Assessment for Learning Research in East Asian Countries

Abstract

Educational agencies in East Asia have heeded the advice of research findings and therefore acknowledge the value of assessment for learning (AfL) practices through policy initiatives. At definitional level, the evolving conceptions and theories of AfL have consistently streamed in from overwhelmingly European and Anglophone-based research. We present a review of mainly peer-reviewed journal articles on selected AfL research in East Asian countries. The findings show that the current implicit and atheoretical approach towards defining and implementing AfL suggests opportunities for further deliberation and theorisation about what constitutes AfL in East Asian countries. It is conceivable that teachers who understand the principles and frequently prepare students for summative assessment in the East Asian classroom are concurrently practising a particular process and practice of AfL. We conclude that the practices of AfL can therefore not just be variable; they will also be very situated and contested.

Keywords: educational assessment; educational policy; student evaluation; student improvement; assessment for learning; formative assessment

Introduction

Drawing further on developments in the fields of sociocultural learning theory, as well as metacognitive and self-regulation theory, a number of researchers (e.g. Black & Wiliam, 1998; Torrance & Pryor, 1998) have argued for even more support of formative assessment in past two decades, particularly regarding how student learning benefits from, rather than being deterred by, classroom assessment. The term 'assessment for learning', or AfL, first used by Black (1986), became part of classroom assessment discourse (Black & Wiliam, 2018; Gardner, 2010; Wiliam, 2011). The term AfL was further made popular through many publications by the Assessment Reform Group (ARG) in the UK as a follow-up to Black and

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Wiliam's (1998) seminal work. Since then, both formative assessment and AfL, as a distinct purpose and practice of classroom assessment of supporting learning and teaching, have been widely quoted, interchangeably, in education policy documents and research papers in several regions, including North America, Europe and Asia-Pacific, as an important educational innovation (Gardner, 2010; Organisation for Economic Co-operation and Development, 2005). More recently, and very notably, Black and Wiliam (2018) attempted to situate formative assessment within the context of a theory of pedagogy. We use the term 'AfL' in this review to include all versions of assessment initiative that privileges the purpose and practice of supporting learning, rather than aggregating or summarising performances.

In East Asia educational agencies have heeded the advice of such research findings and therefore acknowledge the value of formative assessment and AfL practices through policy initiatives. Countries with an explicit mention of formative assessment or AfL in their publicly accessed policy documents, suggesting an attempt to deepen systemic changes in assessment policy within the curriculum, include Hong Kong, Brunei, Malaysia, the Philippines and Singapore.

In Hong Kong assessment is viewed as an integral part of the learning and teaching cycle and systemic measures were introduced to ensure that assessment is valuable to learners (Curriculum Development Council Hong Kong, 2001). Brunei's curriculum emphasises the importance of feedback and the need to diversify pedagogical techniques tailored to student understanding (Ministry of Education Brunei Darussalam, 2013). The Philippines Department of Education has also defined formative assessment as part of the policy guidelines on classroom assessment that should be implemented in schools (Republic of the Philippines Department of Education, 2015). In Malaysia and Singapore reforms include

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

placing greater emphasis on formative assessments, and a move towards more learner-centred approaches (Ministry of Education Malaysia, 2013; Ministry of Education Singapore, 2017).

The AfL debate

In the *Third International Conference on Assessment for Learning,* held in Dunedin, New Zealand (2009), a position paper on AfL representing the views of 31 assessment experts from Asian-Pacific, North American and European countries (interestingly, excluding all East Asian countries) articulated a 'second-generation' definition of AfL that also attempted to capture the day-to-day continual teaching and learning practices of students and teachers:

> Assessment for learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that *enhance on-going learning*. (Klenowski, 2009, p. 264; italics added by author in closing)

Such re-articulation of AfL was seen as necessary, as the experts noted that the ways in which AfL is interpreted and made manifest in educational policy and practice reveal a 'misunderstanding of the principles, and distortion of the practices, that the original ideals sought to promote' (p. 264).

Swaffield (2011) also highlighted how an erroneous interpretation and misrepresentation of AfL in English schools through the National Assessment for Learning Strategy severely undermined what had been reported about an authentic version of AfL to support students' learning and eventually contribute to school improvement. She drew a sharp distinction between AfL that can support important principles such as making learning explicit, promoting learning autonomy and focusing on learning, as opposed to a performance-oriented version that has a 'procedural, ritualistic manner that belies their pedagogical essence' (p. 438). The latter version of AfL is not expected to deliver its full potential effect, based on the underlying principles of the family of practices. Researchers

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

such as Stiggins (2005) and Swaffield (2011) suggested that formative assessment and AfL are not synonymous and that conflating their meanings could lead to difficulties of misappropriation and inefficacious measuring of impacts of either practice.

In reviewing the development of formative assessment and AfL in recent years, we note the following emphasis: the recognition that formative assessment (and subsequently AfL) needs a separate technology within classrooms that has so far not been well understood and is under-developed; such a distinction is necessary and possible, as different inferences need to be drawn from different assessment outcomes. Since the seminal publication of Black and Wiliam's review in 1998, there has been considerable investment in 'practical implementations' of various formative assessment practices (Black & Wiliam, 2018, p. 1); moreover, the basis of such developments of formative assessment was built on pragmatic classroom practices of 'what works' (Baird, Andrich, Hopfenbeck, & Stobart, 2017, p. 337) and not a 'pre-defined theoretical base' (Black & Wiliam, 2018, p. 2.). The last point, in particular, is critical, as it suggests that there is unfinished work when it comes to even the conceptualisation of formative assessment and AfL.

Indeed, other researchers have raised questions about the validity of current definitions of AfL and formative assessment on the grounds of inconsistencies in the many claims about the principles and vagueness of its definition, impacts and effects (Bennett, 2011; Dunn & Mulvenon, 2009; Maxwell, 2004). The difficulties inherent in realising the principles and strategies of formative assessment in the particular context of schools and classrooms have prompted researchers such as Taras (2005, 2009), Carless (2011) and McMillan (2010) to theorise on different possible versions of formative assessment that should be studied closely. Taras (2005) highlighted that formative assessment should not be seen as 'a magic formula' that is separate from, and incompatible with, summative

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

assessment. Citing from Scriven's (1967) earlier conceptualisation of assessment as a single process, she argued that the separation of formative and summative assessment, based on its differing functions, was 'self-destructive and self-defeating', and unfairly demonising the centrality of the summative assessment process in a teacher's day-to-day classroom activities. Her view then is that the same assessment outcome data can potentially serve different functions, as long as the *process* of assessment has been clearly thought through. Controversially, she emphasised the need to appreciate the importance of summative assessment as the missing link in formative assessment (Taras, 2009). Carless and McMillan suggested that formative assessment could not be thought of as a single entity in the classroom but as a *family of practices* that differed in some characteristics or levels of formative-ness. McMillan (2010) and Carless (2011) identified the presence or absence of characteristics for different levels of formative assessment, which could straddle between 'low-level' and 'high-level' (McMillan, 2010) or 'limited' and 'extended' (Carless, 2011).

