Assessment in Education: Principles, Policy & Practice

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/caie20

Formative assessment, growth mindset, and achievement: examining their relations in the East and the West

Zi Yan, Ronnel B. King & Joseph Y. Haw

To cite this article: Zi Yan, Ronnel B. King & Joseph Y. Haw (2021) Formative assessment, growth mindset, and achievement: examining their relations in the East and the West, Assessment in Education: Principles, Policy & Practice, 28:5-6, 676-702, DOI: 10.1080/0969594X.2021.1988510

To link to this article: https://doi.org/10.1080/0969594X.2021.1988510

Published online: 31 Oct 2021.

ASSESSMENT IN EDUCATION: PRINCIPLES, POLICY & PRACTICE
2021, VOL. 28, NOS. 5–6, 676–702
https://doi.org/10.1080/0969594X.2021.1988510

Routledge
Taylor & Francis Group

Formative assessment, growth mindset, and achievement: examining their relations in the East and the West

Zi Yan [a], Ronnel B. King [b] and Joseph Y. Haw [a,c]

[a]Department of Curriculum and Instruction, The Education University of Hong Kong, Hong Kong SAR, China; [b]Centre for the Enhancement of Teaching and Learning, Faculty of Education, The University of Hong Kong, Hong Kong SAR, China; [c]School of Education, Ateneo de Zamboanga, Zamboanga, Philippines

ABSTRACT

Both formative assessment and growth mindset scholars aim to understand how to enhance achievement. While research on formative assessment focuses on external teaching practices, work on growth mindset emphasises internal psychological processes. This study examined the interplay between three formative assessment strategies (i.e. sharing learning progressions, providing feedback, and instructional adjustments) and growth mindset in predicting reading achievement using the PISA2018 data. We focused specifically on samples from the West (the United States, the United Kingdom, Ireland, Canada, Australia, and New Zealand) and the East (Mainland China, Hong Kong SAR, Macau SAR, Chinese Taipei, Japan and Korea) which comprised of 109,204 15-year old students. The results showed that formative assessment strategies were positively, albeit weakly, related to a growth mindset in the East, but not in the West. In contrast, growth mindset was positively related to reading achievement only in the West, but not in the East. The impacts of different formative assessment strategies on reading achievement demonstrated cross-cultural variability, but the strongest positive predictor was instructional adjustments. These findings highlight the potential synergy between formative assessment and growth mindset in enhancing academic achievement as well as the importance of cultural contexts in understanding their roles in student learning.

ARTICLE HISTORY
Received 5 January 2021
Accepted 6 September 2021

KEYWORDS
Formative assessment; growth mindset; reading achievement; PISA 2018; sharing learning progressions; feedback; instructional adjustments

Introduction

Formative assessment has been widely recognised as a crucial facilitator of effective teaching and learning (Bennett, 2011; Black & Wiliam, 1998). Similarly, growth mindset has been found to be related to optimal learning-related outcomes, such as better academic achievement (Claro et al., 2016; Costa & Faria, 2018) and self-regulated learning (Burnette et al., 2012). Although both formative assessment and growth mindset emphasise identifying the gap between a student's current achievement and the expected goal and then finding a way to close the gap (Black & Wiliam, 1998; Dweck, 2015), the research on these two areas have proceeded in parallel with limited cross-over of ideas.

CONTACT Zi Yan zyan@eduhk.hk Department of Curriculum and Instruction, The Education University of Hong Kong, D1-1/F-49,10 Lo Ping Road, Tai Po, Hong Kong, China

Both formative assessment and growth mindset focus on enhancing performance. Formative assessment emphasises the external teaching process, while growth mindset underscores the importance of internal psychological processes. Theoretically, mindsets may function as interpretative lenses that filters what information students attend to and, therefore, could potentially mediate the effectiveness of formative assessment. As one of the hurdles of formative assessment is the failure to activate students' responsibility and agency in the assessment process (Guo & Yan, 2018; Yan & Brown, 2021), investigating the role of growth mindset in the link between formative assessment and academic performance represents a meaningful direction to understand the mechanism through which formative assessment influences learning. Without taking students' mindsets into account, it would be difficult to enhance their responsibility and agency in the formative assessment process.

Hence, this study aims to examine the interplay between formative assessment, growth mindset, and academic achievement in a reading test using data from PISA 2018. We also differentiate the three specific formative assessment strategies (i.e. sharing learning progressions, providing feedback, and instructional adjustments) so as to provide a nuanced picture about the relationships among the key variables.

Formative assessment

Formative assessment has been defined as activities that elicit evidence of student learning and can be used by teachers, students, or others to make decisions about future teaching or learning plans (Black & Wiliam, 1998). As indicated by the definition, formative assessment encompasses different practices in classrooms (Bennett, 2011). However, a lack of consensus on how formative assessment should be designed and implemented is a big challenge for researchers to interpret and compare the effects of formative assessment in different contexts (McMillan et al., 2013). As a response, Wiliam and Thompson (2008) proposed a framework in which various formative assessment practices are categorised into five key strategies and one 'big idea'. The five formative strategies include 1) clarifying and sharing learning intentions and criteria for success; 2) engineering effective classroom discussions, questions, and learning tasks; 3) providing feedback that moves learners forward; 4) activating students as instructional resources for one another; and 5) activating students as the owners of their own learning (Wiliam & Thompson, 2008, p. 64). The 'big idea' refers to instructional adjustments aiming to cater for students' needs that have been identified by formative assessment strategies.

Strategies 1 and 2 have different functions: the former points out what the students are expected to achieve, while the latter identifies students' current performance levels. These two strategies, however, share a common feature – they both reflect and provide information about learning progressions that refer to the learning goal and the coordinated sequential teaching steps towards that goal (Catley et al., 2005; Duschl et al., 2011). Thus, Strategies 1 and 2 are operationalised as *sharing learning progressions* in this study. Strategy 3 (*providing feedback*) and the big idea (*instructional adjustments*) are supposed to indicate what needs to be done for students to achieve their learning goals. Strategies 4 and 5 highlight students' own responsibilities in learning. As the PISA 2018 did not

provide data on Strategies 4 and 5, this study focused on three formative assessment strategies: sharing learning progressions, providing feedback, and instructional adjustments.

The positive effect of formative assessment on academic achievement has been well documented. Black and Wiliam's (1998) meta-analysis reported effect sizes ranging from 0.4 to 0.7 favouring students receiving formative assessment. Other review studies (e.g. Bennett, 2011; Dunn & Mulvenon, 2009) also provided support for the efficacy of formative assessment in enhancing student academic achievement. In particular, formative assessment was found useful in enhancing students' reading achievement. For instance, Förster and Souvignier (2014) found that learning progress assessment, similar to formative assessment, which provides teachers with diagnostic information on students' learning progress had a positive impact on Grade 4 students' reading achievement. Chen et al. (2019) also reported that formative assessment embedded into a reading annotation system significantly improved Grade 5 students' reading comprehension performance through facilitating their self-regulated learning and reflection.

Growth mindset

The core idea of mindset theory is that students hold different 'implicit theories' or 'mindsets' about their intelligence (Dweck, 2017; Yeager & Dweck, 2012). For instance, some students view intelligence as something that they cannot change (a fixed mindset), whereas others view intelligence as something malleable that they can develop over time (a growth mindset). Mindsets are schematic knowledge structures that include beliefs about the stability of key attributes and organise the way students ascribe meaning to what happens to them in school. For example, the same event such as failing an exam can be interpreted in different ways. For a student with a growth mindset, failure may be seen as an opportunity to learn and strive harder next time. In contrast, for a student with a fixed mindset, failure may be interpreted as a verdict that one really is not good at a particular subject.

Because mindsets function as interpretative lenses, they have been found to predict important learning-related outcomes. Growth mindset has been associated with better academic achievement (Claro et al., 2016; Costa & Faria, 2018; Sisk et al., 2018), self-regulated learning (Burnette et al., 2012), mastery response after setbacks (Blackwell et al., 2007), and psychological well-being (King, 2012, 2017; Schleider et al., 2015), and enhanced sensitivity to task-relevant input leading to better learning (Schroder et al., 2014) among others. In contrast, having a fixed mindset is associated with lower levels of achievement, lower psychological well-being and giving up after failure (Smiley et al., 2016). With regard to the correlation between mindset and reading achievement, the findings in the literature are mixed. Petscher et al. (2017) found that both a global and language-specific growth mindset were associated with higher reading comprehension performance for elementary school students. In contrast, McCutchen et al. (2016) studied a sample of Grades 3–6 students and found that mindset was significantly related with maths achievement but not reading performance.

