting with teachers. In contrast, secondary students cared more about examinations and, therefore, preferred teacher-dominated assessment because this provided quicker and achievement-oriented feedback.

Teachers' perception and practice with regard to assessment for learning appears more complicated. Hong Kong teachers generally believe in using assessments for improving student outcomes (Brown et al., 2009; Hui, 2012; Yan & Cheng, 2015). This may reflect the impact of policy promotion, professional development, and teacher education about the positive impact of assessment for learning. It may also reflect a more traditional idea that by testing students, teachers are contributing to the development of personal attributes such as positive learning attitudes and readiness for future challenges (Hui, 2012). However, despite endorsement of the principles of assessment for learning, it is still generally underused in classrooms (Lam, 2018; Yan & Cheng, 2015) and the extent of application varies among schools (Yu, Kennedy, Fok, & Chan, 2006). Many Hong Kong schools are still dominated by traditional examination-oriented assessment practices and are not using assessment for the purpose of supporting learning and teaching (Berry, 2010). Yan and Cheng (2015) surveyed 450 Hong Kong primary teachers and found that even though teachers had quite a positive attitude, higher levels of self-efficacy, and intention regarding formative assessment, there was little implementation of formative assessment in their classrooms. Lam (2016) reviewed the effectiveness of implementation of assessment for learning practices in the English language classrooms during the period 2004–14. He found that although Hong Kong teachers have become more receptive to innovative assessment practices, they remain sceptical of their new role in such a change. He, therefore, concluded that assessment for learning has a long way to go before its full integration into the local classrooms.

Based on the authors' knowledge, together with arguments available in literature (e.g., Hui, Brown, & Chan, 2017; Kennedy, Chan, Fok, & Yu, 2008; Lee & Coniam, 2013; Lam, 2018; Mak & Lee, 2014; Tang, Leung,

Chow, & Wong, 2010; Yan & Cheng, 2015), the reasons explaining the mismatch between teacher beliefs and practices regarding assessment for learning in Hong Kong touch on teacher factors and the contextual constraints they face, the two major prerequisites identified in Heitink et al.'s (2016) and (Yan et al.'s (2021)) reviews regarding the implementation of formative assessment. However, the major obstacle has to do with the overarching surveillance and evaluation culture.

### 3.3. Obstacles: teachers and their context

Successful implementation of assessment for learning places a high expectation on teachers who must not only believe in assessment for learning, but also have sufficient professional knowledge and skills, commitment, and willingness to put assessment for learning into practice (Carless, 2011; Lee & Coniam, 2013; Mak & Lee, 2014). Lack of a well-trained cadre of teachers with commitment was one of the major reasons for the unsatisfactory implementation of assessment for learning, even in its early forms of TAS and TOC. However, the successful change of assessment practices does not only rely on teachers' beliefs and practices, but also require a context that facilitates the changes (Davison, 2013).

Practical constraints in Hong Kong classrooms, such as large class sizes and heavy teaching workload, inhibit teachers' full implementation of assessment for learning. Successful assessment for learning needs deliberate planning, sufficient time for implementation in classroom practice, and individualised follow-up. It is difficult for a teacher to provide individualised feedback to students in a class with more than 30 students (Yan & Cheng, 2015). In order to reconcile the contradictions between the requirements of the assessment for learning initiative and the available resources in reality, there needs to be a common understanding and vision among the different stakeholders as to how to manage those expectations (Hargreaves & Shirley, 2009; Mak & Lee, 2014).

### 3.4. Obstacles: evaluation culture

Possibly the most important obstacle revolves around Hong Kong's systematic assessment culture. This refers to both evaluation of students through public examinations but also the wider social system of evaluating teachers and schools through systematic monitoring and evaluation of student achievement. Throughout Hong Kong there is a strong expectation that schools should place a heavy emphasis on enhancing students' examination scores, which is contradictory with the major principles of assessment for learning (Berry, 2008). The wider social system refers to the examination-driven culture in Hong Kong that challenges the implementation of assessment for learning and trivialises assessment for learning initiatives (Kennedy et al., 2008; Lee & Coniam, 2013). Assessment for learning practice in Hong Kong schools is largely constrained by the pressure to demonstrate good teaching quality through the preparation of students for examinations (Hui et al., 2017). This is exactly what concerned Black and Wiliam (1998b); that is, the requirements for certification and accountability would pose a substantial challenge to conducting formative assessment.

