
Title	Understanding the interaction of assessment, learning and context: Insights from Singapore
Author(s)	Tang T. Heng, Lynn Song and Kelvin H.K. Tan
Source	<i>Educational Research</i> , (2021)

Copyright © 2021 Taylor & Francis

This is an Accepted Manuscript of an article published by Taylor & Francis in *Educational Research* on 22/01/2021, available online:

<http://www.tandfonline.com/doi/pdf/10.1080/00131881.2021.1874248>

Understanding the interaction of assessment, learning and context: Insights from Singapore

Tang T. Heng*, Lynn Song and Kelvin H.K. Tan

Nanyang Technological University-National Institute of Education, Singapore

* Corresponding author: tangtang.heng@nie.edu.sg

To cite this article: Heng, T. T., Song, L., & Tan, K.H.K. (2021). Understanding the interaction of assessment, learning and context: Insights from Singapore. *Educational Research*.
<https://doi.org/10.1080/00131881.2021.1874248>

Abstract

Background: Assessment for Learning (AfL) discourses and practices have gained popularity globally, as educational policies and ideas are transferred across borders. An important area of investigation is how AfL may be construed and enacted differently, according to context.

Purpose: By examining how teachers in Singapore understand and use AfL—a core principle embedded within differentiated instruction—this study explores how context influences assessment, learning and teaching beliefs and practices.

Method: As part of a larger study of how 10 public school teachers in Singapore understand and use differentiated instruction, the research reported here adopted a qualitative approach to elucidate teachers' lived experiences. Data, which were analysed thematically, included 10 questionnaires, 30 semi-structured interviews, 39 lesson observations logs, and pre-and post-lesson observation interviews.

Findings: The analysis identified three themes: (i) teaching and learning for Assessment of Learning (AoL) rather than AfL, (ii) teacher-directed rather than learner-driven AfL practices, and (iii) behaviourist approaches in teaching and assessment practices. Finding suggested that teachers in Singapore face tensions as personal beliefs and practices around teaching, learning, and assessment collide and/or coexist with their sociocultural and educational contexts.

Conclusions: Rather than assuming that educational ideas like AfL are acontextual, or represent a 'best practice' that transfers easily across countries, scholars, policy-makers, and educators would benefit from considering more fully how AfL is a 'situated' concept, and reflecting more deeply on the complex interplay between source and destination contexts.

Keywords: assessment for learning (AfL), formative assessment, learning theory, differentiated instruction, educational policy borrowing, Singapore

Introduction

Assessment for learning (AfL) discourses and practices have been popularised in Europe, Canada, and the United States, as well as in Asia Pacific countries (Klenowski 2009), reflecting the speed and extent that educational policies or ideas are borrowed or transferred across borders. However, scholars contend that, beyond studying the types and trajectories of education policies transferred, there is a need to scrutinise education within the sociocultural, economic, and educational contexts of both source and destination countries (Heng and Song 2020, Steiner-Khamsi 2014, Vavrus and Bartlett 2012). Only then can we begin to understand how assessment, learning and teaching are construed differently across national contexts (Brown et al. 2009, Leong et al. 2018), and therefore gain insight into how AfL may be construed and enacted differently, according to context. To this end, this article reports on a study that sought to examine how teachers in Singapore understand and use AfL.

Background

Assessment for Learning, and Learning in Assessment for Learning

A full review of AfL is not within the scope of our study: here, we offer a background discussion in order to situate our work. Black and Wiliam's (1998) seminal study on formative assessment foregrounded that assessment needs to serve the purpose of improving students' learning, by using evidence to adjust instruction to meet student needs. By contrast, in summative assessment or assessment of learning (AoL), evidence is used for the purpose of determining if learning has taken place, and focus is placed on outcomes rather than process. The phrases 'formative assessment' and 'AfL' have since been used interchangeably worldwide

(Klenowski 2009) and debates continue about definitions (Jönsson 2020). A commonly-used definition of AfL is ‘the process of seeking and interpreting evidence for use by students and their teachers to decide where the students are in their learning, where they need to go, and how best to get there’ (Broadfoot et al. 2002, 2-3). While various definitions have since been proposed (Bennett 2011, Klenowski 2009), overlapping characteristics are evident. For one, intentionality and clarity are needed by both teachers and students in identifying learning goals, where students stand in relation to the goals and how they will get there (Broadfoot et al. 2002, Wiliam 2011). Next, the ongoing assessment evidence collected has to be specific enough to determine the next course of instructional action by revealing the nature of the gap in students’ learning. Also, teachers need to provide constructive guidance to develop students’ capacity for self-assessment and self-directed improvement (Broadfoot et al. 2002). Emphasis is placed on the student’s role in assessment—for example, they are expected to play an active role in using feedback (from teachers and peers) and self-assessment to move their learning forward and become more independent learners (Bennett 2011).