The evolving standards of attending to different purposes, practices and processes of classroom assessment may be useful for further deliberation and theorisation about what constitutes effective AfL that is supportive of students' learning. However, such a distinction of decision-making, including the difference between AfL and formative assessment, may be excessively convoluted and widen what teachers 'need to' and 'are actually' conceptualising and practising. Furthermore, the underpinning theories of any particular version of AfL are not necessarily easily generalisable across different classroom contexts, as they are value-laden and contestable. The over-simplification of categorising a particular practice into just formative versus summative, or AfL versus AoL, has already been noted by researchers (Bennett, 2011).

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

The East Asian context

Indeed, despite the recognition that formative assessment and AfL have received at policy level, concerns about the adoption of such policy in East Asian countries have been raised. Such concerns are particularly salient in examination-driven societies, where the historical valuing of examinations as a means of upward social and economic mobility poses challenges to certain initiatives. For example, the importance of college entrance examinations, perceived as a 'fair' way of selecting students, has caused teachers to resist new ways of teaching ELT in Japanese schools when traditional teaching approaches were deemed more practical to prepare students for the written exams (Hadley, 1997). In Malaysia high-stakes examinations continue to dominate, undermining other functions of assessment not related to selection and certification (Ong, 2010). Similarly, in Hong Kong, in spite of policy that explicitly recognises that a reliance on written tests and examinations as major methods of assessment could have a narrowing effect on learning, 'assessment for selection' practices are a key feature in the curriculum (Curriculum Development Council Hong Kong, 2001).

The implementation of AfL in East Asian classrooms may encounter further barriers such as deeply rooted learning traditions (Thanh Pham & Renshaw, 2015). For instance, East Asian countries are known to be hierarchical with high power distance indexes (Hofstede, Hofstede, & Minkov, 2010). People in these societies accept that there is a social order that they must follow, with some holding more power and authority than others. Within the classroom, this would mean a hierarchical relationship between teacher and student, where the teacher is respected and treated with deference, sometimes even fear (Hofstead et al., 2010). This respect accorded to teachers may be explained by the belief that teachers are responsible for acquiring knowledge from authoritative sources and delivering that knowledge to their students. Thus, students regard teachers as definitive sources of

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

knowledge and easily accept their role as passive listeners, rarely challenging their teachers (Thanh, 2014). Such classrooms also tend to be teacher-centred. For instance, Dang's (2016) study on noticing-based collaborative feedback showed that, although many students in the study preferred collaborative feedback, their beliefs about its effectiveness were split, with a large number still trusting their teachers' corrections more than those of their peers. Dang (2016) suggested that one reason for this is the nature of Vietnamese classrooms, where the teacher's authority is still highly regarded. In Thailand itself the idea of student-centred learning is completely foreign to Thai people, as there is no Thai equivalent for the word (Panhoon & Wongwanich, 2012). Therefore, the idea of adopting a more student-centred approach to formative assessment or AfL may not be easily accepted in East Asian cultures, where teachers' authority is accepted without question, and at times even expected, in the classroom.

The existence of the macro-scale 'cultural learning models' (Wang & Li, 2003), such as Asian models of collectivism, may also play a part in the understanding and implementation of AfL in East Asia. Thanh Pham and Renshaw's (2015) study has shown how students were spurred on to try new ways of learning in order to help their group score well, and how they appreciated the benefits of collective learning. Educational stakeholders in East Asia would do well to tap into, or at least recognise, such cultural factors to enhance students' learning.

At the definitional level, the multiple evolving conceptions and theories of what constitutes AfL have consistently streamed in from overwhelmingly European and Anglophone-based research (Wiliam, 2011). AfL-related conferences, such as the International Invitational Assessment for Learning Symposium, also see a majority of representation by European and Anglophone countries, potentially leaving particular

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

considerations regarding East Asian classrooms (or other parts of the world) out of the conversations. These considerations may limit the wealth of perspectives in the conceptualisation and use of AfL. There is therefore room to question the application of a particular version of AfL to an East Asian school and classroom context.

We purposefully selected the region 'East Asia', as the authors reside in this region; we anticipated that our familiarity with specific countries in this region would help us to interpret the research findings more optimally. Specifically, in the context of East Asian countries, we focussed on the following research questions:

(1) How do the researchers define AfL in their studies?

(2) What are the implementation issues relating to AfL?

Given how AfL is reported to be an evolving and problematic construct, it would be valuable to find out if there was a preference for any particular definition(s) or whether it is necessary to redefine what AfL means in East Asian countries. An awareness of the kind of implementation challenges that may be attributed to a cultural or political context might also help to inform researchers and educators in the region on how to appropriate AfL more productively in their own context. We hope that these two research questions would be a start in the process of understanding the conceptualisation and implementation of AfL in this region.

Methodology

A few sources on literature reviews guided our methodology (Evans & Benefield, 2001; Petticrew & Roberts, 2016; Weed, 2005; Weed, 2008). We started with a broad search and selected our studies based on the criteria that we have developed before synthesising the

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

findings. Rather than using a positivist approach, we adapted Weed's (2008) more interpretivist approach, which focusses on 'locating meanings within the context in which they emerge' (Weed, 2008, p. 19). Our interest lies in studies on AfL within the East Asian context, and we wanted to ensure that, whatever interpretations were made, this was done while keeping the context of the selected studies in mind. Furthermore, as Nisbett (2003) has shown, Asians tend to pay more attention to context and also to favour relationships, so it was important to preserve these when we discussed the findings. In addition, as suggested by Weed (2008), our search was more cyclical and iterative, where we included more search terms that were deemed relevant as we progressed, rather than being too restrictive from the start. This was particularly useful because of the dearth of AfL-specific research in the East Asian context.