Despite the popularity of the growth mindset construct in both the popular press and scholarly circles, there are also challenges to its relevance. For example, a study conducted by Burgoyne et al. (2020) found that growth mindset effects are smaller than

those commonly found for other motivational factors. One potential reason for the mixed or weak effects is the ongoing debate on the operationalisation of the growth mindset construct. In the literature, mindsets are construed both as a domain-general construct focusing on beliefs about intelligence (e.g. Blackwell et al., 2007) and a domain-specific construct focusing on specific subject areas such as maths or language (e.g. Degol et al., 2018; Lou & Noels, 2017). Indeed, Dweck (1996), the pioneer of mindset research, acknowledged that mindsets could be both domain-general and domain-specific. In PISA 2018, mindset was operationalised as a domain-general factor pertaining to students' beliefs about the malleability of their general intelligence.

These challenges, notwithstanding, mindset proponents have made a strong case for the role of mindsets in learning and achievement. However, most of the existing mindset studies have been conducted within one cultural context and using convenience samples. We contribute to this ongoing discussion by drawing on a large-scale international dataset such as PISA that includes student participants from both the East and the West, thereby expanding the generalisability of past research.

The link between formative assessment and growth mindset

The research on formative assessment and growth mindset have proceeded in parallel with different foci: the former reflects a pedagogical approach while the latter mostly focuses on a psychological approach. Both formative assessment and growth mindset research emphasise telling the truth about a student's current achievement and then finding a way to help him or her perform better (Black & Wiliam, 1998; Dweck, 2015). Hence, both are areas with great potential for improved practice to lead to improved learning outcomes (Bibbens, 2018) and there are possible synergies between them.

Theoretically, formative assessment practices directing students' attention to learning process and strategies may nurture a belief that the learning ability is malleable, and the desirable outcomes could be achieved through hard work, effort, and the use of effective strategies. Through a focus on revision and progress towards better achievement, formative assessment has the potential to encourage a growth mindset among students (Sanchez et al., 2017). The learning gains brought by formative assessment might further strengthen such a belief that appreciates and acknowledges the value of hard work, effort, and effective strategies. Past studies argued that growth mindsets can be taught (Blackwell et al., 2007; Yeager et al., 2016). Most importantly, teachers' practices influence their students' mindsets (Haimovitz & Dweck, 2017). For example, Sun (2015) reported that teachers who adopted process-oriented teaching (such as frequent questioning, providing chances for explaining and revisions) promoted a growth mindset among students.

In addition, the way students react to formative feedback is likely to be guided by their mindsets (Sperling & Shapcott, 2012). A growth mindset is likely linked to better use of feedback, while a fixed mindset might result in maladaptive responses to feedback which, in turn, might have negative impact on learning outcomes (Mangels et al., 2006). However, empirical evidence focusing on the link between formative assessment and mindset, especially with large samples, is still limited.

Culture, formative assessment, and growth mindset

Another important factor often neglected in studies on formative assessment and growth mindset is the role of culture. Culture refers to the set of values, beliefs, and traditions that influence the behaviours of a social group (Triandis, 2002). Cultural context could influence both the formative assessment strategies teachers use as well as the implicit beliefs students have about their intelligence. Though the notion of culture is very complex, most education scholars attempt to simplify this complexity by juxtaposing Eastern versus Western cultures (Li, 2002, 2003, 2005). Eastern cultural contexts are usually synonymous to Confucian heritage cultures, while Western culture is synonymous to North America and English-speaking societies.

Culture might impact assessment beliefs and practices (Yan & Brown, 2021; Teasdale & Leung, 2000) which, in turn, not only influence teachers' approaches to assessment (Davison, 2004), but also impact the ways in which students engage with assessment tasks (Carless, 2011) as well as their approaches to learning and studying (Struyven et al., 2005). Thus, the interpretation of the process and consequence of assessment needs to take the role of culture into account. Assessment principles that work well for one culture might not work in another culture if the principles are not well integrated with the local context. As Yan and Brown (2021) reported, the implementation of formative assessment, originated from the Western culture, is far less than satisfactory in Eastern classrooms where the use of public examinations for important decision-making prevails. A particular assessment practice could lead to varied consequences under different social norms. For instance, although Confucian heritage culture is believed to prioritise summative assessment at the expense of formative assessment (Wicking, 2020), preparation for high-stakes, summative assessments often lead to desirable self-regulated learning activities (Brown, 2021). What is regarded as rote learning in the West is actually a combination of memorisation and deeper levels of learning, rather than being simply equivalent to surface learning, in the East (Biggs, 1999; Kember & Wang, 2016).

Culture might also affect mindset and how it relates to key outcomes. Though growth mindsets have been found to be beneficial across different cultures (Grigorenko et al., 2001), there seems to be some cultural differences. For example, the concept of what intelligence is might vary across cultural contexts (Sternberg, 2004). Another cross-cultural difference is the way mindsets are associated with other variables in the nomological network. Studies have found that having a fixed mindset was found to be more detrimental for North American students compared to Chinese students (Burnette et al., 2012). Growth mindset was also associated with both learning/mastery and ego/performance goals among Chinese students (Chen & Wong, 2015). Though the relationship between growth mindset and mastery goal is cross-culturally universal (e.g. Burnette et al., 2012), the positive association between performance approach and growth mindset was largely absent in Western cultural contexts. For Chinese students, it might not be sufficient to merely master one's academic tasks (i.e. mastery goal) but also to demonstrate academic success to one's parents, family members, and teachers, which is more akin to performance goals.

Though there is relatively little research on how cultural forces impinge on formative assessment and growth mindsets, past empirical research has shown that culture might be an important yet neglected aspect of formative assessment and mindset. Motivational

factors that are salient in one cultural context might not be as salient in other cultural contexts (King et al., 2018; King & McInerney, 2014; 2019; Li & Yamamoto, 2020). Bernardo et al. (2021) found that cultural beliefs affect the endorsement of a growth mindset and how mindsets are related to achievement outcomes. In their study, they found that in cultures wherein people believe that there are multiple routes to achieving a given outcome, the importance of growth mindset decreases. This is perhaps due to the understanding that intelligence and effort are not the only routes to academic success. Given that growth mindset research originated in North American Western contexts, there is a need to test the role of growth mindsets in other cultural contexts, which this study aims to address by explicitly focusing on Eastern and Western cultures.

This study

This study aims to bridge the fields of formative assessment and mindset to examine the relationship between formative assessment practices, student mindset, and academic achievement using the PISA 2018 data set. To provide a clear picture of the effect of formative assessment, we examined the role of three specific formative assessment strategies that are available in the PISA 2018 data, i.e. sharing learning progressions, providing feedback, and instructional adjustments. The specific questions guiding this study are:

(1) What is the relationship between student reports of teachers' formative assessment practices, students' mindset and students' reading achievement?
(2) Does mindset mediate the relationship between formative assessment practices and reading achievement?
(3) Is there a cultural difference regarding the relationship between student reports of teachers' formative assessment practices, students' mindset and students' reading achievement?

Method

Data and participants

The study utilised the PISA 2018 data set, which consists of more than 600,000 15-year-old nationally representative students from across 79 countries/economies. We drew our data from Western English-speaking cultures (the United States, the United Kingdom, Ireland, Canada, Australia and New Zealand) and Eastern Confucian cultures (Mainland China, Hong Kong SAR, Macau SAR, Chinese Taipei, Japan and Korea) which has a total of 109,204 students. We excluded student participants with more than 5% missing values on the key variables of the study, limiting our sample to 41,632 students from the English-speaking countries/economies (51% female) and 41,208 students from Confucian countries/economies (49% female).

PISA uses a complex two-stage sampling method in recruiting student participants. This means that PISA samples various types of schools and subsequently samples the students within these schools. This process opens secondary data analysis to possible errors due to variance in (1) the probability of selecting schools and students within

schools across participating countries/economies, (2) the adjustments done in students' non-response across types of schools and characteristics, and (3) oversampling in selected types of school for national-level reporting (OECD, 2009b). The PISA dataset assigns 80 replicate weights based on Fay's version of balanced repeated replication and a final weight for each student to account for these variances and errors. Furthermore, these weight replicates reflect the clustered characteristics of the data (Rutkowski et al., 2010). The current study utilised these replicate weights to ensure more robust statistical estimates (OECD, 2009b).

Measures

PISA is a triennial assessment programme that measures students' achievement scores in three learning domains (i.e. reading, mathematics, and science), with one domain as the specific focus in every cycle. PISA 2018 focused on reading as its primary domain of assessment based on a literacy framework that operationalised reading as 'understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential and to participate in society' (OECD, 2019a, p. 28).

Aside from standardised assessment, PISA 2018 had a survey component that included demographic information and other non-cognitive contextual measures associated with reading literacy. This study examined the questionnaire items related to the students' growth mindset and perceived formative assessment. We also looked at the students' demographic details and their achievement score in reading.

Independent variables

We selected items from the PISA student questionnaire that reflect teachers' formative assessment activities and clustered them into three strategies: (1) *sharing learning targets and progress*, (2) *providing feedback*, and (3) *instructional adjustments*. Please refer to Table 1 for a summary of items, their classification, the scales used, and their internal consistency coefficients. We reversed coded sharing learning progressions to be consistent with the direction of other variables.