Notwithstanding its widespread use and appeal (Klenowski 2009, Taras 2010), AfL and its applications has been subjected to numerous critiques. For a start, implementation and/or approach can be problematic, with educators sometimes reducing AfL to a series of steps focusing on procedures (Baird et al. 2017), using AfL for external validation (Torrance and Pryor 1998), or implementing ‘assessment as learning’ instead of for learning (Hume and Coll 2009, 287). Others point to a lack of engagement with assessment theories (Taras 2010), definitional ambiguity affecting implementation (Tan 2017), and inconclusive effectiveness (Bennett 2011). There has been much scholarly discussion about the challenges of reconciling theories of learning with theories of assessment (Baird et al. 2017, Bennett 2011, Broadfoot 2017, Goldstein

2017). The under-elaboration of ‘learning’—what it looks like, how it takes place, under what conditions—creates ambiguity in how assessment data should be used. Broadly, different theories of learning may exert disparate implications on the purpose and, consequently, practices of assessment (Baird et al. 2017). Engaging with learning theories is critical not simply because of the affiliation with assessment, but because of the interaction with teaching, reflecting a triarchic relationship across learning, teaching, and assessment (Torrance and Pryor 1998). Conceptions of learning affect assessment practices, including how collected assessment data is fed forward into teaching. For example, behaviourist views are associated with teaching strategies that stimulate or reinforce (e.g. drill and practice, and the reward or deduction of participant points) and socio-constructivist views with practices that draw from students (e.g. case studies, collaborations, and discovery learning) (Bransford et al. 2000, Watkins 2003). However, learning theory is not the sole driver for one’s assessment and teaching practices; rather, there is a complex and multi-directional relationship across these elements, further compounded by the pressures of the larger context in which it is set (Jiang 2015, Torrance and Pryor 1998).

Tensions in assessment practices may be especially evident in educational systems that are traditionally associated with high-stakes standardised tests (Brown et al. 2009, Deneen et al. 2019, Ratnam-Lim and Tan 2015) and that are under pressure to perform in international tests (Hume and Coll 2009). To add to the complexity, conceptions of learning (and, relatedly, AfL) differ across sociocultural and educational contexts. For example, it has been suggested that teachers in Asian contexts, where examinations are very much a part of the schooling culture, can face significant barriers understanding and implementing AfL (Brown et al. 2009, Leong et al. 2018, Ratnam-Lim and Tan 2015). As Alexander’s (2001) five-nation comparative study

revealed, learning, teaching, and assessment are shaped by the geography, history, and culture of a country. Implementing education change within a single country is no mean feat to begin with; transferring educational ideas across contexts is even more complex, and can create tensions in respect of conceptions of teaching and learning between source and destination contexts (Heng and Song 2020). It is noteworthy, though, that discourses about AfL rarely foreground the criticality of contexts, particularly in terms of how learning and teaching are seen and operationalised in tandem with assessment.

Assessment for Learning and Differentiated Instruction in Singapore

This article is derived from a larger study investigating teachers' understanding and practices of differentiation in Singapore. In recent years, differentiated instruction has gained prominence in Singapore, as educators seek new educational approaches to respond to increasing diversity in the classrooms (Heng and Song 2020). Differentiated instruction is an educational approach where teachers intentionally adjust curricula, learning, teaching, assessment, and routines to address the range of students' readiness levels, interests, and learning profiles (Tomlinson 2001). Amongst the five core principles, two relate to assessment: assessment informs teaching and learning, and instruction responds to student needs indicated by formative assessment. The triarchic relationship across assessment, teaching, and learning bears out in the former principle, with AfL in the latter (Tomlinson and Moon 2013). Assessment in differentiated instruction therefore offers fertile ground to explore teachers' understanding and use of AfL.

Tomlinson and Moon (2013) define AfL, or formative assessment¹, as ‘gathering, learning from, and using information about student needs to adjust teaching and learning plans in ways that promise to contribute to student success’ (p. 60). They distil AfL into two categories: pre-assessment (administered before the start of a unit) and ongoing (or formative) assessment (administered throughout a unit or units). Both serve the purpose of determining how instruction should be adjusted to support students’ learning and growth in light of curricular goals, and how it should ‘happen *with* students, not *to* them’ (p. 64). Beyond readiness information (i.e., students’ proximity to knowledge, understanding, and skills), Tomlinson and Moon (2013) urged teachers to collect other types of information—e.g., student interest or learning preferences—to paint a holistic picture of how students can be invited into the curriculum. These principles of assessment closely resonate with the views that AfL should champion socio-constructivist theories of learning (Black and Wiliam 1998, James 2006, Shepard 2000, Torrance and Pryor 1998).

However, Singapore’s education system is frequently regarded as examination-oriented, where national examinations perform a purportedly objective and meritocratic sorting mechanism that inadvertently washes back into primary and secondary schools (Deneen et al. 2019, Ratnam-Lim and Tan 2015, Tay et al. 2020). Some observed effects include a curricular emphasis on content acquisition and grades, rather than learning and holistic development (Hogan et al. 2013), and the restriction of AfL usage to examinable subjects for increasing scores rather than for nurturing learner agency and development (Deneen et al. 2019). Although teachers value involving students as active learners and assessors, they expressed low

¹ For the purpose of this paper, we use formative assessment in differentiated instruction interchangeably with assessment for learning.

proficiency or lack of opportunity to practice it (Deneen et al. 2019). Scholars have also attributed the lower emphasis on learner involvement to the influence of Confucian-heritage-culture on teaching and assessment (Heng and Song 2020, Leong et al. 2018). Additionally, the high stakes of examination results have become a defacto authority and institution of assessment purpose in schools, which has resulted in the risk of focusing on standards of performativity of teaching and learning among school leaders (Leong and Tan 2014, Tan 2017).

