Asia-Pacific Edu Res (2023) 32:867–876
https://doi.org/10.1007/s40299-022-00702-0

REGULAR ARTICLE

The Potential of Using Formative Assessment to Enhance Academic Achievement in the Confucian-Heritage Culture: A Comparison between Hong Kong and Shanghai

Yangyu Xiao[1] · Yuyang Cai[2] · Qianwen Ge[3] · Yan Yang[3]

Accepted: 28 November 2022 / Published online: 11 December 2022

Abstract Formative assessment has been considered an essential concept for motivating and improving learning. However, implementing formative assessment has been confronted with challenges in Confucian-heritage culture (CHC) contexts such as Hong Kong and Shanghai, where the atmosphere of high-stakes tests dominates at policy and classroom levels. Whether formative assessment can lead to learning achievement in examination-oriented contexts remains an open issue. To unleash the potential benefits of formative assessment, CHC cities such as Hong Kong have been making great efforts to create conditions that can bring about the positive effect of formative assessment on learning achievement. To test this potential benefit of formative assessment in CHC contexts, the current study compared the impact of formative assessment on reading achievement in Hong Kong and Shanghai by including the mediation of motivation between formative assessment and achievement. Results of multi-group path analysis with the Programme for International Student Assessment (PISA) 2009 data showed: (a) a positive direct effect of formative assessment on achievement in Hong Kong and a negative direct effect in Shanghai, (b) a positive indirect effect through motivation both in Hong Kong and Shanghai, and (c) a total positive effect in both Hong Kong and Shanghai, but this total effect was more substantial in Hong Kong whereas trivial in Shanghai. The positive impact of formative assessment, particularly the relatively more significant total effect on the Hong Kong side, supports the potential of using formative assessment to enhance student learning in CHCs through local policy reform on assessment and teacher training on assessment literacy.

Keywords Formative assessment · Reading achievement · Reading motivation · PISA

Highlights

- Formative assessment strategies produced a positive but small effect on reading achievement in Hong Kong and Shanghai.
- The total effect of formative assessment on reading was significantly larger in Hong Kong than in Shanghai.
- The larger effect of formative assessment in Hong Kong could be attributed to the mediation of reading motivation.
- It is also possible that the efforts invested by Hong Kong in enhancing formative assessment practices have fostered slight outperformance on the Hong Kong side.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40299-022-00702-0.

✉ Yuyang Cai
sailor_cai@hotmail.com

[1] School of Humanities and Social Science, The Chinese University of Hong Kong, Shenzhen, China

[2] Centre for Language Education and Assessment Research (CLEAR) & School of Languages, Shanghai University of International Business and Economics, 1900 Wenxiang Road, Shanghai, Songjiang, China

[3] School of Languages, Shanghai University of International Business and Economics, 1900 Wenxiang Road, Shanghai, Songjiang, China

Introduction

In recent years, an increasing number of formative assessment studies have shown that formative assessment leads to learning gains through identifying students' weaknesses, teaching for understanding, engaging students actively, and increasing students' motivation (e.g., Li, 2016; McMillan et al., 2017; Wiliam et al., 2004). The implementation of formative assessment, however, has met challenges in the Confucian-heritage culture contexts (CHCs), where instructions are usually examination-driven. Under this culture, teachers have been driven to improve students' test scores rather than implement formative assessment practices that have the potential to enhance learning. Even though some teachers intended to incorporate formative assessment in their instructions (e.g., by providing formative feedback to students), these attempts were still constrained by summative assessment. For instance, they had to compromise by assigning a summative mark while giving feedback or generating feedback based on students' practice of summative test papers (Lee & Coniam, 2013).

To unleash the beneficial effect of formative assessment in CHC contexts, assessment experts and classroom teachers attempted a contextualized version of formative assessment (Kennedy et al., 2008). In Hong Kong, the Educational Bureau has been promoting formative assessment across school curricula since 2000, and formative assessment training has been provided to pre-service and in-service teachers (Curriculum Development Council, 2014; 2017). In contrast, in most parts of the Chinese mainland (e.g., Shanghai), the concept of formative assessment is still relevantly new. Comparing the case of Hong Kong with Shanghai would help understand the potential impact of formative assessment in Confucian-heritage contexts.

The purpose of the current study was to compare the effect of formative assessment in two CHC contexts with different implementing schemes of formative assessment in China: Hong Kong and Shanghai. Given the well-documented role of motivation in bridging the relationship between formative assessment and learning achievement (Guthrie et al., 2012), our comparison included the mediation of motivation between formative assessment and learning achievement. The study was conducted using the Programme for International Student Assessment (PISA) 2009 data provided by the Organization of Economic Cooperation and Development (OECD).

Formative Assessment in the Confucian-Heritage Culture (CHC)

There are cross-cultural variations in the understanding of formative assessment. Black and Wiliam (2009) defined formative assessment as classroom procedures where evidence about student learning is elicited, interpreted, and used by teachers, peers, or students themselves to make decisions on their subsequent learning, close the learning gap, and achieve the learning goals. Under such a conception, the optimal formative assessment should be integrated into the instruction process and encourage students' active involvement in the assessment process (Brown, 2019; Carless, 2011). However, in the Confucian-heritage contexts, formative assessment practices are usually carried out in sub-optimal ways, which are featured with test-related practices, such as test preparation, successive quizzes, formative use of summative tests (FUST) (Black et al., 2003; Carless, 2011; Xiao, 2017; Yan et al., 2021). Such formative assessment practices are probably related to the sociocultural traditions in the CHC context, which emphasize test performance and competitions among students and adopt behaviorist orientation toward pedagogy and teacher-centered teaching approaches (Carless, 2011; Poole, 2016).