Dependent variable

Our study focuses on reading achievement as our key dependent variable since PISA 2018's main assessment domain is reading. We used the students' reading proficiency scores in PISA to operationalise reading achievement. These scores were scaled to have a mean of 500 and a standard deviation of 100 using the Rasch model.

The reading proficiency scores, however, were not actual student test scores but were derived values. Given the range of items to be covered within a limited testing time, PISA was administered using a complex test booklet rotation scheme to minimise the cognitive burden on students (Kaplan & Su, 2018; Rutkowski et al., 2010). In this scheme, students only answered test questions from a block of test items representing the whole range of assessments. Then, the students' proficiency scores were drawn from the population's ability distribution derived from individual students' performance on their assigned test

Table 1. Descriptive statistics and internal consistency.

		Over-all Sample		English Speaking		Confucian			
Variables	*α*	M	SD	M	SD	M	SD	Scale	Items
Reading Achievement		512.71 (1.77)	103.70 (0.86)	505.18 (2.83)	106.74 (1.24)	523.96 (1.49)	97.88 (.93)		
Growth Mindset		2.81 (.01)	.91 (0)	2.90 (.01)	.93 (.01)	2.68 (.01)	.86 (0)	1 – *Strongly Agree;* 4 – *Strong Disagree*	'Your intelligence is something about you that you can't change very much.'
Sharing Learning Progressions	.80	3.02 (.01)	.73 (0)	2.97 (.01)	.73 (.01)	3.13 (.01)	.72 (0)	1 – *Every Lesson;* 4 – *Never or Hardly Ever*	ST102Q01TA The teacher sets clear goals for our learning. ST102Q02TA The teacher asks questions to check whether we have understood what was taught. ST102Q03TA At the beginning of a lesson, the teacher presents a short summary of the previous lesson. ST102Q04TA The teacher tells us what we have to learn.
Providing Feedback	.90	2.50 (.01)	.89 (0)	2.60 (.01)	.87 (.01)	2.33 (.01)	.89 (0)	1 – *Never or almost never;* 4 – *Every lesson or almost every lesson*	ST104Q02NA The teacher gives me feedback on my strengths in this subject. ST104Q03NA The teacher tells me in which areas I can still improve. ST104Q04NA The teacher tells me how I can improve my performance.
Instructional Adjustments	.79	2.64 (.01)	079 (0)	2.62 (.01)	.78 (0)	2.67 (.01)	.80 (0)	1 – *Never or almost never;* 4 – *Every lesson or almost every lesson*	ST212Q01HA The teacher adapts the lesson to my class's needs and knowledge. ST212Q02HA The teacher provides individual help when a student has difficulties understanding a topic or task. ST212Q03HA The teacher changes the structure of the lesson on a topic that most students find difficult to understand.

Note: Based on samples from English-speaking and Confucian countries/economies (n = 109,204). Values in parentheses are standard errors.
Source: Programme for International Student Assessment (PISA) 2018.

questions (OECD, 2009b; Rutkowski et al., 2010). These individual proficiency estimates are referred to as plausible values indicating the range of possible scores that the students could reasonably have achieved in the test (OECD, 2009a, 2009b).

The PISA 2018 database provides 10 reading proficiency plausible values. Our study used these 10 plausible values in computing the estimates following the PISA data analysis manual's (OECD, 2009b) procedure for estimation using plausible values.

Mediator

The mediator variable is growth mindset. We took this variable from the students' response to PISA's single-item question, which states: 'Your intelligence is something about you that you can't change very much'. The item was rated on a 4-point Likert scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). We reverse coded the values since the item indicated a fixed mindset such that higher scores reflected a stronger growth mindset.

Covariates

We included the following demographic details as covariates: socioeconomic status, gender, repetition of a grade level, and class size. PISA derived an index of economic, social, and cultural status (ESCS) to operationalise socioeconomic status. ESCS is an index derived from 'the highest educational level of parents, the highest occupational status of parents and possessions in the home' (OECD, 2009a, p. 84). The OECD average was pegged at zero with index values within a standard deviation of one. Consistent with its previous cycles, the PISA 2018 results showed that students with low ESCS have underperformed compared with students with high ESCS. Likewise, female students have a clear performance advantage over male students. Students who had repeated a grade level have also been found to have lower performance than those who did not. PISA also reported that after accounting for student and school ESCS, class size did affect, albeit minimal, student achievement. We controlled for these covariates to account for their potential confounding effects on this study.

Analytic strategy

Missing data

We excluded cases with no available data for the study's variables of interest. Hence, the Canadian sample was not included in the study since it had no information on sharing of learning targets and progress. The rest of the samples have isolated high missing data; we assumed that these are due to PISA's complex survey design. High missing data is expected in large scale assessments, but these can be deemed missing 'by design' and missing completely at random (Grund et al., 2021; Kaplan & Su, 2018). We employed a multiple imputation method using multivariate imputation via changed equations (MICE) package in R (R Core Team, 2019; Rubin, 1996; Van Buuren & Groothuis-Oudshorn, 2011). This approach is aligned with past studies that drew on the PISA dataset (e.g. Chiu et al., 2017, 2016).

Preliminary analysis

We generated the descriptive statistics of the key variables and covariates using the *intsvy* (Caro & Biecek, 2017), *survey* (Lumley, 2004), and *svyPVpack* (Reif & Peterbauer, 2015) packages in R. These statistical software took into account the 10 plausible values, weights and the 80 balanced repeated replication values that come with the PISA dataset in the analysis. We reported the statistical estimates and their standard errors consistent with the PISA data analysis manual (OECD, 2009b). Then, we proceeded with the bivariate correlation analysis and reported the Pearson's *r* coefficient and the associated standard error due to sampling design.

We focused on effect size to determine the magnitude of association. We did not use Cohen's (1988) general guideline for analysing effect sizes since this rule of thumb has been found to be outdated and overly optimistic (Gignac & Szodorai, 2016; Schäfer & Schwarz, 2019). Effect sizes in social sciences and psychology research tend to be smaller but no less meaningful (e.g. Bosco et al., 2015; Ferguson, 2009). For heuristic purposes, we followed Bosco et al.'s (2015) finding that a $|r|$ coefficient between .1 to .25 indicate medium effect size for associations that involve attitudes and behaviours.

Primary analysis

We posited that formative assessment practices (i.e. sharing learning progressions, providing feedback, and instructional adjustments) would positively correlate with student achievement via growth mindset. To test this hypothesis, we created a structural equation model and fitted the model separately with the two sets of data–one for the English sample and another one for the Confucian sample – in R lavaan (R Core Team, 2019; Rosseel, 2012). The first set of data is from the English-speaking countries/economies: the United States, the United Kingdom, Ireland, Australia, and New Zealand. The second set of data is from the Confucian countries/economies: Mainland China, Hong Kong SAR, Macau SAR, Chinese Taipei, Japan, and Korea. We used ESCS, gender (being female), repeating a grade, and class size as covariates to account for their confounding effects. Figure 1 shows our hypothesised mediation model.

Analysis using plausible values and weights. Studies on the use of PISA dataset in secondary analysis have recommended that researchers compute individual estimates for each plausible value and aggregate them using Rubin's method to derive the final estimate (Laukaityte & Wiberg, 2017; Rubin, 2004; Rutkowski et al., 2013). The PISA data analysis manual also emphasised the general rule is to use the 80 weight replicates in the dataset to consider the sampling variances for each estimate (OECD, 2009b, p. 129). This study followed the recommended procedures in computing statistical estimates.

Structural equation model analysis. We conducted a structural equation model analysis using *lavaan* package in R (Rosseel, 2012) and followed Anderson and Gerbing (1988) two-step approach in structural equation modelling for our statistical estimates. First, we conducted a confirmatory factor analysis using a robust maximum likelihood estimator. When the sample size is large, the chi-square statistic tends to be rejected. Hence, we

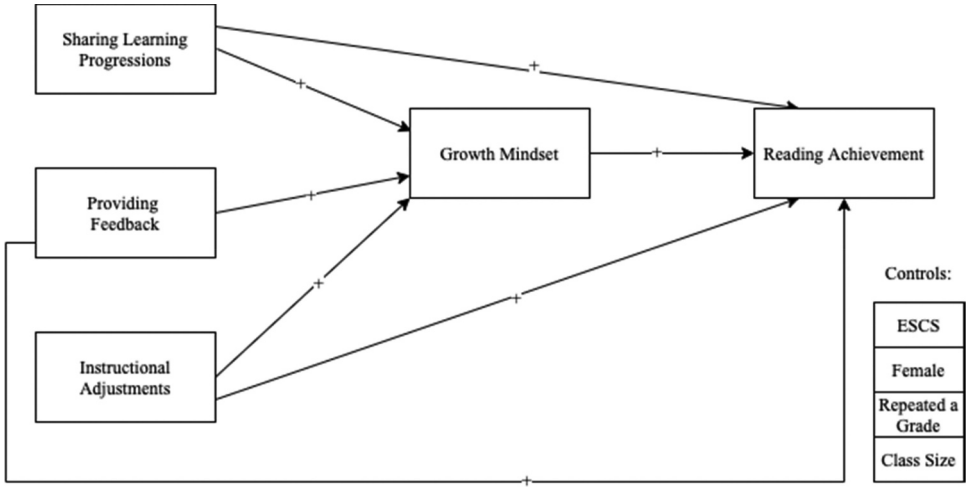

Figure 1. Conceptual framework of association between formative assessment practices and reading achievement as mediated by growth mindset.

focused on the other fit indices using the following conventional goodness of fit criteria: CLI and TLI $\geq$ 0.95 indicates a good fit while CLI/TLI $\geq$ 0.90 is acceptable; RMSEA/SRMR $\leq$ 0.05 indicates a good fit while a value $\leq$ 0.08 is acceptable (Hu & Bentler, 1995).