Such contexts are far removed from the learning environments originally envisaged by AfL and differentiated instruction, highlighting the need to foreground not just the issue of context, but how context interacts with perspectives of learning, and the latter's relationship with AfL. Unique contextual differences challenge assumptions that 'best practices' are a 'globalised form of knowledge' (Vavrus and Bartlett 2012) that can be easily borrowed from a source country and transferred to another destination country. Existing studies highlight the incompatibility of AfL with standardised examination systems, but insufficient attention has been paid to investigating the nuances of learning conceptions in terms of context. Further, teachers' voices within this warrant closer scrutiny, as, in order to gain much-needed insight into classroom practices, there is a need for in-depth studies privileging their voices and lived realities.

Purpose

Against the background set out above, this study sought to examine how a small sample of teachers in Singapore understand and use AfL, exploring how context influences assessment, and learning and teaching beliefs and practices. The following research questions were asked:

What are teachers' understandings of AfL when enacting differentiated instruction? How are teachers implementing AfL when enacting differentiated instruction? How does context interact with teachers' understanding and implementation of AfL?

Method

This study adopted a qualitative approach. As we were concerned with capturing participants' lived experiences, we observed participants in a naturalistic classroom setting, via the conducting of lesson observations, and also interviewed them in order to co-construct a nuanced picture of their understanding and use of assessment (Hatch 2002). Additionally, we integrated a secondary approach into our methodology by 'embedding' (Clark and Ivankova 2017, 140) consensus ratings of the lesson observations into our research design. We presented patterns from our lesson observations to participants during an interview, and elicited their responses so as to member check and enhance the co-construction of knowledge.

Ethical Considerations

Upon approval by the Research Integrity and Ethics Office of our institution as well as the Ministry of Education, we conducted the study according to the guidelines stipulated. We recruited participants from public schools through our network, with participants nominated by colleagues or school principals. All participants were enlisted on a voluntary basis and signed informed consents. Participants were reassured that they could withdraw from the study at any time without consequences and were given details on possible risks, benefits, anonymity, confidentiality, and data security associated with the research process. In accordance with our institution's requirements, only the principal investigators and research associate were involved in data collection and analysis. To ensure participants' anonymity, we assigned them gender-neutral pseudonyms in the reporting.

Participants

Ten public school teachers, one per school, were recruited by criterion sampling in terms of i. gender (including female and male), ii. including core teaching subjects: English (four),

Mathematics (four) and Science (four), iii. including different levels: primary (five) and secondary (five), and iv. participant teaching experience (ranging from 2.5 to 23 years). The sampling was indicative of the Singapore teaching service in terms of gender and level taught (Department of Statistics Singapore 2019, MOE 2019).

Data Collection

At the start of the study, we administered an online questionnaire to elicit background information (e.g., number of classes/students taught). The study was conducted in English. We conducted 39 lesson observation cycles, with each comprising the lesson observation(s) (LO) and a pre-and post-lesson observation interview (LO#Intv). We completed two observation cycles per topic, so that we could observe how participants adjusted instruction. All participants, except for one, completed four LO cycles. At the pre-LO interview, we asked participants about their lesson intentions and concerns (e.g. What are the lesson objectives? What student data were used to inform the lesson design?). During the LO, the first and second author kept separate logs that detailed minute-by-minute observations of participants' practices. At the post-LO interview, we asked about their lesson reflections, concerns, and intended follow-up actions (e.g. What instructional objectives were met? How did you know? What do you intend to do for subsequent lessons?). The observation logs and pre-/post-LO interview data were used to complete a consensus observation rating table for each LO, with scores averaged across 39 LO cycles (Table 1). Included with the consensus ratings was qualitative supporting evidence extracted from our observation logs. We also conducted three semi-structured interviews. The first, a 60-minute beginning-of-study interview (Intv1), asked about participants' perspectives about differentiated instruction, teaching, assessment, and learning (e.g. What makes a good teacher? What do you think is the purpose of schooling?). The second, a 90-minute end-of-first-year interview (Intv2),

conducted upon completion of 39 LOs, asked participants about their views on LO patterns observed (e.g. What are your thoughts regarding why teachers often collect information about readiness, rarely about interest, learner, and learning profile?). The third, a 90-minute end-of-second-year interview (Intv3) sought to understand how their implementation changed (e.g. How did your understanding and implementation of differentiated instruction changed over time?). All interviews were audio-recorded and transcribed verbatim. Finally, after every two LO cycles, the first author kept extended memos detailing thoughts on the data, theory, and method (Miles and Huberman 1984), yielding twenty memos.