Procedure

We began our search in July and ended it in September 2017. The search phrase we used comprised two parts. One part is an AfL-related search term, while the other is the country where the study was conducted, for example, 'assessment for learning and Singapore'. We used the EBSCO research database, as well as Google Scholar, for this purpose. Only full-text peer-reviewed journals were retrieved for consideration. As 'formative assessment' and 'AfL' are often used interchangeably, we used both terms in our search.

The initial results suggested that AfL research could also be found under terms such as 'classroom assessment', 'feedback', 'peer-assessment' and 'self-assessment', so we broadened our search terms to include these.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

We confined our search to contexts that were similar in terms of educational settings. Thus, the titles and abstracts of the initial research results were filtered using the following inclusion criteria:

- Peer-reviewed journal articles, book chapters or reports published from 2009 onwards. Conference papers were excluded.

- They offer empirical evidence.

- They have references to AfL or formative assessment.

- They were conducted in any East Asian regions.

- They were conducted in the primary or secondary school context. E-learning or ICT-based learning was excluded.

- The results are discussed within a broader macro context, either political or cultural.

Only evidence-based research was included in our study, thereby allowing us to discuss the findings based on what is happening in the real school context rather than on an author's opinion. We adopted the Education For All (EFA) classification, as reported by the United Nations Educational, Scientific and Cultural Organization (UNESCO) to define East Asian regions, that is, China, Hong Kong, Taiwan, Japan and South Korea, including Southeast Asian countries such as Singapore, Malaysia, Brunei, Thailand, Vietnam, Indonesia and the Philippines. We chose this classification based on the close geographical proximity of these countries, as well as on similar shared traits. For instance, many of these countries still experience challenges in terms of educational quality and equity (UNESCO, 2015), and many have included AfL as part of their education policy and reform, as discussed above. Thus, it would be interesting to see how AfL is enacted in these countries.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

In terms of school context, studies in vocational and tertiary educational settings were excluded. One reason for this is that such educational settings tend to be different in terms of their mission, structure and curriculum, compared to primary and secondary schools. Studies that involve the use of ICT were excluded for the same reasons. We acknowledge that if AfL has to be 'part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information' (Klenowski, 2009, p. 264), it needs to be understood in typical 'day-to-day' classroom settings. We therefore recommend that such studies merit a review of their own, separate from this one.

Results

In the initial search, 196 articles were selected based on title and abstract. The articles were then read in their entirety and reviewed again based on the inclusion criteria, bringing the number down to 24. Several rounds of discussions were held with all the four authors to decide whether to include an article in the final list. All four authors are in the education field and have either a master's or doctorate's degree in education. An overview of these 24 articles is shown in Table 1 for quick reference.

It is worth noting that almost all of the studies cited feedback as being central to AfL. In these studies, feedback is either for teachers, to help them modify their teaching practices (Sardareh, Saad, Othman, & Me, 2014; Yatab & Shahrill, 2014; Zhao, Van den Heuvel-Panhuizen, & Veldhuis, 2016), or for students to use to improve their learning (Brown, Kennedy, Fok, Chan, & Yu, 2009; Bryant & Carless, 2010; Gu, 2014; Hui, Brown, & Chan, 2017; Koh, Lim, Tan, & Habib, 2015; Lam, 2013; Leong, 2014; Rashid & Jaidin, 2014; Tong, 2011; Yu, 2015), or both (Butler & Lee, 2010; Cagasan, Luo, Robertson, & Care, 2016; Griffin, Cagasan, Care, Vista, & Nava, 2016; Sardareh, 2016; Yin & Buck, 2015).

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

The articles selected were then read thoroughly several times and data relevant to the research questions were extracted into a data extraction form which was shared among the authors. From the form, two categories, "Definitions/conceptions of AfL/FA" and "Cultural/political context discussed in the article" were further analysed. For the first category, we looked for the definitions and terms that were used in the articles and noted their frequency. For the second category, we conducted a thematic analysis on the data that was extracted and came up with several themes.

Definitions of AfL

Of the 24 articles, 11 used the term 'formative assessment' (Bryant & Carless, 2010; Butler & Lee, 2010; Cagasan et al., 2016; Griffin et al., 2016; Gu, 2014; Lin et al., 2013; Ratnam-Lim & Tan, 2015; Tan, 2016; Tong, 2011; Yan & Cheng, 2015; Yin & Buck, 2015). Eight used the terms 'AfL' and 'formative assessment' interchangeably, assuming them to be the same thing (Azis, 2015; Hui, 2012; Hui et al., 2017; Koh et al., 2015; Lam, 2013; Rashid & Jaidin, 2014; Yu, 2015; Zhao et al., 2016). Only Leong (2014) made a distinction between AfL and formative assessment, with formative assessment seen as something broader that adheres less strictly to three principles highlighted by Swaffield (2011), namely, transparency of learning, autonomy of learners and a focus on learning.

In many of the articles, the researchers refer to formative and/or AfL as an assessment practice that has a distinct purpose of improving students' learning during the process of teaching. Black and Wiliam's (1998) definition was the most widely quoted, with nine of the studies referencing their work on AfL (Azis, 2015; Bryant & Carless, 2010; Gu, 2014; Hui et al., 2017; Koh et al., 2015; Sardareh, 2016; Yan & Cheng, 2015; Yatab & Shahrill, 2014; Zhao et al., 2016). Three of the studies (Cagasan et al., 2016; Griffin et al., 2016; Rashid &

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Jaidin, 2014) quoted more recent works by Black and Wiliam but retained the same essentials from their earlier work, that is, that information gleaned from students' learning is used to improve their learning. Meanwhile, Klenowski's (2009) definition was quoted in only two of the articles (Hui et al., 2017; Lam, 2013). Interestingly, Hui et al. (2017) raised concerns about whether assessment can really support learning in the Hong Kong high-stakes examination context. Griffin et al. (2016) preferred to use the term 'assessment for teaching' instead of 'assessment for learning', citing the need for teachers to play a central role in guiding students before any assessment can be of use to the students.

In two of the articles, acronyms were used to indicate the localisation of AfL in some countries. School-based assessment for learning (SBAfL) in Brunei was part of an initiative to equip students with twenty-first-century skills (Rashin & Jaidin, 2014), while holistic assessment (HA) in Singapore was meant to move schooling away from an exam-oriented culture to focus on a more holistic development of the student that includes non-academic aspects (Ratnam-Lim & Tan, 2015). These localised versions of AfL were introduced and promoted by policy-makers in the respective countries.