Next, we conducted the structural equation modelling using robust maximum likelihood to estimate the parameters. While the PISA data showed that students are nested within schools and countries/economies, we did not employ a multilevel structural equation as our parameters of interest are all within the student level. Instead, we applied the design-based method by fitting our structural equation model using the weights and replicates provided by the PISA dataset (Lumley, 2004; Stapleton et al., 2016). This procedure takes into account the nested nature of our data.

Individual structural equation models were created for each plausible value of reading achievement of our two samples. To account for PISA's sampling design in our structural equation model estimates, we relied on Oberski's (2014) *lavaan.survey* package. This package refits the structural equation model using the final weight and its 80 replicates in the PISA dataset. Then, the final parameter estimates were aggregated at the last stage (Little & Rubin, 2019; OECD, 2009b; Rubin, 2004). We computed the final standard error from the combined average sampling and imputation variance of the estimates (OECD, 2009b). Finally, we examined the significance of indirect effects of formative assessment practices on students' reading achievement via growth mindset to confirm our hypothesis.

Results

Descriptive statistics and bivariate correlations

Table 1 shows the descriptive statistics of the independent variables, mediator, and dependent variable for the English-speaking and Confucian samples. Our correlation analysis found some common results among the two sets of samples. First, sharing learning progressions, providing feedback, and instructional adjustments are strongly correlated. Second, these formative assessment variables have weak associations with growth mindset. Third, sharing learning progressions and providing feedback are weakly correlated with reading achievement. Finally, adaptive teaching is moderately correlated to reading achievement in both samples. The key contrast between the two samples is that growth mindset has a strong correlation with reading achievement among English-speaking while they are only weakly correlated with achievement in the Confucian sample. Tables 2 and 3 below show the respective summary of variable correlations for the English-speaking and Confucian countries/economies. We also ran a correlation analysis per country/economy and included the results as an appendix due to space constraints (please refer to Appendix Table A).

Structural equation model

We posited a structural equation model wherein formative assessment practices (i.e. sharing learning progressions, providing feedback, and instructional adjustments) predicted student achievement via growth mindset. Confirmatory factor analysis results showed an excellent model-data fit under robust maximum likelihood estimator for both samples. Fit indices for the English-speaking sample were $\chi^2 = 329.45$, $df = 32$, $p < .001$, scaling correction factor = 7.79; CFI = .99, TLI = 0.99, RMSEA = 0.04 with 90%

Table 2. Bivariate correlations in the English-speaking sample (n = 44,679).

	Reading Achievement	Growth Mindset	Sharing Learning Progressions	Providing Feedback	Instructional Adjustments	ESCS	Class Size	Sex
Growth Mindset	0.24***							
	(0.01)							
Sharing Learning Progressions	−0.07***	0.02***						
	(0.01)	(0.01)						
Providing Feedback	0.05***	0.05***	0.40***					
	(0.01)	(0.01)	(0.01)					
Instructional Adjustments	0.11***	0.06***	0.43***	0.55***				
	(0.01)	(0.01)	(.01)	(0.01)				
ESCS	0.34***	0.12***	−0.02***	0.07***	0.10***			
	(.02)	(0.01)	(.01)	(0.01)	(0.01)			
Class Size	0.03***	0.02**	−0.03***	−0.02***	−0.04***	−0.03***		
	(0.02)	(0.01)	(.02)	(0.02)	(0.02)	(0.02)		
Sex	0.11***	0.03***	−0.05***	−0.06***	0.00	−0.02***	0.00	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(.01)	
Repeated a Grade	−0.27***	−0.10***	0.05***	−0.01**	−0.03***	−0.18***	−0.01*	−0.04***
	(0.01)	(0.01)	(.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)

Source: Programme for International Student Assessment (PISA) 2018.
Note: Standard error estimates are enclosed in parentheses. *** $p < .001$; ** $p < .01$; *$p < .05$

Table 3. Bivariate correlations in the Confucian sample (n = 41,872).

	Reading Achievement	Growth Mindset	Sharing Learning Progressions	Providing Feedback	Instructional Adjustments	ESCS	Class Size	Female
Growth Mindset	.03***							
	(.01)							
Sharing of Learning Targets and Progress	.06***	.08***						
	(.01)	(.01)						
Providing Feedback	.03***	.07***	.42***					
	(.01)	(.01)	(.01)					
Instructional Adjustments	.10***	.06***	.45***	.58***				
	(.01)	(.01)	(.01)	(.01)				
ESCS	.21***	.01	0	.01*	.01*			
	(.01)	(.01)	(.01)	(.01)	(.01)			
Class Size	.22***	.01**	.01**	−.02***	−.02***	−.15***		
	(.03)	(.01)	(.02)	(.02)	(.02)	(.02)		
Female	.09***	−.04***	−.02***	−.11***	0	.02***	0	
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	
Repeated a Grade	−.14***	.02*	−.01	0	−.02***	−.16***	.05***	−.04***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)	(.01)

Source: Programme for International Student Assessment (PISA) 2018.
Note: Standard error estimates are enclosed in parentheses. *** $p < .001$; ** $p < .01$; *$p < .05$

Confidence Interval = [.036, .044] and SRMR = 0.02. The Confucian sample's fit indices were $\chi^2 = 1318.63$, $df = 32$, $p < .001$, scale correction factor = 3.15; CFI = .98, TLI = .97, RMSEA = 0.06 with 90% Confidence Interval = [.05, .06] and SRMR = 0.02.

After ascertaining the model-data goodness of fit, we added the structural component and proceeded with the path analysis. After refitting the structural equation to the survey design, the structural equation models consistently demonstrated good fit both for the English-speaking sample ($\chi^2 = 1027.32$, $df = 86$, $p < .001$, scaling correction factor = 7.39; CFI = .97, TLI = .96, RMSEA = 0.04 with 90% Confidence Interval = [.04, .05], and SRMR = 0.03) and the Confucian sample ($\chi^2 = 1769.76$, $df = 86$, $p < .001$, scaling correction factor = 3.31; CFI = .97, TLI = .96, RMSEA = 0.04 with 90% Confidence Interval = [.04, .05], and SRMR = 0.03). The model fit was consistent across the different SEM repetitions done for each plausible value across both English and Confucian samples so we no longer reported the values here.

Figures 2 and 3 show the structural equation models for the two samples showing the final standardised path estimates and their associated standard errors after controlling for the covariance.

Our structural equation model for the English-speaking sample showed a direct effect of growth mindset on reading achievement ($ß = .20$, SE = 1.39, $p < .001$). The relation between growth mindset and the three formative assessment strategies were not significant. *Sharing learning progressions* had direct negative and significant relation with reading achievement ($ß = -.11$, SE = 2.79, $p < .001$), while *instructional adjustments* had direct positive and significant relation with reading achievement ($ß = .14$, SE = 2.99, $p < .05$). The indirect effects of the three formative assessment strategies via growth mindset were not significant: *sharing learning progressions* ($ß = .004$, SE = .67, $p = .355$), *providing feedback* ($ß = .002$, SE = .59, $p = .615$), and *instructional adjustments* ($ß = .01$, SE = .71, $p = .198$).