In Hong Kong, the government has incorporated several formative elements to mitigate the negative impact of high-stakes tests. The Basic Competence Assessment was introduced in 2004 and provided a web-based item bank where teachers were encouraged to flexibly use the items to identify students' learning needs (Yan et al., 2021). Subsequently, the School-based Assessment (SBA) was introduced to the education system in Hong Kong. The SBA is a test where students are assessed by their course teachers in low-stress conditions. The SBA has the merits of alleviating students' test anxiety allowing the teachers to offer feedback for subsequent learning (Gao, 2009). The teacher's marks in SBA take 15% of the high-stakes HKDSE (Hong Kong Diploma of Secondary Education Examination). Formative assessment has also been written explicitly into local assessment guidelines to promote its use at the classroom level, emphasizing concrete learning objectives, quality feedback, and developing students' understanding (Curriculum Development Council, 2014; 2017). Regular training workshops are provided to ensure teachers have competence in implementing formative assessment in classrooms (Lam, 2015; Lee & Coniam, 2013).

Research studies discussed how formative assessment can be implemented in classroom teaching. Gan and Leung (2020) proposed that formative assessment can be integrated into task-based language teaching in primary English writing classes by sharing sample works with students, allowing students to test their understanding through guided gap-filling practices, and involving students in process-based writing, which requires pre-writing, drafting, evaluating, and revision. Although Gan and Leung (2020) did not collect empirical evidence, their study showed that Hong Kong language teachers have attempted to integrate formative assessment into everyday classrooms. In an empirical study of Hong

Kong secondary school writing classes, teachers shared successful criteria by modeling and deconstructing sample texts and explaining criteria explicitly; they also provided feedback on a self-designed form based on the criteria mentioned. However, due to school policies, teachers corrected every mistake and provided students with a numerical score, and students did not have the chance to write a second draft (Lee & Coniam, 2013). The study indicates that teachers faced the dilemma of adopting good formative assessment practices and meeting the school requirements. In another study, Carless and Lam (2014) reported using students-generated test papers as a revision strategy. Asking students to design their test papers, which may not be considered a desirable strategy in western literature, turned out to be a productive revision strategy in Hong Kong, as it made students aware of what they could anticipate (learning goals), involved students actively in generating and answering test questions, and provided them with ownership of their study. The above studies showed that Hong Kong teachers had attempted to adopt students-centered formative assessment practices, despite the potential contextual constraints such as school policies and exam-oriented cultures that discourage the application of optimal formative assessment.

In contrast to Hong Kong, formative assessment is still a relevant new concept in the Chinese mainland. The Ministry of Education conceptualized formative assessment as self-assessment, peer assessment, and assessment conducted by teachers and school administrators to observe, evaluate and monitor the students' learning effectiveness (Ministry of Education, 2003). This conception regards formative assessment more like well-planned assessment procedures instead of the divergent and student-directed process. Probably because the formative assessment was introduced without the structural change of the education system (Poole, 2016), empirical studies in the Chinese mainland extensively addressed how formative assessment practices were constrained by the social and cultural realities, such as the large class size (Hu, 2002), the lack of teacher training and professional development (Gu, 2014; Tang & Adamson, 2014), the dominant role of high-stakes summative assessment (Chen, 2015; Xiao, 2017), and the authoritative role of teachers as well as the deep-rooted transmissive way of teaching (Chen, 2015).

Facing the social realities, teachers faced the dilemma of implementing the reform and teaching to the test. A study on the formative use of summative tests (FUST) revealed that teachers used test follow-up strategies such as questioning, re-teaching, teacher-student conferences, and student self-reflection to enhance students' understanding. Teachers' practices diverged regarding whether teacher feedback encouraged reflection or imposed teachers' views and whether re-teaching clarified understanding or transmitted the knowledge to students (Xiao, 2017). Whereas the previous study described a situated version of formative assessment by taking into consideration of the high-stakes tests, Yin and Buck (2015) attempted to incorporate formative assessment into daily classrooms and reported the use of "think aloud" tasks, extended-response questions and reflection tasks to enhance students' understanding; however, their actions were constrained by students' attempts to seek for the only correct answer and the pressure for securing high scores in high-stakes test (Yin & Buck, 2015).

Generally, formative assessment practices in the Confucian-heritage cultural contexts are affected by sociocultural traditions. These include the social expectations of high scores in external examinations and the traditional belief that teachers should be the main party to impart knowledge (Carless, 2011; Poole, 2016). Nevertheless, formative assessment practices in Hong Kong seemed to have incorporated more of the state-of-art principles of formative assessment compared with the practices in the Chinese mainland.