[Insert Table 1]

Data Analysis

Preliminary data analysis took place during data collection. After each cycle, the first and second author separately reviewed data from LO and interviews to rate participants' practices, on a scale of 0 (no evidence) to 3 (strong evidence), against descriptors based on the five principles of differentiated instruction. The first and second author subsequently derived a consensus score after discussing their observations of qualitative data. The average scores of the ratings across 39 LO cycles were compiled into Table 1. Our qualitative observations, supplemented with patterns from this table, were shared with participants a few days in advance of Intv2 to verify our observations, as well as to elicit their responses.

Upon completion of data collection, we started by inductively coding qualitative data that addressed assessment in the transcribed interviews (Hatch 2002). We coded recurring ideas such as 'high-stakes exam', 'scoring well', 'remediation', and 'learning gaps'. Subsequently, we

coded other data segments that addressed teaching and learning, and identified emerging ideas such as teacher as ‘expert’, ‘empty vessels’, ‘completing the syllabus’, and ‘meritocracy’. These ideas were then sorted according to their relationships; e.g. X is a kind of, or reason for, Y (LeCompte 2000). For instance, ‘high-stakes exam’ and ‘scoring well’ were examples related to ‘teachers focus on teaching for assessment of learning’, while participants used ‘meritocracy’ to explain why. To strengthen the credibility of our analysis, we compared patterns from qualitative data with the patterns from the rating table and extended memos. Additionally, we used iterative questioning during the interviews to detect inconsistencies. To increase the trustworthiness of our analysis, we conducted member checking, by inviting participants to verify our observations, as well as peer scrutiny, with the first author checking the second author’s coding (Shenton 2004).

Findings and discussion

Our analysis identified three main themes: (i) teaching and learning for AoL not AfL; (ii) teacher-directed (not learner-driven) AfL practices; and (iii) behaviourist approaches in teaching and assessment practices. Selected noteworthy findings from these themes are presented in the subsections below; where relevant, these are illustrated with anonymised quotations from the data. All names are pseudonyms.

(i) Teaching and Learning for Assessment of Learning (not Assessment for Learning)

Firstly, across LOs and interviews, our findings suggested that teachers’ teaching and learning practices were heavily focused on preparing students for AoL rather than AfL. In many lessons, participants were explicit in detailing examination requirements, with discussions of

learning frequently centring on examination goals and ways to attain the highest score. As one participant explained, their ‘training of them [students] would be to talk about the scores’.

Participants also focused on teaching examination techniques, such as the use of checklists and precise answer presentation to reflect procedural accuracy. The purpose and process of providing feedback to students during Science lessons is evident in Sage’s description of teaching process:

...during [error] correction I will highlight the techniques to answer the question.... Let's say I teach graphs. I will write T slab. The graph must have title, the scale must be correct: the line, the axis, and then the points must be plotted correctly. Then I teach the student what it means, then they have to do a self-check. I tell them that it is to ensure they score the five full marks for their graph drawing.

The purpose of feedback, as in this case, is for students to acquire requisite examination ‘techniques’ to answer graph-related questions, rather than the development of conceptual understanding behind graphing. This focus on examination scores and procedures also surfaced in Language lessons; for example, one participant commented that they teach students to be ‘safe in choosing topics [during exams]’. Many perceived an expectation that a good teacher is measured by students’ examination results: as Mak reflected, ‘as a teacher within the school, what makes a good teacher - I think the emphasis is still very much results-based’.

Secondly, participants argued that they were responding to how their students, colleagues, and the wider community view learning: learning was frequently equated to assessment, which, in turn, equated to progressing to the next (educational) phase. However, many participants noted that, as they gained experience teaching, they no longer ‘only chase exam scores’ or remained ‘results-driven’; they continued to struggle to reach the ‘right balance’. Although they expressed belief in more holistic perspectives of learning — for example, ‘understanding as learning’ and learning as ‘understand[ing] more about the world around us ...

for lifelong learning’ — it appeared that they wrestled with society’s instrumental view of learning.

Thirdly, learning was not simply equated to assessment, but to having attained good assessment scores. For example, Micah remarked that teachers often ‘respond to students based on the number we have, the mark that we have given them’. Participants also observed that equating learning with scores was prevalent among students, parents and society, where a good score serves, as one participant put it, as ‘a validation of their [students’] supposed ability or hard work’ and even self-worth. Another echoed this: ‘[students] are equating their achievement standard to their whole academic self-worth...therefore I’m [they’re] an abled learner because they tie it to a number’.

Participants suggested practical reasons for the strong exam focus. For example, Ashley explained that ‘national exams... have to sift out people for different jobs that are integral to allowing the society to function...[with] systems in place to allow people to work on their merit’, whilst Poh observed that the society ‘values meritocracy... [that] academic success equals better chances of getting a good choice’ in life. Some participants, including Poh, noted the ‘pressure that they [students] feel from different stakeholders’ encouraged students to become exam focused: for instance, Micah observed that parents prioritised ‘how he [their child] can get the A1’. Pressure was also mentioned in relation to teachers; for example Jude argued, ‘nobody wants to be known as the teacher who cannot procure results’ since ‘teachers are professionally assessed in part based on their students’ grades’.