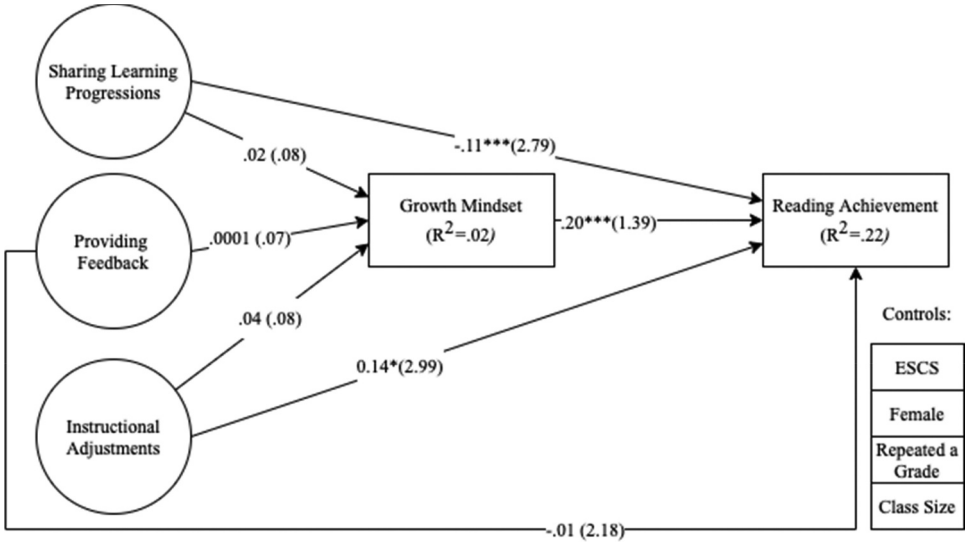

Figure 2. Structural equation model of association between dimensions of formative assessment and reading achievement via growth mindset among English-speaking countries/economies. For clarity of presentation, error terms and estimates of covariances were not shown.***$p < .001$ *$p < .05$.

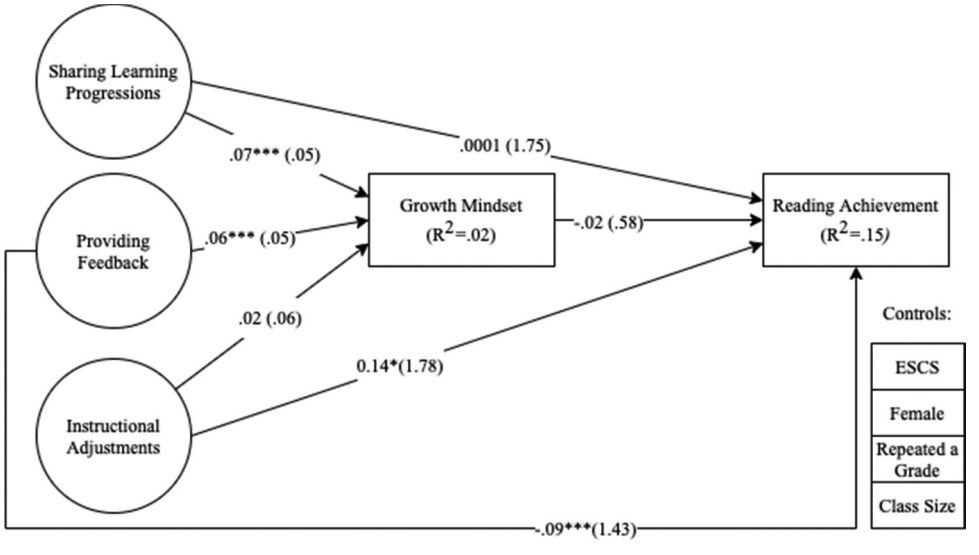

Figure 3. Structural equation model of association between dimensions of formative assessment and reading achievement via growth mindset among Confucian countries/economies. For clarity of presentation, error terms and estimates of covariances were not shown. ***$p < .001$ *$p < .05$.

The structural model for the Confucian sample did not show direct effect of growth mindset on achievement ($ß = -.02$, SE = .58, $p > .05$). Two formative assessment strategies showed significant and positive relation with growth mindset: *sharing learning progressions* ($ß = .07$, SE = .05, $p < .001$), and *providing feedback* ($ß = .06$, SE = .05, $p < .001$). *Instructional adjustments* had direct positive and significant relation with

reading achievement ($ß$ = .14, SE = 1.78, $p < .05$), while *providing feedback* had direct negative and significant relation with reading achievement ($ß$ = −.09, SE = 1.43, $p < .001$). The formative assessment strategies of *sharing learning progressions* ($ß$ = −.001, SE = .014, p = .063), *providing feedback* ($ß$ = −.001, SE = .12, p = .086), and *instructional adjustments* ($ß$ = .002, SE = .08, p = .312) did not have significant indirect effects on reading achievement via growth mindset.

Discussion

The purpose of this study was to investigate the relationship between formative assessment practices (i.e. sharing learning progressions, providing feedback, instructional adjustments), growth mindset, and academic achievement using the PISA 2018 data set. We also examined potential cultural differences regarding the relationship between Western English-speaking cultures (the United States, the United Kingdom, Ireland, Canada, Australia, and New Zealand) and Eastern Confucian cultures (Mainland China, Hong Kong SAR, Macau SAR, Chinese Taipei, Japan, and Korea).

The relationship between formative assessment and growth mindset

The correlational analysis revealed that all formative assessment strategies were positively and significantly related to growth mindset in both the East and the West. The SEM results showed that students who reported receiving more formative assessment practices were more likely to have a growth mindset only in the East but not in the West. In the East, two formative assessment strategies (sharing learning progressions and providing feedback) were positively, and weakly related to growth mindset. In the West, all the associations were weak and non-significant. These results partially support the claim that formative assessment helps develop a growth mindset among students, although the support is weak and constrained in the Eastern samples.

We speculate that the association between formative assessment strategies and growth mindset was underestimated due to the way PISA 2018 assessed these constructs. Growth mindset was assessed with only one item and it would have been better if multi-item measures that were more psychometrically robust were used. This, however, was not possible given the way PISA designed the questionnaire. Second, the growth mindset assessed in PISA was about general, but not domain-specific, growth mindset. However, formative assessment items specifically asked about teachers' strategies in the language lessons. The different levels of specificity in the measurement of the two constructs might reduce the associations. Recent studies have argued the need to measure language-specific growth mindsets although research in this area is still very nascent as past studies have mostly focused on general mindsets (see King & Trinidad, 2020; Lou & Noels, 2016, 2017 for domain-specific examples). Our study is among the first to test the relationships between formative assessment and growth mindset using a large cross-sectional sample, and the findings must be interpreted with caution. To better test the correlation between these two constructs, future studies may consider using domain-specific instruments for growth mindset, using longer multi-item surveys that are more psychometrically valid, and using more rigorous data collection plans (e.g. longitudinal, or experimental designs).

The relationship between formative assessment and reading achievement, and the mediating role of growth mindset

Instructional adjustments directly predicted reading achievement in both the East and the West. The path coefficients (β = .14 for both the Eastern and Western samples) reached the threshold of medium effect size according to Bosco et al.'s (2015) criterion and was larger than that of the other two strategies. Researchers have argued that formative assessment may not be as effective as expected if the practices (e.g. sharing learning progressions, providing feedback) are not accompanied with necessary instructional adjustments to meet students' learning needs (Briggs et al., 2012; McMillan et al., 2013). This finding is in line with these arguments and highlights that instructional adjustments are essential to effective formative assessment (Wiliam, 2010; Wiliam & Leahy, 2007). Although the effect sizes reported in previous reviews of the effect of formative assessment (e.g. Bennett, 2011; Black & Wiliam, 1998; Dunn & Mulvenon, 2009) are usually larger than the effect size we found here, a direction comparison is inappropriate due to two reasons. First, as H. Li (2016) pointed out, the effect size of formative assessment reported in previous reviews are mainly about the differences between the experimental and the control groups in experimental studies, while the relationship revealed in the present study is derived from a cross-sectional sample that were nationally representative. The tight controls in experimental settings may lead to stronger effects but the effects we found might be more ecologically valid given the large sample size and representativeness of our sample. Second, as formative assessment is an umbrella term that covers various forms of strategies, the effect of a specific strategy (i.e. instructional adjustments in this study) is not directly comparable to the effect of formative assessment in general, or another different strategies in particular.

The effects of the other two formative assessment strategies on reading achievement demonstrated cross-cultural variability. The relation between sharing learning progressions and reading achievement was non-significant in the East, but significantly negative in the West. Conversely, providing feedback was negatively related to reading achievement in the East, but the relation was not significant in the West. The weak and non-significant relations (sharing learning progressions and reading achievement in the East; providing feedback and reading achievement in the West) reiterate that sharing learning progressions or providing feedback alone might not guarantee greater achievement (Shute, 2008; Van der Kleij et al., 2015). A general consensus is that formative assessment is useful only if students make sense of the formative feedback generated from the assessment process and actively use it (Carless & Boud, 2018; Hattie & Timperley, 2007). PISA 2018 has examined teachers' formative assessment practices without gauging students' engagement (e.g. interpret and/or use the feedback) in the formative assessment process. This missing information on an important link in the formative assessment loop calls for special caution in interpreting these results.

The negative relations (sharing learning progressions and reading achievement in the West; providing feedback and reading achievement in the East) are even more counter-intuitive, but partially echoes the OECD report (2016) on the PISA 2015 data indicating that students who experienced more formative assessment had lower academic achievement, and the pattern differs across countries/economies. A possible explanation is that

'the relationship runs not from teaching strategy to student success on the items, but in the opposite direction' (OECD, 2016, p. 62), indicating that teachers are more likely to use formative-assessment strategies for those students with lower academic achievement. The results imply that teachers in the West may share more learning progressions with students with lower reading achievement, while teachers in the East tend to provide more feedback to students who are struggling with reading.