Formative Assessment and Academic Achievement

Academic achievement is used interchangeably with learning achievement and learning outcome to refer to students' learning gains or literacy in reading, mathematics, and science (OECD, 2009). As a particular domain of learning achievement, reading achievement is defined as students' capacity to "understand, use, reflect on and engage with written texts, to achieve one's goals, to develop one's knowledge and potential, and to participate in society." (OECD, 2009, p. 14). OECD has used this conceptualization of reading achievement as an operationalizing definition to develop standardized PISA reading achievement tests in all participating countries and regions, including Hong Kong and the Chinese mainland.

Extensive evidence shows that formative assessment leads to learning gains and academic achievement by sharing standards and criteria, eliciting evidence of learning, providing feedback, and activating students in the assessment process (Black & Wiliam, 1998; McMillan et al., 2017; Wiliam, 2010; Wiliam et al., 2004). Black and Wiliam's (1998) seminal review of 250 articles revealed an effect size between 0.40 and 0.70 regarding the efficiency of formative assessment on student learning. A subsequent intervention study further supports this positive impact. Wiliam et al. (2004) revealed that the majority of effect sizes in classes with formative assessment practices were around 0.2–0.3, with a medium value of 0.27; interestingly, the positive effect size was observed in the Year 11 group, whose performance was measured against the large-scale high-stakes GCSE tests, showing that the positive impact can be found in the external assessments. Although researchers did not fully agree on the effect size,

a consensus has been reached that formative assessment improves student academic achievement (Li, 2016).

When examining the impact of formative assessment on reading achievement, Li (2016) analyzed 2009 PISA data of students from the USA and revealed that formative assessment has a direct positive relationship with students' reading achievement ($\beta = 0.085$) and students' attitudes towards reading. Brookhart et al. (2010) investigated the impact of formative assessment practices on reading achievement in a remedial reading class in the USA; the statistical evidence shows that the gains in reading achievement of the formative assessment group were differentially more significant than students who were not in the formative assessment group. The learning gains were believed to be associated with greater motivation and engagement with reading tasks (Brookhart et al., 2010), the chances for teachers to identify weaknesses through assessments (Adamu et al., 2020), or the targeted feedback on students' reading performance (Marchand & Furrer, 2014). Whereas there is clear evidence that formative assessment leads to learning gains, studies investigating the relationship between formative assessment and learning gains in the CHC context are still scarce.

Educators' understanding of theories has been considered to affect their educational practices and students' achievement positively, as teachers with a better understanding of formative assessment would involve students in meaningful and constructive learning activities (Farkas & Jang, 2019). By tracing students' performance change for one and a half years in Swedish schools, Andersson and Palm (2017) revealed that classes taught by teachers with professional training in formative assessment outperformed those taught by teachers without training at a statistically significant level ($p = 0.036$; $d = 0.66$). Yin et al. (2008) also found that teachers' assessment knowledge affected their assessment practices, which in turn affected students' motivation and achievement. Given the different levels of support and training teachers received in different regions in the CHC contexts, comparing the impact would generate implications for formative assessment practices.

Formative Assessment and Motivation

Formative assessment has been found to enhance students' motivation to learn by providing clear learning targets and feedback that give students hope and positive expectations (Cauley & McMillan, 2010; Shepard et al., 2018). In a study of first- and second-year Dutch university students, Leenknecht et al. (2021) found a positive impact of formative assessment on autonomous motivation ($\beta = 0.399$, $p < 0.001$), and the effect was mediated through autonomy satisfaction and competence satisfaction. A qualitative study in Hong Kong also supported the positive effect of formative assessment on motivation. Lee (2011) found that, after receiving teachers' feedback and support for a semester, students' fear and anxiety towards writing changed to enjoying the challenge of writing. Concerning reading motivation, Förster and Souvignier (2014) believed that feedback on the learning progress enhanced the positive self-perception of the ability to read and thus enhanced students' intrinsic motivation, which was most beneficial to achievement.

In addition, students experience close emotional bonds in a social group when doing formative assessments (Lam et al., 2017; Leenknecht et al., 2021). Support from teachers and peers helps build a close emotional relationship (Lam et al., 2017). When analyzing the impact of classroom practices on motivation, Guthrie et al. (2012) reported that students felt more motivated to learn when the instructor considered students' perspectives, acknowledged their feelings, and provided chances for self-directed learning. Reviews above suggest that students feel more motivated to learn because they feel confident in their ability and emotionally supported. Such motivation, as Lam et al. (2017) testified, can go beyond the course and improve students' overall experience of positive emotions.

The motivation to learn also acts as a mediating factor between formative assessment and achievement. Nolen (2011) explained that formative assessment aims not merely to communicate the standards to students but to motivate students to work towards the standard. In this sense, providing praise with informative feedback increases students' motivation to achieve the expected standard. As stated in the PISA 2009 framework, motivated and highly engaged readers spend considerable time reading a wide range of texts. They consider reading an interesting and valuable experience (OECD, 2009). According to Guthrie et al. (2012), reading motivation and engagement mediate the effect of classroom practices on students' learning achievement and lead to better reading competence. Similar findings were revealed by an intervention study with fourth-graders in mid-Atlantic state schools. Guthrie et al. (2007) reported that interest in reading explained 12% of the variance in students' reading comprehension achievement in that study. Although Guthrie et al. (2007)'s study did not explicitly address formative assessment, their study supported that motivation can be a critical mediating factor in students' reading achievement.