Seven of the studies did not explicitly explain the definition of formative assessment and/or AfL (Butler & Lee, 2010; Hui, 2012; Leong, 2014; Ratnam-Lim & Tan, 2015; Tan, 2016; Tong, 2011; Yu, 2015), assuming the reader's familiarity with them. None of the studies have discussed the theoretical basis for choosing one definition over the other, and none seems to problematize the concept of 'AfL' as a useful separate construct from summative assessment. There is also a general assumption that 'improvement of learning' is a well-understood notion.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Implementation issues of AfL

Half of the studies were concerned with policy implementation. These are usually in relation to an educational policy or reform where AfL is introduced, and the researchers were interested to know how these reforms affected schools. The studies either sought participants' perceptions of the implementation of AfL (Hui, 2012; Hui et al., 2017; Rashid & Jaidin, 2014; Ratnam-Lim & Tan, 2015; Tan, 2016; Yatab & Shahrill, 2014) or included observations of how teachers actualise AfL policies in classroom practices (Bryant & Carless, 2010; Cagasan et al., 2016; Griffin et al., 2016; Gu, 2014; Sardareh, 2016; Tong, 2011).

The greatest obstacle to the implementation of AfL, from the studies in this review, appears to come from the pressure of high-stakes external examinations (Bryant & Carless, 2010; Hui et al., 2017; Tan, 2016; Yatab & Shahrill, 2014; Yu, 2015). Teachers who believe in AfL are unable to carry these practices out, or give them up in favour of more summative-oriented assessment practices, because of the need to prepare students for these high-stakes examinations. Students' narrow focus on examinations is also seen as a hindrance when they do not see the link between what they learn and how it will help them in the exams (Koh et al., 2015; Lin et al., 2014). Many of the studies also reflect an emphasis on accountability in East Asian educational institutions (Griffin et al., 2016; Hui et al., 2017; Tong, 2011; Yan & Cheng, 2015). Griffin et al. (2016) suggested the need for more forms of accountability in the Philippines education system. Azis (2015) found that Indonesian teachers' conceptions of assessment had a close affinity with accountability conceptions.

Some studies also highlighted the highly authoritative nature of East Asian classrooms. In such classrooms it is generally accepted that teachers are the ultimate directors and decision-makers (Cagasan et al., 2016; Griffin et al., 2016; Lam, 2013). Research on questioning and feedback in Malaysia showed that teachers had the ultimate authority over

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

knowledge in the classrooms, with teachers still providing the correct answers most of the time, leaving little space for dialogic interactions (Sardareh, 2016; Sardareh et al., 2014). In fact, a few studies showed that students prefer teacher assessment to peer assessment (Bryant & Carless, 2010), sometimes getting frustrated with discussions when they would rather be told the 'correct answer' (Yin & Buck, 2015) by the teachers, who are considered the 'experts' in the classroom. They also suggest that students are afraid to reveal their misconceptions in the classroom for fear of losing the respect of their teacher and peers.

The teachers' mindset, influenced by their own upbringing and beliefs, colours their practice in the classroom. Some teachers believe in AfL as a good learning strategy for students, but are constrained by institutional values and school culture (Yu, 2015), as well as pressure from parents (Hui et al., 2017; Ratnam-Lim & Tan, 2015). There is also the belief, on the part of both students and teachers, that grading and scoring are still a 'valued part of learning', not at all incompatible with AfL (Azis, 2015), and that motivating students through test scores is an accepted way to make students responsible for improving their own learning (Brown et al., 2009; Bryant & Carless, 2010; Yin & Buck, 2015). Furthermore, it is difficult to dislodge the idea that examinations are inherently more fair and objective (Tong, 2011).

Teachers are also limited by their knowledge and competency of AfL (Sarderah, 2016; Sardareh et al., 2014; Yan & Cheng, 2015). Some are unconvinced of the potential benefits of AfL (Koh et al., 2015; Rashin & Jaidin, 2014) in contrast to tried and tested practices (Leong, 2014; Tong, 2011). In certain instances, AfL is perceived as a 'Western construct' (Bryant & Carless, 2010).

Then there are teachers who find ways of making AfL work in high-stakes examination settings, coming up with strategies to ensure that formative assessment is still carried out despite the challenges (Lam, 2013; Leong, 2014). Butler and Lee (2010)

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

suggested that formative and summative assessment should 'meet halfway', that is, to use summative for formative purposes, and vice versa, as well as limiting feedback to being on the process rather than the performance. In Brunei, teachers were already implementing BCATs (Brunei Common Assessment Tasks), which were 'formal, standardised and moderated assessment tasks that familiarise teachers and students with assessment for learning (AfL) best practices' (Ministry of Education Brunei Darussalam, 2013) and had both a formative and a summative component (Yatab & Shahrill, 2014).

Discussion

The findings reveal various mediating influences that inform nation- or school-wide policy, and teachers' perceptions, and inhibit or support the implementation of new assessment practices in selected East Asian countries. All the articles have selected AfL, formative assessment or its derivative as an innovative classroom assessment. Based on the findings, we note there is an ongoing effort in East Asian countries' terms of policy-making, research and professional learning, serving as external mediation to develop or change teachers' AfL practices. These changes in the external domain can precipitate, for instance, changes in teachers' awareness of AfL at the very least.

The evolving rhetoric highlights the problem of drawing definitive parameters regarding what constitutes good AfL, in the context of the difficulties of translating theory into the practice of AfL by teachers, particularly in very different sociocultural contexts, even within East Asia. Partaking in considering the different threads of argument for and against the different conceptualisations brings to bear an understanding of the different practices of AfL in a particular region. Consequently, the practices of AfL can therefore not just be variable; they will also be very situated and contested. We infer from the present set of

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

findings that research on AfL in the East Asian context suffers from a lack of ongoing debates in at least two areas: the current theorisation of AfL, particularly alternative conceptualisations (e.g. Taras, 2009; Carless, 2006); and how 'learning' has (already) taken place efficaciously in the East Asia context based on a theory of pedagogy (Black & Wiliam, 2018) and cultural models (Wang & Li, 2003). While students and teachers in East Asia could be immediately interested in the 'goodness' of the more evident and short-term consequences (e.g. 'good' grades and mastery of a topic) that flow from particular actions of AfL, they are equally concerned with adhering to wider societal fundamentals.