The results showed that growth mindset did not mediate the relationship between formative assessment strategies and reading achievement. Previous studies (e.g. Mangels et al., 2006; Sperling & Shapcott, 2012) argued that mindsets filter what information students attend to and alter the way they react to formative assessment and, therefore, potentially influence the effectiveness of formative assessment. Students with a growth mindset have a greater likelihood to conceive and use formative feedback in adaptive ways (Sperling & Shapcott, 2012). Our finding does not seem to support such a mediating role of growth mindset. To some extent, this result is not a surprise given that the growth mindset only weakly, or non-significantly, related to both formative assessment and reading achievement. However, this result should be interpreted with caution because it is possibly due to the way growth mindset was measured and the potential mismatch between the levels of specificity in the measurement of the constructs, as discussed earlier.

The relationship between growth mindset and reading achievement

There is also cultural difference regarding the direct relation between growth mindset and reading achievement. Growth mindset was positively related to reading achievement ($\beta = .20$, $p < .001$) in the West, corroborating past studies in terms of positive impact of growth mindset on learning-related outcomes (e.g. Blackwell et al., 2007; DeBacker et al., 2018; Yeager et al., 2016). In the East, however, growth mindset was not significantly related to reading achievement. We posit three potential reasons, from the sociocultural perspective, to explain why a growth mindset is more powerful in predicting reading achievement in the West, but less so in the East. Firstly, the non-significant finding may be due to ceiling effects of hard work in the Eastern context where Confucian notions of effort and perseverance prevail (Hau & Salili, 1991). Cross-national studies have suggested that East Asians perceive failure to be due to lack of effort more than Westerners (Yan & Gaier, 1994). Westerners, on the other hand, have been found to value ability more (Ho, 2004). A growth mindset brings about learning gains by motivating a student to learning from failure and strive harder next time. When a student has been working very hard already, for example, students in Mainland China reported spending 57 hours per week studying (OECD, 2019b), a growth mindset may not increase hours of studying or learning performance any further perhaps due to ceiling effects (Yeager & Dweck, 2020). As Zhao (2020) surmised, 'perhaps the Chinese students do not care or are not allowed to care if their intelligence is fixed or malleable because they have to study hard regardless'.

Secondly, a recent PISA report indicates that the gap in mindset between high and low achieving students was much smaller in the East compared to the OECD average suggesting that both high and low achieving students hold a more similar type of mindset (OECD, 2019b). This report is echoed by what we found in this study as the standard

deviation of growth mindset in the East is smaller than that in the West (see Table 1). The smaller variance in the Eastern growth mindset data might also contribute to a non-significant correlation.

Clearly, culture might play a role in how growth mindsets are associated with achievement outcomes as demonstrated by our findings. The role of culture was also alluded to in previous studies (Bernardo et al., 2021; Costa & Faria, 2018). However, our understanding of these cultural intricacies remains incomplete. More research is needed to delve into the cultural universality and specificity of the growth mindset construct. Although work on mindsets and culture is still at its nascency, researchers can take cue from research on other motivational and non-cognitive factors which have been found to demonstrate both culturally universal and specific patterns (Lee, 2009; Lee & Chen, 2019; Liou, 2017).

Limitations and future directions

Similar to other studies applying the secondary analysis method, this study has limitations related to the structure of the PISA data. Firstly, PISA 2018 has only one item measuring mindset. Under most conditions, multi-item scales outperform single-item scales in terms of predictive validity (Diamantopoulos et al., 2012). A single-item often has more measurement error and, hence, less statistical power, than a multi-item index, so we have less confidence in its non-significant results although retain confidence in its significant results (Kennedy, 2008). Future empirical studies are suggested to use a multiple-item instrument to produce a better measure of mindset.

Secondly, this study relied on student-reported data. Whether student reports can authentically reflect teachers' instructional practices is questionable (Beecham, 2009) and particular cautions are warranted to interpret the PISA student questionnaire data (Hopfenbeck, 2016; OECD, 2010). For example, Hopfenbeck and Maul (2011) found that a non-trivial portion of Norwegian students in the PISA 2006 might not have provided valid responses to the questionnaire, especially for students with poor academic performance. To provide a reliable depiction of teachers' teaching practice, future investigations may consider more direct and objective data sources (e.g. observation data).

Third, the formative assessment practices covered in the PISA 2018 data are not comprehensive. Some important components of formative assessment, such as self- and peer-assessment, are missing. Compared with teacher-directed formative assessment, self- and peer-assessment require a higher level of student engagement wherein they need to continuously generate internal feedback, through an internal psychological process, to achieve a better performance (Yan & Brown, 2017; Yan et al., 2020; Nicol, 2020). Thus, that subset of formative assessment may have a closer relationship with a growth mindset underscoring that learning is incremental as opposed to just getting it or not (Sanchez et al., 2017). This could be an interesting field to explore the link between formative assessment and mindset.

Finally, since the data used in this study is cross-sectional, the application of SEM is insufficient to determine causality in non-experimental designs (Beatty & Brew, 2005; Kline, 2011). Future studies could test the causal relationship between formative assessment, growth mindset, and academic performance with experimental or longitudinal designs.

Conclusion

The findings of this study provide insights into the relationship between specific formative assessment strategies, growth mindset, and reading achievement. In general, formative assessment strategies were positively related to a growth mindset only in the East, but not in the West. Comparatively, instructional adjustments, among the three formative assessment strategies, had the strongest and positive impact on reading achievement. Growth mindset, in turn, positively predicted reading achievement in the West but not in the East. The mediating effect of growth mindset in the relationship between formative assessment strategies and reading achievement was not significant.

This study has two unique contributions to the literature. First, it bridges two interrelated strands of research, i.e. formative assessment and mindset. The interplay between these two strands, as demonstrated in this study, has the potential to strengthen the synergy between external pedagogical strategies and internal psychological processes. More research along this line could provide a more sophisticated understanding of the mechanism by which formative assessment and mindset impact student learning and how such impact could be maximised. Second, the cross-cultural variability in the relationship between specific formative assessment strategies, growth mindset, and academic achievement revealed in this study highlights the need to consider cultural contexts when interpreting the results of large-scale international assessment projects, like PISA, and designing formative assessment or mindset intervention programmes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Zi Yan is an Associate Professor in the Department of Curriculum and Instruction at The Education University of Hong Kong. His publications and research interests focus on two related areas, i.e., educational assessment in the school and higher education contexts with an emphasis on student self-assessment; and Rasch measurement, in particular its application in educational and psychological research. A recent book co-edited with Lan Yang is entitled Assessment as learning: Maximising opportunities for student learning and achievement.

Ronnel B. King is an Associate Professor in the Centre for the Enhancement of Teaching and Learning at The University of Hong Kong. His research interests are on student motivation, well-being, and positive psychology/education.

Joseph Y. Haw is a doctoral student specializing in curriculum and instruction at The Education University of Hong Kong. He is also a research fellow at the Ateneo de Zamboanga University, Philippines. He has been working in an integrated basic education school in the Philippines for seven years as School President prior to his admission to the doctoral program. As an administrator, he focused on teacher and learner well-being, teacher formation, curriculum design, and organizational development. His current research focus is on teacher and student motivation.

ORCID

Zi Yan http://orcid.org/0000-0001-9305-884X

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411

Beatty, B., & Brew, C. (2005). Measuring student sense of connectedness with school: The development of an instrument for use in secondary schools. *Leading and Managing*, *11(2)*, 103–118.

Beecham, R. (2009). Teaching quality and student satisfaction: Nexus or simulacrum? *London Review of Education*, *7*(2), 135–146. https://doi.org/10.1080/14748460902990336

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678

Bernardo, A. B. I., Cai, Y., & King, R. B. (2021). Society-level social axiom moderates the association between growth mindset and achievement across cultures. *British Journal of Educational Psychology*, e12411. https://doi.org/10.1111/bjep.12411

Bibbens, T. (2018). Learning-driven data: Tracking improvement within a formative assessment cycle in English. *English in Australia*, *53*(1), 33–41. https://eric.ed.gov/?id=EJ1183095

Biggs, J. (1999). *Teaching for quality learning at university: What the student does*. 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

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, *100*(2), 431–449. https://doi.org/10.1037/a0038047

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., & Shepard, L. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, *31*(4), 13–17. https://doi.org/10.1111/j.1745-3992.2012.00251.x

Brown, G. T. L. (2021). Assessments cause and contribute to learning: If only we let them. In Z. Yan & L. Yang (Eds.), *Assessment as learning: Maximising opportunities for student learning and achievement* (pp. 38–52). Routledge.