To sum up, formative assessment leads to increased motivation and academic achievement, and motivation can work as a mediating factor. We summarize the relationship as our conceptual framework in Fig. S1. To further testify to this framework in the CHC contexts, the current study examined the extent to which formative assessment affects reading achievement through reading motivation using the PISA 2009 data provided by students from Shanghai and Hong Kong. Two research questions are put forward as follows:

Question 1. To what extents do formative assessment strategies affect the reading achievement of Shanghai and Hong Kong adolescent students?

Question 2. To what extent does reading enjoyment mediate the relationships between formative assessment strategies and reading achievement across Shanghai and Hong Kong adolescent students?

Method

Data Source

The current study used PISA2009 data (https://www.oecd.org/pisa/data/pisa2009database-downloadabledata.htm). Our dataset included 9809 15-year-old adolescents from Shanghai (51% girls) and Hong Kong (48% girls), each occupying 52% and 48% of the total sample. The mean age for Shanghai was 15.77 ($SD = 0.30$) and 15.75 ($SD = 0.28$) for Hong Kong. Students' social-economic status (SES) was represented using an index of economic, social, and cultural status (ESCS). The ESCS captured information regarding students' family backgrounds (OECD, 2009). The mean of ESCS was $M = -0.48$, $SD = 1.06$ for Shanghai and $M = -0.76$, $SD = 1.21$ for Hong Kong.

Measures

Reading motivation: Reading motivation was operationalized in terms of reading enjoyment using a four-point scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). Eleven items were used for students to report their attitudes towards different statements about reading. Examples of statements included: "Reading is one of my favorite hobbies" and "I enjoy going to a bookstore or a library." For detailed information regarding this motivation scale, please refer to OECD (2009). The overall mean of reading motivation was $M = 2.97$, $SD = 0.46$ for Shanghai, and $M = 2.79$, $SD = 0.50$ for Hong Kong. The internal consistency of the scale was $\alpha = 0.84$ for Shanghai and 0.87 for Hong Kong.

Formative assessment strategies: Teachers' use of formative assessment strategies was measured using a nine-item scale. Each item was operationalized using a four-point scale, ranging from 1 (*Never or hardly ever*) to 4 (*In all lessons*). The formative assessment strategies did not address strategies particular to reading but general for learning in all domains. An example was, 'The teacher tells students in advance how their work is going to be judged.' The overall mean of formative assessment strategies was $M = 2.60$, $SD = 0.56$ for Shanghai, and $M = 2.56$, $SD = 0.54$ for Hong Kong. The internal consistency of the scale was $\alpha = 0.79$ for Shanghai and 0.85 for Hong Kong.

Reading achievement: PISA 2009 defined reading literacy as students' ability to understand, use, reflect on, and engage with written texts to achieve their purposes (OECD, 2009). The mean of reading was $M = 0.556$, $SD = 80.06$ for Shanghai and $M = 534.16$, $SD = 83.06$ for Hong Kong.

Data Analysis

Cases with complete missing values on reading motivation, formative assessment strategies, or reading achievement were excluded. PISA 2009 provided Rasch-calibrated composite scores for multiple-indicator variables such as reading motivation and formative assessment strategies (see OECD, 2012). Given the advantage of Rasch-calibrated scores in controlling for measurement errors, Rasch scores instead of raw scores were used for primary data analysis. To explore the direct and indirect effect of formative assessment strategies on reading achievement, we conducted a two-group path analysis across the Shanghai and Hong Kong data.

The model was computed on M*plus* 7.4 (Muthén & Muthén, 1998–2015) with the estimator of Maximum Likelihood Robust (MLR) (Satorra & Bentler, 1994). Multiple indices were used to determine a good model-data fit: Comparative Fit Index (CFI) and Tucker–Lewis index (TLI) values no smaller than 0.95, and Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) values no larger than 0.05 (Mueller & Hancock, 2012).

Findings

Correlations between Reading Motivation, Formative Assessment, and Achievement

Table S1 presents the correlations among reading motivation, formative assessment strategies, and reading achievement across the Shanghai and Hong Kong groups.

As shown, for both cities, the highest correlation was between reading motivation and reading achievement: $r = 0.36$, $p < 0.01$ for Shanghai and $r = 0.37$, $p < 0.01$ for Hong Kong. Followed was the correlation between formative assessment strategies and reading motivation: $r = 0.14$, $p < 0.01$ for Shanghai and $r = 0.17$, $p < 0.01$ for Hong Kong. The correlation between formative assessment strategies and reading achievement was $r = 0.10$, $p < 0.01$ for Hong Kong but not significant for Shanghai.

Two-Group Path Analysis

We conducted a two-group path analysis to examine the direct effects of formative assessment strategies and reading motivation on reading achievement and the indirect effect of

formative assessment through reading motivation on reading achievement. This model had a perfect fit with the data: *TLI* = 1.00, *CFI* = 1.00, *RMSEA* = 0.000, and *SRMRs* = 0.000.