The varying dynamics of how AfL is enacted in East Asian classrooms are therefore dependent not only on external models of principled practices, but also, internally, on an individual's perspectives, values, beliefs and attitudes towards learning, teaching and possibly what constitutes an educated and useful person, as defined by a particular East Asian country context. These culturally specific macro-models of education are less likely to be changed in the foreseeable future.

From the findings in the articles, the current implicit and atheoretical approach to defining and implementing AfL (Baird et al., 2017) suggests opportunities for further deliberation and theorisation about what constitutes AfL in East Asian countries. Arguably, most East Asian teachers can see that all assessment can eventually help students to learn in some way. It is conceivable that teachers who frequently prepare students for summative assessment in East Asian classrooms are concurrently practising a particular process and practice of AfL (Black & Wiliam, 2018; Taras, 2009). Another framework that can be considered is the learning-oriented assessment (LOA) introduced in Hong Kong (Carless, Joughin, & Liu, 2006). LOA seeks to emphasise the important role that assessment plays 'in promoting productive student learning… regardless of any other functions it may be

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

performing'. Within this framework, assessments that are considered 'summative' may also be useful in promoting learning, as they 'equally have a positive influence on what students focus on in their study and how they learn as they prepare for and undertake assessment for grades' (Carless et al., 2006, pp. 7–8). Three main elements are emphasised in LOA to ensure that the focus is on quality of student learning outcomes: a) assessment tasks are designed as learning tasks; b) students are involved in evaluating their work; and c) feedback is always fed forward to the next step in students' learning.

Several of the findings discussed above are not unique to East Asia. The centrality of feedback in AfL has been emphasised in other studies (Heitink, Van der Kleij, Veldkamp, Schildkamp, & Kippers, 2016). The popularity of Black and Wiliam's (1998) seminal article on formative assessment, and particularly their subsequent works on practical implementations of various classroom formative assessment practices, have been highly influential in the current set of literature reviewed. Countries outside this region also face similar challenges to AfL, such as a lack of assessment literacy among teachers and deep-seated beliefs and practices that are resistant to change (De Lisle, 2016; Flórez Petour, 2015). Accountability pressures and the backwash effect of summative assessments are also commonly acknowledged elsewhere (Black & Wiliam, 2005).

While we recognise the commonalities that exist regarding these issues, there are important nuances that make East Asia different. First, the high-stakes nature of its examinations has long-standing cultural importance. To many East Asians, doing well in exams will help increase the social and economic standing of one's family. When a student does well, it is not only for the sake of self-advancement, but also for the economic and social benefit and pride of his/her family, community and even country (Nisbett, 2003). Second, there is the belief that examinations provide an impartial system that allows anyone to

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

succeed regardless of social background (Carless, 2011; Poocharoen & Brillantes, 2013). In the hierarchical and pluralistic societies of East Asia, this perceived fairness is important, as it reduces any tendencies to advance a person based on personal connections or prejudices. Third, examinations and grading are integral to both the education system and life, and teachers and students have accepted it as a valued part of learning and improvement. Thus, high-stakes examinations may not lose their importance in East Asia, even if countries such as Singapore are trying to reverse this trend (Ng, 2017) by introducing a particular conceptualisation of AfL (Ministry of Education Singapore, 2009) and other aggressive policy interventions (Ministry of Education Singapore, 2018).

Then there are other cultural factors that are prevalent in East Asia, such as the highly authoritative, teacher-centred nature of East Asian classrooms. This brings its own challenges when trying to promote learner autonomy in the classroom. Perhaps Biggs' (1996) explanation of how Asian learners learn would be helpful. Whereas in the West children are encouraged to explore and create first, and develop their skills later, teachers in China, for example, focus on helping Chinese learners develop and master basic skills first, before being allowed to explore and create (Biggs, 1996). Thus, it is natural for East Asian classrooms (especially at the primary and secondary level) to be teacher-centred, as teachers feel it is their responsibility to equip their students with what are considered necessary foundational knowledge and skills at the beginning. This does not necessarily mean that only surface learning has taken place (Biggs, 1996) or that students are not central to the teaching (Hui et al., 2017). We should also consider the possibility that AfL has also taken place outside what is observed in the classroom (none of which was explored in the studies reviewed above). Strategies more suitable to East Asian students' learning, such as the use of more written

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

instead of oral–dialogic formative assessment (to draw out misconceptions without embarrassing the students), as suggested by Yin and Buck (2015), should also be considered.

Limitations

There were some analyses that we were unable to conduct, for example, on the learning orientations or underlying theories of learning that the 24 articles have used in their research or conception of AfL. This was because there was not enough information in the articles that allowed us to do so. We also acknowledge and agree with researchers such as Baird et al. (2018) that the relationships between any underlying theories of learning and a particular AfL conceptualisation are currently over-simplified. We envisage that more work needs to be done on further conceptualisation of AfL that is sensitive to social–cultural and situated theories of learning.

Conclusions

We conclude that AfL has to be better understood from a contextually grounded approach in East Asian countries, and that the embedding of new AfL practice (which takes reference mainly from research from outside East Asia) needs to be sensitive to the existing indigenised and idiosyncratic conceptions and practices. We concur with Black and Wiliam's (2005) conclusion that the way forwards lies in finding the 'sweet spots'; and where arguments for a particular version of AfL are least offensive to existing social–cultural values while still being able to catalyse a shift in them, to privilege the well-being of the collective needs of both individual students and society. This is consistent with what researchers have highlighted as the deeply embedded sociocultural codes of assessment that could be highly resistant to change (Kennedy et al., 2008; Roos & Hamilton, 2005). More research is needed on how AfL and summative-oriented assessment practices have (already) been harmonised

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

productively in certain East Asian classrooms. Conceptions of AfL that take into account what 'improving learning' means to East Asians *and* how accountability affects assessment conceptions should also be considered when defining AfL (Hui et al., 2017). This will entail a critical awareness of the social–cultural values underlying assessment and learning relationships, and therefore how further theorising of AfL needs to take place in a more situated context.

Funding

References

Azis, A. (2015). Conceptions and practices of assessment: A case of teachers representing improvement conception. *TEFLIN Journal, 26*(2), 129–154.