Burgoyne, A. P., Hambrick, D. Z., & Macnamara, B. N. (2020). How firm are the foundations of mind-set theory? The claims appear stronger than the evidence. *Psychological Science*, *31*(3), 258–267. https://doi.org/10.1177/0956797619897588

Burnette, J. L., O'Boyle, E., VanEpps, E., Pollack, J., & Finkel, E. (2012). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, *139*(3), 655–701. https://doi.org/10.1037/a0029531

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

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Caro, D. H., & Biecek, P. (2017). intsvy: An R package for analyzing international large-scale assessment data. *JSS Journal of Statistical Software*, *81*(7), 1-44. https://doi.org/10.18637/jss.v081.i07

Catley, K., Lehrer, R., & Reiser, B. (2005). *Tracing a proposed learning progression for developing understanding of evolution*. Paper Commissioned for the Committee on Test Design for K-12 Science Achievement Center for Education, National Research Council.

Chen, C.-M., Chen, L.-C., & Horng, W.-J. (2019). A collaborative reading annotation system with formative assessment and feedback mechanisms to promote digital reading performance. *Interactive Learning Environments, 29*(5), 848-865. https://doi.org/10.1080/10494820.2019.1636091

Chen, -W.-W., & Wong, Y.-L. (2015). Chinese mindset: Theories of intelligence, goal orientation and academic achievement in Hong Kong students. *Educational Psychology*, *35*(6), 714–725. https://doi.org/10.1080/01443410.2014.893559

Chiu, M. M., Chow, B. W. Y., & Joh, S. W. (2017). Streaming, tracking and reading achievement: A multilevel analysis of students in 40 countries. *Journal of Educational Psychology*, *109*(7), 915–934. https://doi.org/10.1037/edu0000188

Chiu, M. M., Wing-Yin Chow, B., Mcbride, C., & Mol, S. T. (2016). Students' sense of belonging at school in 41 countries: Cross-cultural variability. *Journal of Cross-Cultural Psychology*, *47*(2), 175–196. https://doi.org/10.1177/0022022115617031

Claro, S., Paunesku, D., Dweck, C. S., Thompson, R. A., & Wilson, T. D. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences*, *113*(31), 8664–8668. https://doi.org/10.1073/pnas.1608207113

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

Costa, A., & Faria, L. (2018). Implicit theories of intelligence and academic achievement: A meta-analytic review. *Frontiers in Psychology*, *9*, 829. https://doi.org/10.3389/fpsyg.2018.00829

Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, *21*(3), 305–334. https://doi.org/10.1191/0265532204lt286oa

DeBacker, T. K., Heddy, B. C., Kershen, J. L., Crowson, H. M., Looney, K., & Goldman, J. A. (2018). Effects of a one-shot growth mindset intervention on beliefs about intelligence and achievement goals. *Educational Psychology*, *38*(6), 711–733. https://doi.org/10.1080/01443410.2018.1426833

Degol, J. L., Wang, M. Te, Zhang, Y., & Allerton, J. (2018). Do growth mindsets in math benefit females? Identifying pathways between gender, mindset, and motivation. *Journal of Youth and Adolescence*, *47*(5), 976–990. https://doi.org/10.1007/s10964-017-0739–8

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434–449. https://doi.org/10.1007/s11747-011-0300-3

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessments: The limited scientific evidence of the impact of formative assessments in education. *Practical Assessment & Research and Evaluation*, *14*(7), 1–11. https://doi.org/10.7275/jg4h-rb87

Duschl, R., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education*, *47*(2), 123–182. https://doi.org/10.1080/03057267.2011.604476

Dweck, C. S. (1996). Implicit theories as organizers of goals and behavior. In P. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: The relation of cognition and motivation to behavior* (pp. 69–90). New York: Guilford.

Dweck, C. S. (2015, September 22). *Revisits the 'Digital Nomad Life Lessons & Mindset'*.

Dweck, C. S. (2017). *Mindset: Changing the way you think to fulfil your potential*. Robinson.
Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538. https://doi.org/10.1037/a0015808
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
Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069
Grigorenko, E. L., Geissler, P. W., Prince, R., Okatcha, F., Nokes, C., Kenny, D. A., Bundy, D. A., & Sternberg, R. J. (2001). The organisation of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavioral Development*, *25*(4), 367–378. https://doi.org/10.1080/01650250042000348
Grund, S., Lüdtke, O., & Robitzsch, A. (2021). On the treatment of missing data in background questionnaires in educational large-scale assessments: An evaluation of different procedures. *Journal of Educational and Behavioral Statistics*, *46*(4), 430-465. https://doi.org/10.3102/1076998620959058
Guo, W. Y., & Yan, Z. (2019). Formative and summative assessment in Hong Kong primary schools: Students' attitudes matter. *Assessment in Education: Principles, Policy & Practice, 26*(6), 675-699. https://doi.org/10.1080/0969594X.2019.1571993
Haimovitz, K., & Dweck, C. S. (2017). The origins of children's growth and fixed mindsets: New research and a new proposal. *Child Development*, *88*(6), 1849–1859. https://doi.org/10.1111/cdev.12955
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487
Hau, K.-T., & Salili, F. (1991). Structure and semantic differential placement of specific causes: Academic causal attributions by Chinese students in Hong Kong. *International Journal of Psychology*, *26*(2), 175–193. https://doi.org/10.1080/00207599108247885
Ho, I. T. (2004). A comparison of Australian and Chinese teachers' attributions for student problem behaviors. *Educational Psychology*, *24*(3), 375–391. https://doi.org/10.1080/0144341042000211706
Hopfenbeck, T. N. (2016). The power of PISA – Limitations and possibilities for educational research. *Assessment in Education: Principles, Policy & Practice*, *23*(4), 423–426. https://doi.org/10.1080/0969594X.2016.1247518
Hopfenbeck, T. N., & Maul, A. (2011). Examining evidence for the validity of PISA learning strategy scales based on student response processes. *International Journal of Testing*, *11*(2), 95–121. https://doi.org/10.1080/15305058.2010.529977
Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Sage Publications, Inc.
Kaplan, D., & Su, D. (2018). On imputation for planned missing data in context questionnaires using plausible values: A comparison of three designs. *Large-Scale Assessments in Education*, *6*(1), 6. https://doi.org/10.1186/s40536-018-0059-9
Kember, D., & Wang, D. (2016). Why do Chinese students out-perform those from the West? Do approaches to learning contribute to the explanation? *Cogent Education*, *3*(1), 1–15. https://doi.org/10.1080/2331186X.2016.1248187
Kennedy, P. (2008). *Guide to econometrics* (6th ed.). Wiley–Blackwell.
King, R. B. (2012). How you think about your intelligence influences how adjusted you are: Implicit theories and adjustment outcomes. *Personality and Individual Differences, 53*(5), 705–709. https://doi.org/10.1016/j.paid.2012.05.031
King, R. B. (2017). A fixed Mindset leads to negative affect: The relations between implicit theories of intelligence and subjective well-being. *Zeitschrift Fur Psychologie*, *225*(2), 137–145. https://doi.org/10.1027/2151-2604/a000290

King, R. B., & McInerney, D. M. (2014). Culture's consequences on student motivation: Capturing cross-cultural universality and variability through personal investment theory. *Educational Psychologist*, *49*(3), 175–198. https://doi.org/10.1080/00461520.2014.926813

King, R. B., & McInerney, D. M. (2019). Family-support goals drive engagement and achievement in a collectivist context: Integrating etic and emic approaches in goal research. *Contemporary Educational Psychology, 58*, 338–353. https://doi.org/10.1016/j.cedpsych.2019.04.003

King, R. B., McInerney, D. M., & Pitliya, R. J. (2018). Envisioning a culturally imaginative educational psychology. *Educational Psychology Review*, *30*(3), 1031–1065. https://doi.org/10.1007/s10648-018-9440-z

King, R. B., & Trinidad, J. E. (2020). Growth mindset predicts achievement only among rich students: examining the interplay between mindset and socioeconomic status. *Social Psychology of Education*, *24*(3), 635–652. https://doi.org/10.1007/s11218-021-09616-z

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Laukaityte, I., & Wiberg, M. (2017). Using plausible values in secondary analysis in large-scale assessments. *Communications in Statistics - Theory and Methods*, *46*(22), 11341–11357. https://doi.org/10.1080/03610926.2016.1267764

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences*, *19*(3), 355–365. https://doi.org/10.1016/j.lindif.2008.10.009

Lee, J., & Chen, M. (2019). Cross-country predictive validities of non-cognitive variables for mathematics achievement: evidence based on TIMSS 2015. *EURASIA Journal of Mathematics, Science and Technology Education*, *15*(8), em1725. https://doi.org/10.29333/ejmste/106230

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

Li, J. (2002). A cultural model of learning: Chinese "Heart and mind for wanting to learn." Journal of Cross-Cultural Psychology, 33(3), 248–269. https://doi.org/10.1177/0022022102033003003

Li, J. (2003). U.S and Chinese cultural beliefs about learning. *Journal of Educational Psychology*, *95*(2), 258–267. https://doi.org/10.1037/0022-0663.95.2.258

Li, J. (2005). Mind or Virtue. *Current Directions in Psychological Science*, *14*(4), 190–194. https://doi.org/10.1111/j.0963-7214.2005.00362.x

Li, J., & Yamamoto, Y. (2020). Western and East Asian sociocultural learning models: Evidence from cross-cultural and immigrant research. *Asian Journal of Social Psychology*, *23*(2), 174–186. https://doi.org/10.1111/ajsp.12384

Li, X., Han, M., Cohen, G. L., & Markus, H. R. (2021). Passion matters but not equally everywhere: Predicting achievement from interest, enjoyment, and efficacy in 59 societies. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(11). https://doi.org/10.1073/PNAS.2016964118

Liou, P. Y. (2017). Profiles of adolescents' motivational beliefs in science learning and science achievement in 26 countries: Results from TIMSS 2011 data. *International Journal of Educational Research*, *81*, 83–96. https://doi.org/10.1016/j.ijer.2016.11.006

Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (Vol. 793). John Wiley & Sons.