The standardized estimates of the path analysis with each group are shown in Fig. S2. The direct effect of formative assessment strategies on reading was negative for Shanghai ($\beta = -0.04$, $p < 0.01$, 95% CI [− 0.06, − 0.02]), but positive for Hong Kong ($\beta = 0.04$, $p < 0.01$, 95% CI [0.01, 0.06]). A Wald test indicated a significant difference between the two estimates ($x^2 = 17.10$, *degree of freedom* = 1, $p < 0.001$).

The direct effect of reading motivation on reading was $\beta = 0.36$, $p < 0.01$, 95% CI [0.34, 0.39] for Shanghai and $\beta = 0.36$, $p < 0.01$, 95% CI [0.34, 0.38] for Hong Kong. The direct effect of formative assessment strategies on reading motivation was $\beta = 0.12$, $p < 0.01$, 95% CI [0.12, 0.17] for Shanghai, and $\beta = 0.17$, $p < 0.01$, 95% CI [0.14, 0.19].

The indirect effect of formative assessment strategies on reading is, therefore, represented by the product of the path coefficient from formative assessment strategies to reading motivation and that from reading motivation to reading. The indirect effect of formative assessment strategies by way of reading motivation was $\beta = 0.05$, $p < 0.01$, 95% CI [0.04, 0.06], and that for Hong Kong was $\beta = 0.06$, $p < 0.01$, 95% CI [0.05, 0.07]. A Wald test indicated a significant difference between the two indirect effects: $x^2 = 15.57$, *degree of freedom* = 1, $p < 0.001$).

The total effect of formative assessment strategies on reading is the sum of the direct and indirect effects. The total effect was $\beta = 0.10$, $p < 0.01$ for Hong Kong and $\beta = 0.01$, $p < 0.01$ for Shanghai.

Discussion

Question 1. To what extents do formative assessment strategies affect the reading achievement of shanghai and Hong Kong adolescent students?

The total effect of formative assessment strategies on reading achievement was positive in Hong Kong ($\beta = 0.10$, $p < 0.01$), meaning that a standard deviation unit increase in formative assessment was associated with a 0.10 standard deviation unit increase in reading achievement. This effect was also positive for Shanghai, but the effect size was trivial ($\beta = 0.01$, $p < 0.01$), such that formative assessment can be considered to almost have no meaningful effect on reading achievement. Compared with the impact of formative assessment on reading achievement found in Li (2016)'s study with the US counterparts using the same PISA 2009 data ($\beta = 0.19$, $p < 0.001$), the total effect of formative assessment on reading achievement in both Hong Kong and Shanghai was small or even trivial. As the learning gains achieved through formative assessment are associated with students' engagement with assessment tasks and more targeted feedback on students' performance (Brookhart et al., 2010; Marchand et al., 2014), it is reasonable to assume that formative assessment in China, which are more teacher-directed and involve a lot of test-related practices, leads to learning gains to a less optimal degree. However, the comparatively small effect size in both cities still sheds a positive light on the potential benefits of formative assessment in the CHC context. As students' performance in the PISA test has been used as the indicator of reading achievement, our study supports the contention that formative assessment could lead to academic achievement in external examinations (McDonald & Boud, 2003; Wiliam et al., 2004). This would possibly encourage the implementation of formative assessment in the Confucian-heritage contexts and a broader educational context where high-stakes external examinations matter. As Yan and Cheng (2015) point out, teachers' intention to conduct the formative assessment is strongly predicted by their instrumental attitudes towards formative assessment (i.e., the positive consequence of formative assessment on student learning).

The positive findings become more evident when it comes to the direct effect. The direct effect of formative assessment on reading achievement was negative for Shanghai ($\beta = -0.04$, $p < 0.01$) but positive for Hong Kong ($\beta = 0.04$, $p < 0.01$). The differences between the two cities are not surprising, as the existing literature indicates that formative assessment practices in Hong Kong involved students more actively and adopted more state-of-the-art principles of formative assessment, compared with their counterparts in the Chinese mainland (e.g., Lee & Coniam, 2013; Poole, 2016; Xiao, 2017). Such findings reinforce the previous contention that the impact of formative assessment could be affected by assessment practices, which are influenced by the support offered at the policy, school, and classroom levels (Liu & Xu, 2017). Although we did not collect direct evidence of formative assessment classroom practices, our interpretation is still valid, considering we have measured the relationship between the strength of formative assessment and reading achievement through a large and representative cross-sectional sample (Li, 2016).

The negative direct effect of formative assessment on reading achievement in Shanghai and the comparatively small positive effect size in Hong Kong indicate the need for teacher training to help teachers develop formative assessment activities that meet their local needs. As Yan and Cheng (2015) pointed out, despite the good intention of the Hong Kong government in implementing formative assessment, teachers' classroom practices are affected by internal factors such as teachers' intention, their self-efficacy, and external factors such as class size, workload, class climate, and school policies which are beyond the teachers' control. Thus, assessment training needs to help teachers adapt and integrate the assessment practices into contextualized

classroom teaching (Carless, 2011; Wiliam, 2010; Yan & Chan, 2015).

Question 2. To what extent does reading enjoyment mediate the relationships between formative assessment strategies and reading achievement across shanghai and Hong Kong adolescent students?