Baird, J. A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: fields apart? *Assessment in Education: Principles, Policy & Practice, 24*(3), 317–350.

Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25.

Biggs, J. (1996). Western misinterpretations of the Confucian-Heritage Learning in Culture. In D. A. Watkins & J. B. Biggs (Eds.), *The Chinese learner: cultural, psychological and contextual influences*. Hong Kong: Comparative Education Research Centre; Melbourne, Australia: Australian Council for Educational Research.

Black, H. (1986). Assessment for learning. In D. L. Nuttall (Ed.), *Assessing educational achievement* (pp. 7–18). London, England: Falmer Press.

Black, P., McCormick, R., James, M., & Pedder, D. (2006) Learning how to learn and assessment for learning: a theoretical inquiry. *Research Papers in Education, 21*(2), 119–132.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–73.

Black, P., & Wiliam, D. (2005). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices. *The Curriculum Journal, 16*(2), 249–261.

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 1–25.

Brown, G. T. L., Kennedy, K. J., Fok, P., 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.

Bryant, D. A., & Carless, D. R. (2010). Peer assessment in a test-dominated setting: empowering, boring or facilitating examination preparation? *Educational Research for Policy and Practice, 9*, 3–15.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing, 27*(1), 5–31.

Cagasan, L., Luo, R., Robertson, P. & Care E. (2016). *Formative assessment project: phase 2 research report*. Assessment, Curriculum and Technology Research Centre (ACTRC). Melbourne and Manila. Retrieved from https://minerva-access.unimelb.edu.au/bitstream/handle/11343/129791 /2016%20Care%20et%20al%20-%20Formative%20assessment%20project%20- %20Phase%202%20research%20report.pdf?sequence=1

Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. New York: Routledge.

Carless, D., Joughin, G., & Liu, N.-F. (2006). A conceptual framework for learning-oriented assessment. In D. Carless, G. Joughin & N. F. Liu (Eds.), *How Assessment Supports Learning: Learning-oriented Assessment in Action* [E-book version] (pp. 7–15). doi: 10.5790/hongkong/9789622098237.001.0001

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438–481.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Curriculum Development Council Hong Kong. (2001) *Learning to learn: the way forward in curriculum* [Web edition]. Retrieved from http://www.edb.gov.hk/en/curriculum-development/cs-curriculum-doc-report/wf-in-cur/index.html

Dang, T. T. D. (2016). Vietnamese EFL students' perceptions of noticing-based collaborative feedback on their writing performance. *English Language Teaching, 9*(5), 141–153.

Darling-Hammond, L. (1988). Policy and professionalism. In A. Lieberman (Ed.), *Building a professional culture in schools* (pp. 55–77) New York, NY: Teachers College Press.

De Lisle, J. (2015). The promise and reality of formative assessment practice in a continuous assessment scheme: The case of Trinidad and Tobago. *Assessment in Education: Principles, Policy & Practice, 22*(1), 79–103.

Dunn, K. E., & Mulvenon, S. W. (2009). A Critical Review of Research on Formative Assessment: The Limited Scientific Evidence of the Impact of Formative Assessment in Education. *Practical Assessment, Research & Evaluation, 14*(7), 2.

Evans, J., & Benefield, P. (2001). Systematic Reviews of Educational Research: Does the Medical Model Fit? *British Educational Research Journal, 27*(5), 527–541.

Flórez Petour, M. T. (2015). Systems, ideologies and history: A three-dimensional absence in the study of assessment reform processes. *Assessment in Education: Principles, Policy & Practice, 22*(1), 3–26.

Gardner, J. (2010). Developing teacher assessment: An introduction. In J. Gardner, W. Harlen, L. Hayward & G. Stobart (Eds.), *Developing teacher assessment* (pp. 1–11). Maidenhead, UK: Open University Press.

Griffin, P., Cagasan, L., Care, E., Vista, A., & Nava, F. (2016). Formative assessment policy and its enactment in the Philippines. In D. Laveault & L. K. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation* (Vol. 4, pp. 75–92). Switzerland: Springer International Publishing.

Gu, P. Y. (2014). The unbearable lightness of the curriculum: What drives the assessment practices of a teacher of English as a foreign language in a Chinese secondary school? *Assessment in Education: Principles, Policy & Practice, 21*(3), 286–305.

Hadley, G. (1997). A survey of cultural influences in Japanese ELT (Bulletin of Keiwa College No. 6). Japan: Keiwa College.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

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.

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: software of the mind* (3rd ed.). New York, NY: McGraw-Hill.

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., & 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., 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.

Koh, K., Lim, L., Tan, C., & Habib, M. (2015). Building teachers' capacity in formative assessment: the Singapore example. *The New Educational Review, 40*, 211–221.

Lam, R. (2013). Formative use of summative tests: using test preparation to promote performance and self-regulation. *The Asia-Pacific Education Researcher, 22*, 69–78.

Leong, W. S. (2014). Knowing the intentions, meaning and context of classroom assessment: A case study of Singaporean teacher's conception and practice. *Studies in Educational Evaluation, 43*, 70–78.

Lin, T. J., Lee, M. H., & Tsai, C. C. (2013). The commonalities and dissonances between high-school students' and their science teachers' conceptions of science learning and conceptions of science assessment: a Taiwanese sample study. *International Journal of Science Education, 36*(3), 382–405.

Lomax, R. G., West, M. M., Harmon, M. C., Viator, K. A., & Madaus, G. F. (1995). The impact of mandated standardized testing on minority students. *Journal of Negro Education*, 171–185.

Maxwell, G. S. (2004). *Progressive assessment for learning and certification: Some lessons from school-based assessment in Queensland*. Paper presented at the Third

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

        Conference of the Association of Commonwealth Examination and Assessment Boards, Nadi, Fiji.

McMillan, J. H. (2010). The practical implications of educational aims and contexts for formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 41–58). New York, NY: Routledge.

McNeil, L. M. (1988). Contradictions of control, Part 3: Contradictions of reform. *Phi Delta Kappan, 69*(7), 478–485.