(ii) Teacher-Directed (not Learner-Driven) Assessment for Learning Practices

When teachers did implement AfL, they, rather than their students, tended to determine the frequency and type of assessment as well as instructional pacing. We observed that, prior to

teaching a unit, participants demonstrated weak evidence conducting AfL (0.6 out of 3)—or what Tomlinson and Moon (2013) termed ‘pre-assessment’—to understand students’ prior knowledge about the upcoming unit. Participants explained that constant AfL was not necessary; as one put it, teachers ‘would have an idea of the readiness of our students, based on our assessment of their daily work and daily interaction’. So ‘at the back of my [teachers’] minds, we [they] know which students are good, which are weak’.

During instruction, participants demonstrated moderate evidence incorporating in-unit AfL (1.9 out of 3), or ‘formative assessment’ (Tomlinson and Moon 2013). However, we observed that participants showed weaker evidence adjusting instruction (1.2 out of 3) based on AfL data. When we probed participants on how they planned to modify or have modified instruction based on AfL data, they frequently remarked that minimal changes were necessary; however, analysis of lesson observation data did not tend to support this.

Responding to our observation that there was weaker evidence in terms of adjusting instruction based on AfL data, participants pointed out their challenges with a high teacher-student ratio. They explained that it was hard to adjust instruction, as large class sizes of around 40 rendered conducting AfL a ‘struggle’ to begin with. For example, Jude described the challenges and affordances of class size:

With 31 students right now, I can do my formative assessment more easily.... I can assess their written responses faster, and because there are fewer students, I can also require them to write more... I can go around to these 31 students, probably not every single one... But with 41, it’s quite impossible’.

Additionally, participants considered that there was a perceived duty to teach regardless of AfL data because teachers felt that, as one explained, ‘no matter their [students’] level of readiness, no matter how much they know, they will all sit for the same exam’. In the same vein, another

noted that teaching regardless of AfL data was, in a sense, a way of being responsible to all students:

Whether the child has learnt it or not, we have to teach the topic, because there will be children who have not gone through the topic..., it is not fair just because 99% of the class have gone through the topic for me to skip it because that one child who does not have any support at home is going to suffer.

Finally, we observed that participants demonstrated low evidence in promoting learner independence by encouraging self-directed learning and self-assessment during their lessons (1.0 out of 3), suggesting that epistemic authority remained with teachers. The excerpt below outlines a segment in Mak's first lesson observation, and is reflective of lessons taught by other participants, where students often waited for teachers to provide feedback, rather than check with peers, or self-evaluate.

Figure 1: Excerpt from a lesson observation

Time	Detail
0830	Students are working on their tasks. Teacher checks on a group of students. Student A raises hand. Teacher walks over, answers Student A's question. Teacher walks to another pair. Teacher marks each student's work and explains to students.
0835	[This continues with four more pairs of students.]
0843	Student C waves her worksheet over her head to get the teacher's attention. She puts it down after a while and gives up. Teacher continues working with Student B going through the question step by step.
0845	Student E raises hands to call for teacher's attention. Student F calls out. Teacher addresses Student E's question and moves to Student F's group.

Participants explained that the predominant culture of learning is one where there was, in the words of one, a 'mindset that the teacher is the expert' : in the words of another, the teacher represented 'the sole source of power in the classroom, [rendering the student] subordinate to the teacher' In essence, as one observed, 'we don't have this idea that the student is playing a role and the students are involved in co-teaching each other in the classroom'.

(iii) Behaviourist Approach in Teaching and Assessment Practices

Our analysis indicated that teachers tended to perceive learning as a behaviour that needs to be managed or trained. This suggests some alignment with a behaviourist perspective on teaching and learning, which assumes that learning is knowledge acquisition that occurs in response to stimuli (see further Bransford et al. 2000; Watkins 2003). For example, participants felt that teachers believed that students need to be ‘train[ed]’ ‘...for classroom routines’. It is interesting to note that, according to our lesson observation analysis, participants demonstrated moderate to strong evidence in classroom management (2.0 out of 3). Some of the strategies to establish classroom routines that were observed included: clapping signals (one silent clap), chants (teacher: ‘eyes on me’; student: ‘eyes on you’) and rewards (such as early recess and point system, as evident in the excerpt below).

[Teacher to class] If you are ready for this lesson, you should be sitting up straight. (Students start to settle down. Some continue to talk and drink water.) I see that Student A is ready for my lesson, I’m going to give student A’s group 1000 points.

While some participants disagreed with such approaches, they noticed, as one participant put it, that ‘rewards’ like ‘a trip, a gift, a promise of something’ were often used by parents to motivate students to do well for assessments, so much so that students ‘have been socialised’ into this way of thinking and behaviour, making it hard for teachers to undo.

Another observation was that participants appeared to see learning as unidirectional, with teachers determining what ought to be learnt, as opposed to co-constructing learning with students. During lesson observations, it seemed to be the case from our analysis that participants rarely helped students draw links between what they learn and real life, scoring low (1 out of 3) in ensuring that the curriculum was authentic and relevant to them. Similarly, participants rarely tailored instruction to student interest (0.6 out of 3). When we elicited possible reasons for these

observations, participants suggested that teachers perceived students as ‘empty vessels’ to be filled and are ‘very focused on teaching the subject...not teaching the child’.