Our findings show that the indirect effect of formative assessment on reading achievement by way of reading motivation was positive both in Hong Kong and in Shanghai, with a slightly larger effect in Hong Kong ($\beta = 0.06$, $p < 0.001$) than in Shanghai ($\beta = 0.05$, $p < 0.001$). In both cities, formative assessment enhanced students' motivation to read, which was positively associated with reading achievement. The direct impact of motivation on reading achievement appeared to be the same in both Shanghai and Hong Kong ($\beta = 0.36$, $p < 0.001$), whereas the direct impact of formative assessment on reading motivation was larger in Hong Kong ($\beta = 0.17$, $p < 0.001$) than that in Shanghai ($\beta = 0.14$, $p < 0.001$).

The positive albeit small effect of formative assessment on reading motivation in both cities supports that students' motivation is enhanced when their psychological needs for competence are satisfied (Deci & Ryan, 2008). Formative assessment enhances students' perception of competence by articulating clear learning goals and providing instructive feedback so students can view their progress and become more confident (e.g., Cauley & McMillan, 2010; Nolen, 2011). As students develop trust in their competence in reading, they become willing to read, engage in reading, and develop an understanding of reading skills (Förster & Souvignier, 2014; Proctor et al., 2014; Marchand et al., 2014).

In addition, our study supports previous literature that reading motivation mediated the impact of formative assessment on reading achievement (Guthrie et al., 2007, 2012; Nolen, 2011). This is not surprising as students became more motivated to read and clearly understood reading skills. Their reading achievement will be improved, as intrinsic motivation is positively associated with higher achievement (Deci & Ryan, 2008).

In general, our study reinforces the positive impact of formative assessment on motivation and the potential mediating effects of motivation. The comparatively stronger direct impact of formative assessment on motivation in Hong Kong further supports our earlier assumption that formative assessment is better advocated in Hong Kong than in Shanghai, as in the former formative assessment is better supported by government policies and professional training.

Conclusion and Implications

Using PISA 2009 data from Hong Kong and Shanghai, the current study contributes to understanding the impact of formative assessment on reading achievement using a representative data set and rigorous analysis. It is one of the few studies that compared the effect of formative assessment on reading achievement in two cities, both situated in an examination-oriented culture but with education systems directed under different education policies. The comparison regarding the impact of formative assessment on reading achievement between Hong Kong and Shanghai showed that formative assessment appeared to work more efficiently, though with small effect sizes, predicting reading achievement in Hong Kong than in Shanghai. This could probably be a result of the Hong Kong government's constant support in creating an assessment for learning culture and localization of formative assessment in the past decade, for example, explicitly encouraging a culture of assessment for learning in the school curriculum, reducing the number of high-stakes tests (Yan & Brown, 2021), and providing localized formative assessment training to in-service and pre-service teachers (Lam, 2015).

Although our findings are bound to reading achievement, the domain-general feature of formative assessment practices measured in PISA 2009 and the high correlations between reading and the other domains (Cai & Yang, 2022) all suggest the significant implications of our findings for formative assessment practices in other domains.

Most valuably, this comparison of the two cities supports the necessity and possibility of localizing formative assessment in the Confucian-heritage culture contexts. Our findings suggested that despite the influence of examination-oriented culture, formative assessments still have the potential to enhance learning achievement. Yan and Brown (2021) point out that the good intention of introducing formative assessment can be achieved when evidence shows that formative assessment leads to improved academic achievement and achievement in external examinations. In this way, teachers would see how formative assessment is compatible with the examination-oriented culture that they must face (Yan & Brown, 2021).

The comparatively small effect size in both cities was not unpredictable, considering practical constraints such as the examination-oriented learning context, large class size, and the presence of influential public examinations dominating the education systems on both sides. It was also the case that continuous professional training and school support were not sufficiently provided across all schools (Yan & Brown, 2021).

This comparative study generated implications for theoretical research in formative assessment and teacher professional development. Given the contextual influence identified in this study, we highlight the importance of further research to investigate the features of formative assessment in a specific social and cultural context. Additionally, professional teacher training needs to support the localization of

the western concept of formative assessment and develop teachers' assessment literacy. To successfully promote formative assessment in the CHC context, the training needs to focus more on helping teachers compromise their knowledge of assessment with the context to which they belong, instead of providing one-fits-all assessment training (Yan & Brown, 2021). The training, which is still less common in the Chinese mainland, is essential for teachers to become resourceful and reflective assessors (Xu & Brown, 2016). In this way, teachers can think more about how they can accommodate assessment policies and adapt state-of-the-art formative assessment practices to meet the needs of their students.

Finally, this study's limitation is that it relied on cross-sectional data, and the causal relationship could not be fully decided. Most of the interpretations made in the current study were based on the previous literature. How formative assessment practices affect reading achievement can be more complex than the study presented. However, the comparison between the two cities confirms the need to study the impact of formative assessment on reading achievement by considering various influencing factors, such as educational policies and the professional training teachers receive. Further experimental studies are needed to add more evidence to explore the conditions that might facilitate the effect of formative assessment on reading achievement.

Acknowledgements This work was supported by The Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (Code: TP2018068) given to the second author.

Author Contributions Conceptualization: YC; Methodology: YC; Formal analysis and investigation: YC; Writing—original draft preparation: YX, YC; Writing—review and editing: YC, YX, QG, YY; Funding acquisition: YC , YX; Resources: YC; Supervision: YC.