Ministry of Education Brunei Darussalam. (2013). *The national education system for the 21st century.* Retrieved from http://www.moe.gov.bn/spn21dl/SPN21%20ENG%20(2013)%20COMPLETE.pdf

Ministry of Education Malaysia. (2013). *Malaysia education blueprint 2013-2025 (preschool to post-secondary education).* Retrieved from https://www.moe.gov.my/images/dasar-kpm/articlefile_file_003108.pdf

Ministry of Education Singapore. (2009). *Report of the Primary Education Review and Implementation Committee.* Retrieved from http://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/singapore_peri_2009.pdf

Ministry of Education Singapore. (2017). *Assessment concepts in assessment portal.* Abstract retrieved from MOE Singapore intranet website OPAL. Accessed February 1, 2017.

Ministry of Education Singapore (2018). *'Learn for life' – Preparing our students to excel beyond exam results* [Press release]. Retrieved from https://www.moe.gov.sg/news/press-releases/-learn-for-life—preparing-our-students-to-excel-beyond-exam-results.

Ng, C. M. (2017, March 7). *MOE FY 2017 committee of supply debate speech by Minister of Education (Schools), Singapore.* Retrieved from https://www.moe.gov.sg/news/speeches/moe-fy-2017-committee-of-supply-debate-speech-by-minister-of-education-schools-ng-chee-meng

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently... and why.* New York, NY: The Free Press.

Organisation for Economic Co-operation and Development. (2005). *Formative Assessment: Improving learning in secondary classrooms*. [Policy Brief]. Retrieved from http://www.oecd.org/education/ceri/35661078.pdf

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Ong, S. L. (2010). Assessment profile of Malaysia: high-stakes external examinations dominate. *Assessment in Education: Principles, Policy & Practice, 17*(1), 91–103.

Panhoon, S. & Wongwanich, S. (2012). Strategies and consequences of first decade education reforms: the lesson learned from school practices in Chonburi province, Thailand. *International Journal of Learning, 18*(9), 263–288.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: a practical guide.* Oxford, UK: Blackwell Publishing.

Poocharoen, O. O., & Brillantes, A. (2013). Meritocracy in Asia Pacific: Status, issues, and challenges. *Review of Public Personnel Administration, 33*(2), 140–163.

Rashid, R. A., & Jaidin, J. H. (2014). Exploring primary school teachers' conceptions of 'assessment for learning'. *International Education Studies, 7*(9), 69–83.

Ratnam-Lim, C. T. L., & Tan, K. H. K. (2015). Large-scale implementation of formative assessment practices in an examination-oriented culture. *Assessment in Education: Principles, Policy & Practice, 22*(1), 61–78.

Republic of the Philippines Department of Education. (2015). *Policy guidelines on classroom assessment for the K to 12 basic education program*. Retrieved from http://www.deped.gov.ph/sites/default/files/order/2015/DO_s2015_08.pdf

Roos, B., & Hamilton, D. (2005). Formative assessment: A cybernetic viewpoint. *Assessment in Education: Principles, Policy and Practice*, *12*(1), 7–20.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144.

Sardareh, S. A. (2016). Formative feedback in a Malaysian primary school ESL context. *Malaysian Online Journal of Educational Sciences, 4*(1), 1–8.

Sardareh, S. A., Saad, M. R. M., Othman, A. J., & Me, R. C. (2014). ESL teachers' questioning technique in an assessment for learning context: promising or problematic? *International Education Studies, 7*(9), 161–174.

Scriven, M. (1967). The methodology of evaluation. In R. Tyler, R. Gagné & M. Scriven (Eds.), *Perspectives of Curriculum Evaluation* (AERA Monograph Series on Curriculum Evaluation, Vol. 1, pp. 39–83). Chicago: Rand McNally.

Stiggins, R. J. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *The Phi Delta Kappan, 87*(4), 324–328.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice, 18*(4), 433–449.

Tan, C. (2016). Tensions and challenges in China's education policy borrowing. *Educational Research, 58*(2), 195–206.

Taras, M. (2005). Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies, 53*(4), 466–478.

Taras, M. (2009). Summative assessment: the missing link for formative assessment. *Journal of Further and Higher Education, 33*(1), 57–69.

Thanh, P. T. H. (2014). *Implementing cross-culture pedagogies: cooperative learning at Confucian heritage cultures*. Singapore: Springer.

Thanh Pham, T. H., & Renshaw, P. (2015). Formative assessment in Confucian heritage culture classrooms: activity theory analysis of tensions, contradictions and hybrid practices. *Assessment & Evaluation in Higher Education, 40*(1), 45–59.

Tong, S.Y. (2011). Assessing English language arts in Hong Kong secondary schools. *The Asia-Pacific Education Researcher, 20*(2), 387–394.

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom.* UK: McGraw-Hill Education.

United Nations Educational, Scientific and Cultural Organization. (2015). *EFA Global Monitoring Report Regional Overview: East Asia and the Pacific*. Retrieved from http://en.unesco.org/gem-report/sites/gem-report/files/regional_overview_EAP_en.pdf

Wang, Q., & Li, J. (2003). Chinese children's self-concepts in the domains of learning and social relations. *Psychology in the Schools, 40*(1), 85–101.

Weed, M. (2005). 'Meta Interpretation': A method for the interpretive synthesis of qualitative research. *Forum: Qualitative Social Research*, *6*(1), Art. 37. Retrieved from http://nbn-resolving.de/urn:nbn:de:0114-fqs0501375.

Weed, M. (2008). A potential method for the interpretive synthesis of qualitative research: issues in the development of 'meta-interpretation.' *International Journal of Social Research Methodology, 11*(1), 13–28.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3–14.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education, 45*, 128–136.

Yatab, R. S., & Shahrill, M. (2014). The differing views in using the common assessment tasks in secondary school science. *International Journal of Science and Research, 3*(7), 685–693.

Yin, X. & Buck, G. A. (2015). There is another choice: an exploration of integrating formative assessment in a Chinese high school chemistry classroom through collaborative action research. *Cultural Studies of Science Education, 10*(3), 719–752.

Yu, W. M. (2015). Teacher leaders' perceptions and practice of student assessment reform in Hong Kong: a case study. *Planning and Changing, 46*(1/2), 175–192.