### 3.5. Unsuccessful assessment for learning initiatives

In addition to the less satisfactory implementation of assessment for learning in Hong Kong classrooms, the major assessment for learning initiatives (e.g., the school-based assessment (SBA) in public examinations) have not been successful despite their good intentions. It was expected that the introduction of SBA would de-emphasise the summative examinations and treat assessment as an integral part of the learning and teaching cycle (Yan, 2014b). However, in practice the summative function of the SBA overwhelms its formative role, thus, not conforming to the ideals of assessment for learning. Yan (2014a) found that Hong Kong teachers held a negative attitude towards SBA although

they had confidence in completing SBA tasks. Many studies investigating SBA or TAS in Hong Kong (e.g., Cheung & Yip, 2004; Davison, 2007; Qian, 2014; Yip & Cheung, 2005) identified the major challenges in implementing SBA, such as lack of sufficient support (e.g., assessment materials, teacher training), the heavy workload, the negative backwash effect on student learning, and potential unfairness. Furthermore, students also demonstrated a generally negative attitude towards SBA. Gao (2009) reported in a qualitative study that students' perceptions on the usefulness of SBA varied. Some appreciated SBA, while others complained that there was insufficient formative feedback from their teacher. In a study using both survey and interview, Tong and Adamson (2015) found that secondary students considered SBA as less useful than its intention and they were unable to make full use of the feedback. Teachers' and students' resistance to SBA raised public doubt about its validity (Davison, 2007).

Similar to SBA, the Territory-wide System Assessment (TSA) has become another controversial issue in Hong Kong. Although TSA was designed as a low-stakes assessment tool for the sake of enhancing learning and teaching quality, some key stakeholders have serious doubts about its rationale, purpose, and usefulness (Lam, 2018). Despite the government reiterating the formative use of TSA, many school principals perceived that TSA was playing an evaluative rather than pedagogical role (Ngan, Lee, & Brown, 2010). The leaders believed that the TSA results would be a tool to sanction and/or punish underperforming schools. Accordingly, some schools had introduced after-school tutorials for preparation for the TSA. Similarly, for many teachers, TSA has been regarded as a high-stakes test because it acts as a tool to punish rather than assist schools that had unsatisfactory results (Hong Kong Professional Teachers' Union, 2015). As a high degree of trust between stakeholders is a prerequisite for successful implementation of assessment reform (Hopfenbeck, Tolo, Florez, & El Masri, 2013), the unsatisfactory implementation of TSA in Hong Kong is not a surprise.

As a response to these critical voices about TSA, especially as it applied to young children in Primary 3, the TSA policy was reviewed and some adjustments were made in 2018 with a purpose of relieving accountability pressure and motivation to drill students. For example, no student names and school names are required, no collection of school reports by the government, and only 10 % of Primary 3 students from each school are sampled for the assessment, although schools can opt in to whole-level participation. Nevertheless, the public concerns seem not fully resolved and a boycott against participating in TSA continues at the time of writing this paper.

Generally, it would seem teachers and school leaders do not perceive the BCA and TSA systems as improvement or learning-oriented (Hui et al., 2017; Ngan et al., 2010). This conclusion is what Kennedy et al. (2011) have analysed by concluding that assessment for learning is a soft policy, in contrast to the hard policy of the continued emphasis on public examinations in Hong Kong.