Participants noted that teachers were always looking out for, in Mak’s words, ‘learning gaps’ to be closed or bridged’ in their assessment practices. It was interesting that participants also discussed their own learning needs in terms of deficits — for example, one commented that teachers ‘don’t have much training’, and another felt that there was a need to ‘rigorously equip staff’ — revealing a sense that learning was viewed as unidirectional.

Further discussion, limitations and implications

Of course, there are inevitably limitations to our small scale, qualitative study. Given the small sample size and design of the study, there is no intention to claim that these findings are representative of teaching, learning and assessment practices of Singapore teachers. Instead, through in-depth analysis of our interview and lesson observation data, we aim to paint rich descriptions of these teachers’ beliefs and practices, interpreted within the sociocultural and educational context. That our findings indicate these ten Singapore teachers were more inclined to assessment *of* learning—not *for* learning—practices is unsurprising, as it echoes other scholars’ observations (Deneen et al. 2019, Hogan et al. 2013, Ratnam-Lim and Tan 2015).

What this study illuminates, however, are some of the tensions teachers face as their personal beliefs and practices regarding teaching, learning, and assessment collide and/or coexist with their sociocultural and educational contexts: for example, teachers’ attempts to reposition colleagues’, parents’, and students’ behaviouristic attitudes towards learning. The sense that students’ self-worth and teachers’ professional reputation were equated to examination results further dampened attempts to decouple learning from examinations: some teachers felt that emphasising AoL over AfL was justifiable, and was the ‘responsible’ thing to do. This study thus

provides an interesting connection with existing studies describing AfL resistance (e.g. Hume and Coll 2009).

The Singaporean sociocultural and educational context shaped the teachers' AfL understanding and practices in other ways. As opposed to student-driven practices privileged in AfL scholarship (Bennett 2011, Broadfoot et al. 2002, Klenowski 2009), we observed strong teacher-directedness and low student-driven teaching and assessment practices, resonating with findings from previous quantitative studies (Deneen et al. 2019, Hogan et al. 2013) but also underscoring teachers' interpretations. Teachers explained that the low evidence in supporting self-directed learning and self-assessment stemmed from assumptions that students were 'subordinate' to the teacher and were 'too young', suggesting hierarchical perspectives of teaching and learning associated with Confucian-heritage or Asian societies (Heng and Song 2020, Brown et al. 2009, Leong et al. 2018). Further, teachers raised the issue of what was considered to be a high teacher-student ratio in Singapore, arguing that large class sizes and subject specialisation rendered AfL a challenge, as it reduces teachers' opportunities to get to know their students well. Teachers perceived that socio-constructivist approaches, which involve students in co-construction of learning, and require evidence of individual student's learning to feed forward into teaching, rendered the teacher-student ratio paramount. While flexible classroom management can facilitate groupings (Tomlinson and Moon 2013), the expectation that students drive their own learning seems to require different structural affordances, especially in regards to teacher-student ratio. It may therefore be instructive to consider AfL's compatibility to the structural conditions in an educational system that does not prioritise student autonomy, and also where teachers' understanding and use of AfL are often subservient to preparing students for high-stakes national examinations. Perhaps most importantly, having clarity in the

meaning of ‘learning’ in AfL is paramount, as its absence can create confusion in terms of implementation, especially if structural and sociocultural conditions appear to sometimes be at odds with those expected of socio-constructivist learning approach.

This study also draws attention to the problematic assumption that educational ideas such as AfL are acontextual, transferring easily across countries (Vavrus and Bartlett 2012, Steiner-Khamsi 2014). There is a need for scholars, policy-makers, and educators to analyse how these ideas interact with various contextual elements in source and destination contexts, and for greater understanding of how learning is construed, not just in relation to assessment (Baird et al. 2017, James 2017, Taras 2010), but in relation to contexts. Sociocultural norms, national narratives, institutional structures, and social policies compel people to relate to educational ideas differently (Alexander 2001), exerting pressures on how educational ideas are understood and enacted. In this regard, interesting questions arise more broadly. For example, the UK and the USA have borrowed assessment practices from Asian educational systems highly-ranked in PISA tests (You and Morris 2016) – however, to what extent would they be prepared for teachers’ and students’ relationships with learning perhaps shifting to more behaviouristic patterns, or what Zhao (2017) referred to as ‘side effects’ in borrowing? There are also implications from this study for teacher education: supporting teachers in discussing assessment practices in school-based professional learning communities can reposition them as owners of their own professional learning (Jiang 2017), helping them experience and, hopefully consequently, enact self-directed learning.