Funding This study was funded by Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning, TP2018068, Yuyang Cai, Shenzhen Educational Sciences Planning Scheme (The 14th Five-Year Plan), YBZZ21019, Yangyu Xiao

Data Availability The data that support the findings of this study are openly available in OECD PISA 2009 dataset at: https://www.oecd.org/pisa/data/2009database/.

Declarations

Conflict of interest None.

Ethical Approval Not applicable.

Consent to Participate Not applicable.

Consent to Publication The author consents to publish this article in *The Asia Pacific Education Researcher*.

Appendix

(A) Formative assessment items

a)	The teacher explains beforehand what is expected of the students
b)	The teacher checks that students are concentrating while working on the reading assignment
c)	The teacher discusses students' work, after they have finished the reading assignment
d)	The teacher tells students in advance how their work is going to be judged
e)	The teacher asks whether every student has understood how to complete the reading assignment
f)	The teacher marks students' work
g)	The teacher gives students the chance to ask questions about the < reading assignment >
h)	The teacher poses questions that motivate students to participate actively
i)	The teacher tells students how well they did on the < reading assignment > immediately after

(B) Reading motivation items

a)	I read only if I have to
b)	Reading is one of my favorite hobbies
c)	I like talking about books with other people
d)	I find it hard to finish books
e)	I feel happy if I receive a book as a present
f)	For me, reading is a waste of time
g)	I enjoy going to a bookstore or a library
h)	I read only to get information that I need
i)	I cannot sit still and read for more than a few minutes
j)	I like to express my opinions about books I have read
k)	I like to exchange books with my friends

References

Adamu, A., Tsiga, A. U., & Zuilkowski, S. S. (2020). Teaching reading in northern Nigeria: The challenges of large class size. *Pedagogy, Culture & Society*. https://doi.org/10.1080/14681366.2020.1794948

Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction, 49*, 92–102. https://doi.org/10.1016/j.learninstruc.2016.12.006

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Open University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment educational assessment. *Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.4135/9781446250808.n13

Brookhart, S. M., Moss, C. M., & Long, B. A. (2010). Teacher inquiry into formative assessment practices in remedial reading classrooms. *Assessment in Education: Principles, Policy & Practice, 17*(1), 41–58. https://doi.org/10.1080/09695940903565545

Brown, G. T. (2019). Is assessment for learning really assessment? *Frontiers in Education*. https://doi.org/10.3389/feduc.2019.00064

Cai, Y., & Yang, Y. (2022). The fluid relation between reading strategies and mathematics learning: A perspective of the Island Ridge Curve. *Learning and Individual Differences*. https://doi.org/10.1016/j.lindif.2022.102180

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

Carless, D., & Lam, R. (2014). Developing assessment for productive learning in confucian-influenced settings. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 167–179). Springer.

Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 83*(1), 1–6. https://doi.org/10.1080/00098650903267784

Chen, J. (2015). Formative assessment as a vehicle for changing classroom practice in a specific cultural context. *Cultural Studies of Science Education, 10*(3), 753–762. https://doi.org/10.1007/s11422-014-9599-7

Curriculum Development Council (CDC). (2014). *Basic education curriculum guide: To sustain, deepen and focus on learning to learn (Primary 1–6)*. Curriculum Development Council.

Curriculum Development Council (CDC). (2017). *Secondary education curriculum guide*. Curriculum Development Council.

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology/Psychologie Canadienne, 49*(1), 14–23. https://doi.org/10.1037/0708-5591.49.1.14

Farkas, W. A., & Jang, B. G. (2019). Designing, implementing, and evaluating a school-based literacy program for adolescent learners with reading difficulties: A mixed-methods study. *Reading & Writing Quarterly, 35*(4), 305–321. https://doi.org/10.1080/10573569.2018.1541770

Förster, N., & Souvignier, E. (2014). Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction, 32*, 91–100. https://doi.org/10.1016/j.learninstruc.2014.02.002

Gan, Z., & Leung, C. (2020). Illustrating formative assessment in task-based language teaching. *ELT Journal, 74*(1), 10–19. https://doi.org/10.1093/elt/ccz048

Gao, M. (2009). Students' voice in school-based assessment in Hong Kong. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about assessment for learning* (pp. 107–130). Information Age Publishing.

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. https://doi.org/10.1080/0969594X.2013.836076

Guthrie, J. T., Hoa, A. L. W., Wigfield, A., Tonks, S. M., Humenick, N. M., & Littles, E. (2007). Reading motivation and reading comprehension growth in the later elementary years. *Contemporary Educational Psychology, 32*(3), 282–313. https://doi.org/10.1016/j.cedpsych.2006.05.004

Guthrie, J. T., Wigfield, A., & You, W. (2012). Instructional contexts for engagement and achievement in reading. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 601–634). Springer.