## 4. Discussion

Our analysis shows that Hong Kong's borrowing of the assessment for learning policy is ambitious but largely futile, at least at the current stage, because of official use of public examinations for important decision-making (e.g., entry to university, evaluation of school quality, assignment of students to secondary school bands, etc.). Moreover, the implementation of assessment for learning is also hindered by practical constraints such as insufficient teacher training and school-based support, big class sizes, and heavy teacher workload. These constraints are amplified by cultural and social systems that normalise academic success and norm-referenced competition. As argued by Mok (2007), the move from examination-driven culture to embracing assessment for learning requires major paradigm shifts. On one hand, assessment should be regarded as a tool for support effective learning and teaching rather than a selection and accountability mechanism. On the other hand, proper infrastructure needs to be built in schools so as to facilitate

changing assessment practice.

The results challenge the notion of universal borrowing of educational policy without consideration of cultural, societal, and historic norms. As suggested by Carless (2011), in encouraging teachers to use summative tests formatively, a whole-scale shift to a pure assessment for learning approach in Hong Kong could only happen through a more gradual implementation of the assessment for learning practices. At the system level, despite some positive moves (e.g., reducing the number of public examination; introducing SBA), assessment for learning desires of Hong Kong teachers and policy makers are unlikely to be fulfilled unless there is a change in access and support for access to higher or further education. The relatively recent provision of 12 years' universal schooling in Hong Kong suggests that the possibility of even greater access to higher and further education exists.

In Hong Kong, the borrowing of the assessment for learning policy aimed to achieve a balance between classroom formative assessment and external public examinations. The assessment for learning policy was borrowed because its principles were consistent with local concerns and issues already identified in the Hong Kong assessment system long before assessment for learning was invented. Hence, in one sense the adoption of assessment for learning was not a borrowing, but rather an evolutionary step amongst Hong Kong educational elites who sought to diversify the assessment context by de-emphasising the power of external, formal, summative examinations. So while assessment for learning seems to be a policy that is exported around the globe, Hong Kong's situation seems better understood as a case of a localised globalism. The local goals and priorities tapped into a global trend and adapted it to fit local conditions and long-standing priorities (Lingard & Lewis, 2016).

Unfortunately, the tension between assessment for learning, or formative assessment, and assessment of learning, or summative assessment, has always been an issue. Black and William (2018) noted that teachers from different jurisdictions shared the same concern: formative assessment is not feasible because they have to "teach to test" due to the pressure to raise their students' test scores. They commented that "This is a particularly ironic finding given the evidence of the impact of formative assessment practices on student achievement on standardised tests" (p. 552). The challenges associated with localising assessment for learning in Hong Kong are much more significant compared to other western countries/regions because the policy was borrowed into a context that was fundamentally antithetical. Unlike USA or England where the top-down, test-based accountability was just introduced in 1980s, China, including Hong Kong, has indulged in this mode of assessment for accountability usage for over 3000 years (China Civilisation Centre, 2007). The examination-driven culture has been deeply rooted in the social system as well as its members' mind-sets for not only a long-time but also for good reasons. For some Hong Kong teachers, "teach to test" means the spoon-feeding method of teaching and drilling which is strongly thought to be effective in upgrading student performance in summative assessment (Yan & Cheng, 2015). An interesting question to be asked is whether drilling is really superior to assessment for learning in terms of enhancing performance on summative examinations? Many people believe so but the empirical evidence supporting such a belief is surprisingly little (Carless, 2011). This issue taps into the parallel issue as to whether Chinese learners practice memory-based approaches to learning because they are intrinsically programmed to do so or because that approach is a strategic rational response to how assessments are implemented and used (Kennedy, 2016). These are interesting areas for future research.

However, at the classroom level, the implementation of assessment for learning is likely to be successful as a beneficial teaching/learning strategy, if three issues could be addressed. First, in addition to the merits of enhancing students' learning competence in the long term, assessment for learning has to cater for teachers', students', and parents' short-term goals (i.e., increasing examination scores for the benefits such bring). In other words, how to achieve the synergy between

formative and summative assessment, as advocated by Black and William (2018), would be a prerequisite for localising assessment for learning in Hong Kong. Given the existing examination-driven culture, teachers won't give priority to any teaching practices, including assessment for learning, if they think that practice fails to enhance the performance of students on high-stakes assessments. Thus, more attention should be put on how to indigenise assessment for learning through a proper adaptation of pedagogical formulations of assessment tasks and processes so as to accommodate localised expectations (Tang et al., 2010). As long as the quality of schools is associated with examination performance and as long as post-schooling opportunities for students are restricted by examination performance, it will be hard to persuade Hong Kong residents that examinations are not the real thing.