Conclusion

This small scale, in-depth study from Singapore has provided insight into the participant teachers' understanding and use of AfL, highlighting the need to investigate AfL as a 'situated' concept, and reflect more deeply on the complex interplay between source and destination contexts . However, it is important to recognise that, over time, teachers' relationship with assessment may change. Towards the end of this study, the Ministry of Education in Singapore announced assessment policy changes to reduce the emphasis on AoL and enhance students' 'joy of learning' and 'learning for life' (MOE 2018). These recent changes suggest that conditions would be enabled in schools for pedagogical instruction and assessment practices that enhance, rather than measure student learning; the effects remain to be seen. More widely, evidence in this study suggests that, for approaches such as AfL to be implemented in systems with cultures and structures similar to Singapore's, contextual factors need to be identified and enactment needs to be carefully considered.

Funding : This project was supported by the National Institute of Education, Office of Education Research [Grant Number OER02/17HTT].

Acknowledgements: Our gratitude goes out to the study's participants for their generous commitment to the project.

POSTPRINT

References

- Alexander, Robin J. 2001. *Culture and pedagogy: International comparisons in primary education*: Blackwell publishing.
- Baird, Jo-Anne, David Andrich, Therese N Hopfenbeck, and Gordon Stobart. 2017. "Assessment and learning: Fields apart?" *Assessment in Education: Principles, Policy & Practice* 24 (3):317-350.
- Bennett, Randy Elliot. 2011. "Formative assessment: A critical review." *Assessment in Education: principles, policy & practice* 18 (1):5-25.
- Black, Paul, and Dylan Wiliam. 1998. "Inside the black box: Raising standards through classroom assessment." *Phi Delta Kappan* 80 (2):139-148.
- Bransford, John D, Ann L Brown, Rodney R Cocking, MS Donovan, and JW Pellegrino. 2000. "How people learn: Brain, mind, experience, and school (Expanded edition). Commission on Behavioral and Social Sciences and Education National Research Council." *Washington: National Academy Press. Retrieved September 2:2006.*
- Broadfoot, Patricia. 2017. "Time to tame the leviathan? Perspectives and possibilities for a new era in assessment." *Assessment in Education: Principles, Policy & Practice* 24 (3):415-423.
- Broadfoot, Patricia M, Richard Daugherty, John Gardner, Wynne Harlen, Mary James, and Gordon Stobart. 2002. "Assessment for learning: 10 principles." *Cambridge, UK: University of Cambridge School of Education.*
- Brown, Gavin TL, Kerry J Kennedy, Ping Kwan Fok, Jacqueline Kin Sang Chan, and Wai Ming Yu. 2009. "Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment." *Assessment in education: principles, policy & practice* 16 (3):347-363.
- Clark, Vicki L Plano, and Nataliya V Ivankova. 2017. *Mixed methods research: A guide to the field*. Vol. 3: Sage publications.
- Deneen, C, Gavin W Fulmer, Gavin TL Brown, Kelvin Tan, Wei Shin Leong, and Hui Yong Tay. 2019. "Value, practice and proficiency: Teachers' complex relationship with assessment for learning." *Teacher and Teacher Education* 80:39-47.
- Department of Statistics Singapore. 2019. Population Trends, 2019. Singapore: Ministry of Trade & Industry, Republic of Singapore.

- Goldstein, Harvey. 2017. "A response to 'assessment and learning: fields apart?'." *Assessment in Education: Principles, Policy & Practice* 24 (3):388-393.
- Hatch, Amos J. 2002. *Doing qualitative research in education settings*. Albany, NY: State University of New York Press.
- Heng, Tang T., and Song, Lynn. 2020. A proposed framework for understanding educational change and transfer: Insights from Singapore teachers' perceptions of differentiated instruction. *Journal of Educational Change* 21: 595–622. <https://doi.org/10.1007/s10833-020-09377-0>
- Hogan, David, Melvin Chan, Ridzuan Rahim, Dennis Kwek, Khin Maung Aye, Siok Chen Loo, Yee Zher Sheng, and Wenshu Luo. 2013. "Assessment and the logic of instructional practice in Secondary 3 English and mathematics classrooms in Singapore." *Review of Education* 1 (1):57-106.
- Hume, Anne, and Richard K Coll. 2009. "Assessment of learning, for learning, and as learning: New Zealand case studies." *Assessment in Education: Principles, Policy & Practice* 16 (3):269-290.
- James, Mary. 2008. "Assessment and learning." In *Unlocking assessment: Understanding for reflection and application* edited by Sue Swaffield, 20-35. Routledge.
- James, Mary. 2006. "Assessment, teaching and theories of learning." *Assessment and learning* 47:60.
- James, Mary. 2017. "(Re) viewing assessment: changing lenses to refocus on learning." *Assessment in Education: Principles, Policy & Practice* 24 (3):404-414.
- Jiang, Heng. 2015. *Learning to teach with assessment: A student teaching experience in China*: Springer.
- Jiang, Heng. 2017. "Teacher Learning with Assessment in Singapore: Classroom Assessment as Searching and Researching into Teaching." *Search and Research: Teacher Education for Contemporary Contexts*:279-287.
- Jönsson, Anders. 2020. "Definitions of formative assessment need to make a distinction between a psychometric understanding of assessment and "evaluative judgment"." *Frontiers in Education* 5(2): 1-4.
- Klenowski, Val. 2009. "Assessment for learning revisited: An Asia-Pacific perspective." *Assessment in Education: Principles, Policy and Practice* 16 (3):263-268.