Hu, G. (2002). Potential cultural resistance to pedagogical imports: The case of communicative language teaching in China. *Language, Culture and Curriculum, 15*(2), 93–105. https://doi.org/10.1080/07908310208666636

Kennedy, K., Chan, J., Fok, P., & Yu, W. (2008). Forms of assessment and their potential for enhancing learning: Conceptual and cultural issues. *Educational Research for Policy and Practice, 7*(3), 197–207. https://doi.org/10.1007/s10671-008-9052-3

Lam, B.-H., Cheng, R.W.-Y., & Yang, M. (2017). Formative feedback as a global facilitator: Impact on intrinsic and extrinsic motivation and positive affect. In S. C. Kong, T. L. Wong, M. Yang, C. F. Chow, & K. H. Tse (Eds.), *Emerging practices in scholarship of learning and teaching in a digital era* (pp. 265–288). Springer.

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169–197. https://doi.org/10.1177/0265532214554321

Lee, I. (2011). Bringing innovation to EFL writing through a focus on assessment for learning. *Innovation in Language Learning and Teaching, 5*(1), 19–33. https://doi.org/10.1080/17501229.2010.502232

Lee, I., & Coniam, D. (2013). Introducing assessment for learning for EFL writing in an assessment of learning examination-driven system in Hong Kong. *Journal of Second Language Writing, 22*(1), 34–50. https://doi.org/10.1016/j.jslw.2012.11.003

Leenknecht, M., Wijnia, L., Köhlen, M., Fryer, L., Rikers, R., & Loyens, S. (2021). Formative assessment as practice: The role of students' motivation. *Assessment & Evaluation in Higher Education, 46*(2), 236–255. https://doi.org/10.1080/02602938.2020.1765228

Li, H. (2016). How is formative assessment related to students' reading achievement? Findings from PISA 2009. *Assessment in Education: Principles, Policy & Practice, 23*(4), 473–494. https://doi.org/10.1080/0969594X.2016.1139543

Liu, J., & Xu, Y. (2017). Assessment for learning in English language classrooms in China: Contexts, problems, and solutions. In H. Reinders, D. Nunan, & B. Zou (Eds.), *Innovation in language learning and teaching: The case of China* (pp. 17–37). Palgrave Macmillan.

Marchand, G. C., & Furrer, C. J. (2014). Formative, informative, and summative assessment: The relationship among curriculum-based measurement of reading, classroom management, and reading performance. *Psychology in the Schools, 51*(7), 659–676. https://doi.org/10.1002/pits.21779

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on

performance in external examinations. *Assessment in Education: Principles, Policy & Practice, 10*(2), 209–220. https://doi.org/10.1080/0969594032000121289

McMillan, J. H., Andrade, H. L., & Heritage, M. (2017). *Using formative assessment to enhance learning, achievement, and academic self-regulation*. Routledge.

Ministry of Education. (2003). *National curriculum standards for high school English*. The People's Education Press.

Mueller, R. O., & Hancock, G. R. (2012). Structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 371–384). Routledge.

Muthén, L. K., & Muthén, B. Q. (1998–2015). *Mplus 7.4 [Computer software]*. In Muthén & Muthén

Nolen, S. B. (2011). The role of educational systems in the link between formative assessment and motivation. *Theory into Practice, 50*(4), 319–326. https://doi.org/10.1080/00405841.2011.607399

OECD. (2009). *PISA 2009 assessment framework: Key competencies in reading, mathematics and science*. OECD Publishing.

OECD. (2012). *PISA 2009 technical report*. OECD Publishing.

Poole, A. (2016). 'Complex teaching realities' and 'deep rooted cultural traditions': Barriers to the implementation and internalisation of formative assessment in China. *Cogent Education, 3*(1), 1156242. https://doi.org/10.1080/2331186X.2016.1156242

Proctor, C. P., Daley, S., Louick, R., Leider, C. M., & Gardner, G. L. (2014). How motivation and engagement predict reading comprehension among native English-speaking and English-learning middle school students with disabilities in a remedial reading curriculum. *Learning and Individual Differences, 36*, 76–83. https://doi.org/10.1016/j.lindif.2014.10.014

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. Von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Sage.

Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice, 37*(1), 21–34. https://doi.org/10.1111/emip.12189

Tang, S., & Adamson, B. (2014). *Student-centredness in urban schools in China*. London Review of Education.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18–40). Routledge.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice, 11*(1), 49–65. https://doi.org/10.1080/0969594042000208994

Xiao, Y. (2017). Formative assessment in a test-dominated context: How test practice can become more productive. *Language Assessment Quarterly, 14*(4), 295–311. https://doi.org/10.1080/15434303.2017.1347789

Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education, 58*, 149–162. https://doi.org/10.1016/j.tate.2016.05.010

Yan, Q., Zhang, L. J., & Cheng, X. (2021). Implementing classroom-based assessment for young EFL learners in the Chinese context: A case study. *The Asia-Pacific Education Researcher, 30*(6), 541–552. https://doi.org/10.1007/s40299-021-00602-9

Yan, Z., & Brown, G. T. L. (2021). Assessment for learning in the Hong Kong assessment reform: A case of policy borrowing. *Studies in Educational Evaluation, 68*, 100985. https://doi.org/10.1016/j.stueduc.2021.100985

Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education, 45*, 128–136. https://doi.org/10.1016/j.tate.2014.10.002

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. https://doi.org/10.1007/s11422-014-9572-5

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., Tomita, M. K., & Young, D. B. J. A. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education, 21*(4), 335–359. https://doi.org/10.1080/08957340802347845