Second, students' active role in assessment needs to be highlighted in assessment practices in Hong Kong (Lee & Coniam, 2013; Yan, 2018a, 2018b). Hence, in addition to assessment for learning, assessment as learning should also be promoted as a further step in maximising student responsibilities in the assessment process (Yan & Boud, in press). On one hand, assessment as learning itself, is a productive teaching strategy that can provide ample learning opportunities for students during an assessment-like process; that is, the processes of task and criteria specification, task completion, task feedback by teachers and students, and revision opportunities are seen as a way of teaching. Through regular assessment as learning, students are able to monitor their own learning and gradually develop into self-regulating and life-long learners. On the other hand, regular assessment as learning can enhance students' engagement in assessment and develop their evaluative judgment (Harris & Brown, 2018). For example, they could have a better understanding of the nature and requirements of the learning tasks, the quality criteria, the meaning of feedback, and the way to take action on feedback. These features, in turn, can facilitate the process of and optimise the effect of teacher-directed assessment for learning. Furthermore, the practical constraints encountered in assessment for learning (e.g., big class sizes, heavy teaching workload) might be relieved since, in assessment as learning, the teacher is not the solely source of feedback and every individual student is a learning resource for themselves and one another (Yan, Brown, Lee, & Qiu, 2020). More importantly, in this approach the assessment-like activities are serving pedagogical purposes rather than assessment purposes (Brown, 2019). Fortunately, the government has realised the importance of assessment as learning and put it into a policy document (CDC, 2014). Nevertheless, it calls for a paradigm shift in terms of conceptualisation and implementation of educational assessment which as yet has not been achieved.

Third, assessment literacy has to be enhanced across the board; that is teachers, school administrators, parents, students, curriculum makers, policy makers, and teacher educators all have to value educational processes that focus on quality learning and be capable of strategically using assessment data for promoting learning. It is no doubt that the implementation of assessment for learning largely depends on individual teacher capacity and assessment literacy both of which need professional development and consistent support (Black & William, 2018). It should also be acknowledged that appropriate levels of assessment literacy among other stakeholders including students, principals, school personnel, and parents are important for the successful implementation of assessment for learning initiatives (Engelsen & Smith, 2014). Unless there is widespread public acceptance that chasing good examination scores is not the be-all and end-all of schooling, it will be difficult for teachers, even with the support of government policy, to instantiate a policy that flies in the face of public resistance..

## 5. Conclusion

This article critically reviews and discusses the major challenges arising from the educational assessment reform in Hong Kong. It provides a contextualised and critical analysis of borrowing assessment for learning, a western-originated idea, into a jurisdiction dominated by an

examination culture. In particular, the tensions between using assessment for learning as a pedagogical strategy and retaining a strong commitment to formal examination and testing are highlighted.

The borrowing of the assessment for learning policy was a natural response to the concern Hong Kong educators and officials had for decades about the dominance of test-based modes of accountability. School teachers demonstrate a general awareness and acceptance of the importance of assessment for learning in classroom practice. However, it is far from true that the assessment culture in Hong Kong has been changed. Despite the theoretical merits of assessment for learning in support of productive learning, it appears underused in classrooms.

The major challenge lies in the fact that the policy was borrowed into a fundamentally incompatible context and without addressing sufficiently the competing narrative of the hard policy of examinations. At the system level, despite some positive moves (e.g., reducing the number of public examinations; introducing SBA), the good intentions of assessment for learning policy are unlikely to be fulfilled until the superior position of examinations is reduced in Hong Kong. Nevertheless, at the classroom level, useful implementation of assessment for learning could still be possible if evidence and strategies were available to support the idea that assessment for learning leads to better examination performance. The formative use of summative testing strategy advocated by Carless (2011) is one piece of a much larger puzzle which has as yet not been completed. Without convincing proof that involving students in assessment and enhancing assessment literacy of all stakeholders leads to better results than the status quo, success of this policy is unlikely.