Zhao, X., Van den Heuvel-Panhuizen, M., & Veldhuis, M. (2016). Teachers' use of classroom assessment techniques in primary mathematics education – an explorative study with six Chinese teachers. *International Journal of STEM Education, 3*, 1–18.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

Table 1. Overview of the 24 articles analysed

No.	Source	Setting	Research design	Data collected	Sample	Research focus	Research objectives
1	Azis, A. (2015)	Indonesia; secondary school	MM	Interviews, questionnaires	107 teachers	Assessment, including AfL	Examine teachers' conceptions of assessment and how they relate to their assessment practices.
2	Brown, G. T. L., Kennedy, K. J., Fok, P., Chan, J. K. S., & Yu, W. M. (2009)	Hong Kong; primary school	Quanti.; survey	Questionnaires	288 teachers	Assessment, including AfL	Validate instrument in Hong Kong context and examine linkages between conceptions and practices of teachers.
3	Bryant, D. A. & Carless, D. R. (2010)	Hong Kong; primary school	Quali.	Classroom observations, teacher and student group interviews	34 students, 2 teachers	Peer assessment	Examine teachers' and students' perceptions of peer assessment, and the tensions and opportunities in its implementation.
4	Butler, Y. G. & Lee, J. (2010)	S. Korea; primary school	MM; experimental design	Interviews, surveys, pre- and post-test scores	2 teachers, 254 students	Self-assessment	Examine effectiveness of self-assessment among young learners in FLES programme.
5	Cagasan, L., Luo, R., Robertson, P., & Care E. (2016)	Philippines; primary and secondary school	Quali.	Observations, lesson narratives	65 teachers	AfL in general	Explore the implementation of formative assessment in the Philippine classroom.
6	Griffin, P., Cagasan, L.,	Philippines; primary	Quali.	Observations, interviews	61 lessons, no. of	AfL in general	Study the link between assessment and teaching.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, 59, 270-277.

	Care, E., Vista, A., & Nava, F. (2016)	and secondary school			teachers not mentioned		
7	Gu, P. Y. (2014)	China; secondary school	Quali.; case study	Video-recorded observations, interview	1 teacher	AfL in general	Explore how policy is translated into practice by analysing the place of assessment in one unit of teaching.
8	Hui, S. K. F., Brown, G. T., & Chan, S. W. M. (2017)	Hong Kong; primary school	Quali.	Interviews, assessment tasks	4 curriculum leaders	AfL in general	Investigate in-depth the conceptions of assessment in the Chinese context not covered by the Teachers' Conceptions of Assessment (TCoA) questionnaire.
9	Hui, S. K. F. (2012)	Hong Kong; primary school	Quali.	Interviews	5 principals and curriculum leaders	Assessment, including AfL	Examine the types of assessment task used, the nature of AfL in the Hong Kong primary school context and views of how the tasks could support learning improvement in the accountability era.
10	Koh, K., Lim, L., Tan, C., & Habib, M. (2015)	Singapore; secondary school	Quali.; case study	FGD	3 teachers	AfL in general	Explores teachers' experiences with and perceptions of formative assessment.
11	Lam, R. (2013)	Hong Kong; primary school	Quali.; case study	Observations, interviews	2 teachers, 14 students	Assessment, including AfL	Investigate teachers' and students' perspectives of how 'the formative use of summative test' (i.e. test preparation) may impact a student's performance in internal school tests and develop self-regulated learning.
12	Leong, W. S. (2014)	Singapore; primary school	Quali.; case study	Observations, interviews	1 teacher	Assessment, including AfL	Illustrate how the research involved one case-study teacher in some form of 'practical reasoning', verbalising and

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

							enacting espoused conceptions and practices of classroom assessment.
13	Lin, T. J., Lee, M. H., & Tsai, C. C. (2013)	Taiwan; secondary school	Quanti.; survey	Questionnaires	59 teachers, 1048 students	Assessment, including AfL	Investigate high school students' and their science teachers' conceptions of learning science (COLS) and conceptions of science assessment (COSA).
14	Rashid, R. A., & Jaidin, J. H. (2014)	Brunei; primary school	Quali.	Interviews	15 teachers	AfL in general	Explore ways in which assessment for learning (SBAfL) is understood and implemented in government primary schools.
15	Ratnam-Lim, C. T. L., & Tan, K. H. K. (2015)	Singapore; primary school	Quanti.; survey	Open-ended questionnaires	30 teachers, 13 parents	AfL in general	Explore perceptions and experiences of teachers and parents on 'holistic assessment'.
16	Sardareh, S. A. (2016)	Malaysia; primary school	Quali.	Observations, FGD	3 teachers	Feedback	Investigate how formative feedback is being implemented.
17	Sardareh, S. A., Saad, M. M., Othman, A. J., & Che Me, R. (2014)	Malaysia; primary school	Quali.	Observations, interviews	3 teachers	Questioning	Examine the process of classroom questioning during AfL in the primary school ESL context.
18	Tan, C. (2016)	China; primary and secondary school	Quali.; survey	Open-ended questionnaires, interviews	119 principals, vice principals and teachers; 47 interviewed	Assessment, including AfL	Examine education policy borrowing in mainland China through the new curriculum reform (NCR).

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.

19	Tong, S.Y. (2011)	Hong Kong; secondary school	Quali.; case study	Observations, interviews	14 teachers	Performance-based assessment	Examine various assessment tasks and their criteria, student language arts (LA) work and the feedback received.
20	Yan, Z., & Cheng, E. C. K. (2015)	Hong Kong; primary school	Quanti.; survey	Questionnaires	450 teachers	AfL in general	Explore the relationships among teachers' attitudes, intentions and practices regarding formative assessment under the framework of the Theory of Planned Behaviour.
21	Yatab, R. S., & Shahrill, M. (2014)	Brunei; secondary school	MM	Questionnaires, artefacts	33 students, 4 teachers	AfL in general	Investigate the effectiveness of BCATs in lower secondary science.
22	Yin, X. & Buck, G. A. (2015)	China; secondary school	Quali.	Interviews, artefacts, observations, open-ended questionnaires	1 teacher, 48 students	AfL in general	Explore integrating formative assessment to a Chinese high school chemistry classroom through a collaborative action research.
23	Yu, W. M. (2015)	Hong Kong; primary school	Quali.	Interviews, observations, artefacts	16 teacher leaders	Assessment, including AfL	Explore teachers' perceptions of student assessment reform in classroom practice.
24	Zhao, X., Van den Heuvel-Panhuizen, M., & Veldhuis, M. (2016). (2016)	China; primary school	Quali.	Interviews, feedback forms, observations, artefacts	6 teachers	Assessment, including AfL	Report on the use of classroom assessment techniques (CATs) by primary school mathematics teachers.

Leong, W. S., Ismail, H., Costa, J. S., & Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation*, *59*, 270-277.