- LeCompte, Margaret D. 2000. "Analyzing qualitative data." *Theory into practice* 39 (3):146-154.
- Leong, Wei Shin, Haslinda Ismail, Jolene Sonia Costa, and Hong Boon Tan. 2018. "Assessment for learning research in East Asian countries." *Studies in Educational Evaluation* 59:270-277.
- Leong, Wei Shin, and Kelvin Tan. 2014. "What (more) can, and should, assessment do for learning? Observations from 'successful learning context' in Singapore." *Curriculum Journal* 25 (4):593-619.
- Miles, Mattheu B, and A Michael Huberman. 1984. "Qualitative data analysis: A sourcebook of new methods." In *Qualitative data analysis: a sourcebook of new methods*. Sage publications.
- MOE. 2018. Ministry of Education: Committee of Supply Debate 2018. Singapore.
- MOE. 2019. "Ministry of Education, Singapore." Ministry of Education, Last Modified 13 Mar 2019, accessed 14 July 2018. <https://www.moe.gov.sg/home>.
- Ratnam-Lim, Christina Tong Li, and Kelvin Heng Kiat Tan. 2015. "Large-scale implementation of formative assessment practices in an examination-oriented culture." *Assessment in Education: Principles, Policy & Practice* 22 (1):61-78.
- Shenton, Andrew K. 2004. "Strategies for ensuring trustworthiness in qualitative research projects." *Education for information* 22 (2):63-75.
- Shepard, Lorrie A. 2000. "The role of assessment in a learning culture." *Educational researcher* 29 (7):4-14.
- Steiner-Khamsi, Gita. 2014. "Cross-national policy borrowing: Understanding reception and translation." *Asia Pacific Journal of Education* 34 (2):153-167.
- Tan, Kelvin Heng Kiat. 2017. "Asking questions of (what) assessment (should do) for learning: the case of bite-sized assessment for learning in Singapore." *Educational Research for Policy and Practice* 16 (2):189-202.
- Taras, Maddalena. 2010. "Assessment for learning: assessing the theory and evidence." *Procedia-Social and Behavioral Sciences* 2 (2):3015-3022.
- Tay, HY, KHK Tan, CC Deneen, WS Leong, Gavin W Fulmer, and Gavin TL Brown. 2020. "Middle leaders' perceptions and actions on assessment: the technical, tactical and ethical." *School Leadership & Management* 40 (1):45-63.

- Tomlinson, Carol A, and Tonya R Moon. 2013. *Assessment and student success in a differentiated classroom*: ASCD.
- Tomlinson, Carol Ann. 2001. "How to Differentiate in Mixed-Ability classrooms. 2nd." *Alexandria, VA: ASCD*.
- Torrance, Harry, and John Pryor. 1998. *Investigating formative assessment: Teaching, learning and assessment in the classroom*: McGraw-Hill Education (UK).
- Vavrus, Frances, and Lesley Bartlett. 2012. "Comparative pedagogies and epistemological diversity: Social and materials contexts of teaching in Tanzania." *Comparative Education Review* 56 (4):634-658.
- Watkins, Chris. 2003. *Learning: a sense-maker's guide*: Association of Teachers and Lecturers.
- Wiliam, Dylan. 2011. "What is assessment for learning?" *Studies in educational evaluation* 37 (1):3-14.
- Zhao, Yong. 2017. "What works may hurt: Side effects in education." *Journal of Educational Change* 18 (1):1-19.

POSTPRINT

Table 1: Summary of Lesson Observation Consensus Ratings*

Differentiated Instruction Principle		LO rating [#]
I.	Creating environments that are catalysts for learning	1.5
	1. Teacher builds a community of learning	1.7
	2. Teacher builds a safe and inclusive environment (establishes trust)	2.2
	3. Teacher appreciates each child as an individual and encourages learner independence	0.7
II.	Building a foundation of quality curriculum	2.0
	1. Teacher intentionally aligns lesson goals to the curriculum	2.8
	2. Teacher organises conceptual ideas, essential facts and skills to help learners' learning	2.1
	3. Teacher ensures curriculum is authentic and relevant	1.0
III.	Using assessment to inform learning and teaching	1.1
	1. Teacher assesses Readiness	1.8
	2. Teacher assesses Interest	0.6
	3. Teacher assesses Learning Profile	0.8
	4. Teacher assesses Learner Profile	1.3
IV.	Tailoring instruction to learner variance	0.5
	1. Teacher considers learner readiness	0.8
	2. Teacher considers learner interests	0.6
	3. Teacher considers learning profiles	0.4
	4. Teacher considers learner profile	0.2
V.	Leading and managing a flexible classroom	1.8
	1. Teacher implements a flexible classroom	1.6
	2. Teacher exhibits good classroom management	2.0

LO = Lesson observation 0 = No evidence; 3 = Strong evidence; [#] Average of 39 lesson observations *This is an abbreviated version for illustrative purposes, as interview data were primarily used for analysis.