The assessment for learning reform is still ongoing in Hong Kong as this paper is written. Thus, our analysis does not provide an end-of-game evaluation. Instead, our aim is to present a realistic understanding of the status quo about the assessment for learning policy borrowing in Hong Kong. Our analysis describes factors impacting the full implementation of assessment for learning from the perspective of insiders who teach and research in the system. Although the analysis might not offer a solution on how best assessment for learning practice could be pursued in the Hong Kong context, we do hope it can give an overall picture of how assessment for learning is being conducted and the challenges all stakeholders are facing. Such a picture is crucial for drawing more attention and meaningful discussion among various stakeholders to arrive at a feasible solution.

## Acknowledgements

The first author was supported by a General Research Fund (GRF) (Project No.: EDUHK 18607118) from the Research Grants Council of Hong Kong SAR.

## References

- Adamson, B., & Davison, C. (2003). Innovation in English language teaching in Hong Kong: One step forward, two steps sideways? *Prospect*, 18(1), 27–41.
- Assessment Reform Group. (1999). *Assessment for learning: Beyond the black box*. Assessment Reform Group.
- Assessment Reform Group. (2002). *Assessment for learning: 10 Principles. Research-based principles to guide classroom practice*. Assessment Reform Group.
- Berry, R. (2008). *Assessment for learning*. Hong Kong: University Press.
- Berry, R. (2010). Teachers' orientations towards selecting assessment strategies. *New Horizons in Education*, 58(1), 96–107.
- Berry, R. (2011a). Assessment trends in Hong Kong: Seeking to establish formative assessment in an examination culture. *Assessment in Education: Policy, Principles and Practice*, 18(2), 199–211.
- Berry, R. (2011b). Educational assessment in Mainland China, Hong Kong and Taiwan. In R. Berry, & B. Adamson (Eds.), *Assessment reform in education: Policy and practice* (pp. 49–62). Springer.
- Berry, R., & Adamson, B. (2011). Assessment reform past, present and future. In R. Berry, & B. Adamson (Eds.), *Assessment reform in education: Policy and practice* (pp. 3–14). Springer.
- Biggs, J. (1996). The assessment scene in Hong Kong. In J. Biggs (Ed.), *Testing: To educate or to select* (pp. 3–12). Hong Kong Educational Publishing Co.
- Biggs, J. (1998). The assessment scene in Hong Kong. In P. Stimpson, & P. Morris (Eds.), *Curriculum and assessment for Hong Kong: Two components, one system* (pp. 315–324). Open University of Hong Kong Press.
- Black, P., & William, D. (1998a). *Inside the black box: Raising standards through classroom assessment*. School of Education, King's College.
- Black, P., & William, D. (1998b). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 5–74.
- Black, P., & William, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning* (pp. 9–25). SAGE Publication.
- Black, P. J., & William, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575.
- Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In R. W. Tyler (Ed.), *Educational evaluation: New roles, new means. The 63rd yearbook of the National Society for the Study of Education, part 2* (Vol. 69, pp. 26–50). University of Chicago Press.
- Bloom, B., Hastings, J., & Madaus, G. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw Hill.
- Brown, G. T. L. (2013). Assessing assessment for learning: Reconsidering the policy and practice. In M. East, & S. May (Eds.), *Making a difference in education and social policy* (pp. 121–137). Auckland, NZ: Pearson.
- Brown, G. T. L. (2019). Is assessment for learning really assessment? *Frontiers in Education*, 4(64). <https://www.frontiersin.org/articles/10.3389/educ.2019.00064/full>.
- Brown, G. T. L., & Ngan, M. Y. (2010). *Contemporary educational assessment: Practices, principles, and policies*. Pearson Education South Asia.
- Brown, G. T. L., & Wang, Z. (2013). Illustrating assessment: How Hong Kong university students conceive of the purposes of assessment. *Studies in Higher Education*, 38(7), 1037–1057.
- Brown, G. T. L., & Wang, Z. (2016). Understanding Chinese university student conceptions of assessment: Cultural similarities and jurisdictional differences between Hong Kong and China. *Social Psychology of Education*, 19(1), 151–173.
- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347–363.
- Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 12(1), 39–54.
- Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-Heritage settings*. Routledge.
- Chan, W. (1986). School examinations in Hong Kong. *Studies in Educational Evaluation*, 12(3), 355–357.
- Cheung, T. K.-Y. (2008). An assessment blueprint in curriculum reform. *Journal of Quality School Education*, 5, 23–37.
- Cheung, D., & Yip, D. Y. (2004). How science teachers' concerns about school-based assessment of practical work vary with time: The Hong Kong experience. *Research in Science & Technological Education*, 22(2), 153–169.
- China Civilisation Centre. (2007). *China: Five thousand years of history and civilization*. City University of Hong Kong Press.
- Choi, C.-C. (1999). Public examinations in Hong Kong. *Assessment in Education: Principles, Policy & Practice*, 6(3), 405–417.
- Conlon, T. (2006). Formative assessment of classroom concept maps: The reasonable fallible analyser. *Journal of Interactive Learning Research*, 17(1), 15–36.
- Corcoran, T., Rowling, L., & Wise, M. (2015). The potential contribution of intermediary organizations for implementation of school mental health. *Advances in School Mental Health Promotion*, 8(2), 57–70. <https://doi.org/10.1080/1754730X.2015.1019688>.
- Curriculum Development Council. (2001). *Learning to learn: The way forward in curriculum*. Author.
- Curriculum Development Council. (2002). *Basic education curriculum guide: Building on strengths (Primary 1 – Secondary 3)*. Author.
- Curriculum Development Council. (2009). *Senior secondary education curriculum guide: Building on strengths (Secondary 4–6)*. Author.
- Curriculum Development Council. (2014). *Basic education curriculum guide: To sustain, deepen and focus on learning to learn (Primary 1–6)*. Author.
- Curriculum Development Council. (2017). *Secondary education curriculum guide*. Author.
- Davison, C. (2007). Views from the chalkface: English language school-based assessment in Hong Kong. *Language Assessment Quarterly*, 4(1), 37–68.
- Davison, D. (2013). Innovation in assessment: Common misconceptions and problems. In K. Hyland, & L. L. C. Wong (Eds.), *Innovation and change in English language education* (pp. 263–275). Abingdon: Routledge.
- Education Commission. (2000). *Learning for life – Learning through life: Reform proposals for the education system in Hong Kong*. Author.
- English, F. W. (2002). On the intractability of the achievement gap in urban schools and the discursive practice of continuing racial discrimination. *Education and Urban Society*, 34(3), 298–311.
- Engelsen, K. S., & Smith, K. (2014). Assessment literacy. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 91–107). Springer.
- Gao, M. (2009). Students' voices in school-based assessment of Hong Kong: A case study. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about assessment for learning* (pp. 107–130). Information Age Publishing.
- Guthrie, J. T. (2002). Preparing students for high-stakes test taking in reading. In A. E. Farstrup, & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 370–391). Newark, DE: International Reading Association.
- Guo, W. Y., & Yan, Z. (2019). Formative and summative assessment in Hong Kong primary schools: Students' attitudes matter. *Assessment in Education: Principles, Policy & Practice*, 26(6), 675–699.

- Hargreaves, A., & Shirley, D. (2009). *The fourth way: The inspiring future for educational change*. Corwin Press.
- Harris, L. R., & Brown, G. T. L. (2018). *Using self-assessment to improve student learning*. Routledge.
- Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review*, 17, 50–62.
- Hopfenbeck, T., Tolo, A., Florez, T., & El Masri, Y. (2013). *Balancing trust and accountability. The assessment for learning programme in Norway: A governing complex education systems case study*. Oslo: OECD.
- Hue, M. T., Leung, C. H., & Kennedy, K. J. (2015). Student perception of assessment practices: Towards 'no loser' classrooms for all students in the ethnic minority schools in Hong Kong. *Educational Assessment, Evaluation and Accountability*, 27(3), 253–273.
- Hui, S. K. F. (2012). Missing conceptions of assessment: Qualitative studies with Hong Kong curriculum leaders. *The Asia-Pacific Education Researcher*, 21(2), 375–383.
- Hui, S. K. F., Brown, G. T. L., & Chan, S. W. M. (2017). Assessment for learning and for accountability in classrooms: The experience of four Hong Kong primary school curriculum leaders. *Asia Pacific Education Review*, 18(1), 41–51.
- Kennedy, K. J. (2005). *Changing schools for changing times: New directions for the school curriculum in Hong Kong*. The Chinese University of Hong Kong Press.
- Kennedy, K. J. (2016). Exploring the influence of culture on assessment: The case of teachers' conceptions of assessment in Confucian heritage cultures. In G. T. L. Brown, & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 404–419). New York: Routledge.
- Kennedy, K. J., Chan, J. K. S., & Fok, P. K. (2011). Holding policy-makers to account: Exploring 'soft' and 'hard' policy and the implications for curriculum reform. *London Review of Education*, 9(1), 41–54.
- Kennedy, K. J., Chan, J. K. S., Fok, P. K., & Yu, W. M. (2008). Forms of assessment and their potential for enhancing learning: Conceptual and cultural issues. *Educational Research for Policy and Practice*, 7(3), 197–207.
- Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy & Practice*, 16(3), 263–268.
- Lam, R. (2016). Implementing assessment for learning in a confucian context: The case of Hong Kong 2004–14. In D. Wyse, L. Hayward, & J. Pandya (Eds.), *The sage handbook of curriculum, pedagogy and assessment* (pp. 756–771). London: Sage Publications.
- Lam, R. (2018). Testing, drilling and learning: What purpose does the Grade 3 Territory-wide System Assessment serve? *Asia Pacific Education Review*, 19(3), 363–374.
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, 81, 78–89.
- Lee, I. (2011a). Bringing innovation to EFL writing through a focus on assessment for learning. *Innovation in Language Learning and Teaching*, 5(1), 19–33.
- Lee, I. (2011b). Formative assessment in EFL writing: An exploratory case study. *Changing English*, 18(1), 99–111.
- Lee, I., & Coniam, D. (2013). Introducing assessment for learning for EFL writing in an assessment of learning examination-driven system in Hong Kong. *Journal of Second Language Writing*, 22(1), 34–50.
- Lingard, B., & Lewis, S. (2016). Globalization of the Anglo-American approach to top-down, test-based educational accountability. In G. T. L. Brown, & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 387–403). Routledge.
- Mak, P., & Lee, I. (2014). Implementing assessment for learning in L2 writing: An activity theory perspective. *System*, 47, 73–87.
- Mok, M. M. C. (2007). Quality assurance and school monitoring in Hong Kong. *Educational Research for Policy and Practice*, 6(3), 187–204.
- Mok, M. M. C. (2010). *Self-directed learning oriented assessment: Assessment that informs learning & empowers the learner*. PACE Publishing.
- Moore, P. J., Mok, M. M. C., Chan, L. K. S., & Lai, P. Y. (2006). The Development of an indicator system for the affective and social schooling outcomes for primary and secondary students in Hong Kong. *Educational Psychology*, 26(2), 273–301.
- Morris, P. (1985). Teachers' perceptions of the barriers to the implementation of a pedagogic innovation: A South East Asian case study. *International Review of Education*, 31, 3–18.
- Morris, P., Lo, M. L., & Adamson, B. (2000). Improving schools in Hong Kong: Lessons from the past. In B. Adamson, T. Kwan, & K. K. Chan (Eds.), *Changing the curriculum: The impact of reform on Hong Kong's primary schools* (pp. 245–262). Hong Kong University Press.
- Ng, E. M. W. (2014). Using a mixed research method to evaluate the effectiveness of formative assessment in supporting student teachers' wiki authoring. *Computers & Education*, 73, 141–148.
- Ngan, M. Y., Lee, J. C. K., & Brown, G. T. L. (2010). Hong Kong principals' perceptions on changes in evaluation and assessment policies: They're not for learning. *Asian Journal of Educational Research and Synergy*, 1, 36–46.
- Paré, G., & Kitsiou, S. (2017). Methods for literature reviews. In F. Lau, & C. Kuziemsky (Eds.), *Handbook of eHealth evaluation: An evidence-based approach* (pp. 157–180). Victoria, BC: University of Victoria.
- Perry, L. B., & Tor, G.-H. (2008). Understanding educational transfer: Theoretical perspectives and conceptual frameworks. *Prospects*, 38(4), 509–526. <https://doi.org/10.1007/s11125-009-9092-3>.
- Qian, D. D. (2014). School-based english language assessment as a highstakes examination component in hong kong: Insights of frontline assessors. *Assessment in Education: Principles, Policy & Practice*, 21(3), 251–270.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Rand McNally.
- Smith, T. B. (1973). The policy implementation process. *Policy Sciences*, 4(2), 197–209. <https://doi.org/10.1007/BF01405732>.
- Stobart, G. (2006). The validity of formative assessment. In J. Gardner (Ed.), *Assessment and Learning* (pp. 133–146). SAGE Publication.
- Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy and Practice*, 18(4), 433–449.
- Tang, S., Leung, P., Chow, A., & Wong, P. (2010). A case study of teacher learning in an assessment for learning project in Hong Kong. *Professional Development in Education*, 36(4), 621–636.
- Tong, S. A., & Adamson, B. (2015). Student voices in school-based assessment. *Australian Journal of Teacher Education*, 40(2), 15–28.
- William, D. (2018). *Embedded formative assessment* (2nd ed.). Bloomington, IN: Solution Tree Press.
- Yan, Z. (2014a). Predicting teachers' intentions to implement school-based assessment using the Theory of Planned Behaviour. *Educational Research and Evaluation*, 20(2), 83–97.
- Yan, Z. (2018a). Student self-assessment practices: The role of gender, year level, and goal orientation. *Assessment in Education: Principles, Policy & Practice*, 25(2), 183–199.
- Yan, Z. (2014b). School-based assessment in secondary schools. In C. Marsh, & J. C.-K. Lee (Eds.), *Asia's high performing education systems: The case of Hong Kong* (pp. 274–287). Routledge.
- Yan, Z. (2018b). The Self-assessment Practice Scale (SaPS) for students: Development and psychometric studies. *The Asia-Pacific Education Researcher*, 27(2), 123–135.
- Yan, Z. (2020). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & Evaluation in Higher Education*, 45(2), 224–238.
- Yan, Z., Boud, D. (in press) Conceptualising assessment-as-learning. In Z. Yan & L. Yang (Eds.), *Assessment as learning: Maximising opportunities for student learning and achievement*. London & New York: Routledge.
- Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education*, 45, 128–136.
- Yan, Z., Brown, G. T. L., Lee, C. K. J., & Qiu, X. L. (2020). Student self-assessment: Why do they do it? *Educational Psychology*, 40(4), 509–532.
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice*. <https://doi.org/10.1080/0969594X.2021.1884042>.
- Yip, D. Y., & Cheung, D. (2005). Teachers' concerns on school-based assessment of practical work. *Journal of Biological Education*, 39(4), 156–162.
- Yu, Y. M., Kennedy, K., Fok, P. K., & Chan, K. S. (2006). Assessment reform in basic education in Hong Kong: The emergence of assessment for learning. *Paper Presented at the 32nd Annual Conference of the International Association for Educational Assessment, Singapore*.