Lou, N. M., & Noels, K. A. (2016). Changing language mindsets: Implications for goal orientations and responses to failure in and outside the second language classroom. *Contemporary Educational Psychology*, *46*, 22–33. https://doi.org/10.1016/j.cedpsych.2016.03.004

Lou, N. M., & Noels, K. A. (2017). Measuring language mindsets and modeling their relations with goal orientations and emotional and behavioral responses in failure situations. *Modern Language Journal*, *101*(1), 214–243. https://doi.org/10.1111/modl.12380

Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, *9*(8), 1–19. https://doi.org/10.18637/jss.v009.i08

Mangels, J. A., Butterfiels, B., Lamb, J., Good, C., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective Neuroscience*, *1*(2), 75–86. https://doi.org/10.1093/scan/nsl013

McCutchen, K. L., Jones, M. H., Carbonneau, K. J., & Mueller, C. E. (2016). Mindset and standardized testing over time. *Learning and Individual Differences*, *45*, 208–213. https://doi.org/10.1016/j.lindif.2015.11.027

McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the effect of formative assessment on student achievement: So much more is needed. *Practical Assessment, Research & Evaluation*, *18*, Article 2. https://doi.org/10.7275/tmwm-7792

Nicol, D. (2020). The power of internal feedback: Exploiting natural comparison processes. *Assessment & Evaluation in Higher Education, 46*(5), 756-778. https://doi.org/10.1080/02602938.2020.1823314

Oberski, D. (2014). Lavaan.survey: An R package for complex survey analysis of structural equation models. *Journal of Statistical Software*, *57*(1), 1–27. https://doi.org/10.18637/jss.v057.i01

OECD. (2009a) . *Education at a glance 2009: OECD indicators*. OECD.

OECD. (2009b). *PISA data analysis manual SPSS* (2nd ed.). https://doi.org/10.1787/9789264056275-en

OECD. (2010). *PISA 2009 results: Learning to learn student engagement, strategies and practices* (Vol. III).

OECD. (2016). *How teachers teach and students learn: Successful strategies for school* (OECD Education Working Paper No. 130). https://doi.org/10.1787/19939019

OECD. (2019a). *PISA 2018 Assessment and analytical framework*. Retrieved July 15, 2020, from https://doi.org/10.1787/b25efab8-en

OECD. (2019b). *PISA 2018 Insights and interpretations*. https://www.oecd.org/pisa/PISA%202018%20Insights%20and%20Interpretations%20FINAL%20PDF.pdf

Petscher, Y., Al Otaiba, S., Wanzek, J., Rivas, B., & Jones, F. (2017). The relation between global and specific mindset with reading outcomes for elementary school students. *Scientific Studies of Reading*, *21*(5), 376–391. https://doi.org/10.1080/10888438.2017.1313846

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Reif, M., & Peterbauer, J. (2015). *svyPVPack*.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, *91*(434), 473–489. https://doi.org/10.1080/01621459.1996.10476908

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81). John Wiley & Sons.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, *39*(2), 142–151. https://doi.org/10.3102/0013189X10363170

Rutkowski, L., von Davier, M., & Rutkowski, D. (2013). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. CRC Press.

Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H., & Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. *Journal of Educational Psychology*, *109*(8), 1049–1066. https://doi.org/10.1037/edu0000190

Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, *10*(APR), 813. https://doi.org/10.3389/fpsyg.2019.00813

Schleider, J. L., Abel, M. R., & Weisz, J. R. (2015). Implicit theories and youth mental health problems: A random-effects meta-analysis. *Clinical Psychology Review*, *35*, 1–9. https://doi.org/10.1016/j.cpr.2014.11.001 .

Schroder, H. S., Moran, T. P., Donnellan, M. B., & Moser, J. S. (2014). Mindset induction effects on cognitive control: A neurobehavioral investigation. *Biological Psychology*, *103*(1), 27–37. https://doi.org/10.1016/j.biopsycho.2014.08.004

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychological Science*, *29*(4), 549–571. https://doi.org/10.1177/0956797617739704

Smiley, P. A., Buttitta, K. V., Chung, S. Y., Dubon, V. X., & Chang, L. K. (2016). Mediation models of implicit theories and achievement goals predict planning and withdrawal after failure. *Motivation and Emotion*, *40*(6), 878–894. https://doi.org/10.1007/s11031-016-9575-5

Sperling, C., & Shapcott, S. (2012). Fixing students' fixed mindsets: Paving the way for meaningful assessment. *Legal Writing: Journal of Legal Writing Inst., 18*, 39–84. Retrieved from http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/jlwriins18§ion=6 .

Stapleton, L. M., McNeish, D. M., & Yang, J. S. (2016). Multilevel and single-level models for measured and latent variables when data are clustered. *Educational Psychologist*, *51*(3–4), 317–330. https://doi.org/10.1080/00461520.2016.1207178

Sternberg, R. J. (2004). Culture and Intelligence. *American Psychologist*, *59*(5), 325–338. https://doi.org/10.1037/0003-066X.59.5.325

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in Higher Education: A review. *Assessment & Evaluation in Higher Education*, *30*(4), 325–341. https://doi.org/10.1080/02602930500099102

Sun, K. L. (2015). *There's no limit: Mathematics teaching for a growth mindset*. Doctoral dissertation, Stanford University.

Teasdale, A., & Leung, C. (2000). Teacher assessment and psychometric theory: A case of paradigm crossing? *Language Testing*, *17*(2), 163–184. https://doi.org/10.1177/026553220001700204

Triandis, H. C. (2002). *Subjective culture, Unit 2*. Online Readings in Psychology and Culture. http://scholarworks.gvsu.edu/orpc/vol2/iss2/6

Van Buuren, S., & Groothuis-Oudshorn, K. (2011). MICE: Multivariate imputation by chained equation. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes. *Review of Educational Research*, *85*(4), 475–511. https://doi.org/10.3102/0034654314564881

Wicking, P. (2020). Formative assessment of students from a Confucian heritage culture: Insights from Japan. *Assessment & Evaluation in Higher Education*, *45*(2), 180–192. https://doi.org/10.1080/02602938.2019.1616672

Wiliam, D., & Leahy, S. (2007). A theoretical foundation for formative assessment. In J. H. McMillan (Ed.), *Formative assessment: Theory into practice* (pp. 29–42). Teachers College Press.

Wiliam, D., & Thompson, M. (2008). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer Ed.) *The future of assessment: Shaping teaching and learning* (pp. 53–82). Erlbaum.

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

Yan, W., & Gaier, E. L. (1994). Causal attributions for college success and failure. *Journal of Cross-Cultural Psychology*, *25*(1), 146–158. https://doi.org/10.1177/0022022194251009

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education, 42*(8), 1247-1262. https://doi.org/10.1080/02602938.2016.1260091

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., Chiu, M. M., & Ko, P. Y. (2020). Effects of self-assessment diaries on academic achievement, self-regulation, and motivation. *Assessment in Education: Principles, Policy & Practice, 27* (5), 562-583. https://doi.org/10.1080/0969594X.2020.1827221

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302–314. https://doi.org/10.1080/00461520.2012.722805

Yeager, D. S., & Dweck, C. S. (2020). What can be learned from growth mindset controversies? *American Psychologist*, *75*(9), 1269–1284. https://doi.org/10.1037/amp0000794

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., Lee, H. Y., O'Brien, J., Flint, K., Roberts, A., Trott, J., Greene, D., Walton, G. M., & Dweck, C. S. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology*, *108*(3), 374–391. https://doi.org/10.1037/edu0000098

Zhao, Y. (2020). *PISA peculiarities: Why doesn't growth mindset work for Chinese students?* Retrieved from https://nepc.colorado.edu/blog/pisa-peculiarities

Appendix

Table A. Descriptive statistics and bivariate correlations by countries/economies.

CNT	n	Reading Achievement		Sharing Learning Progressions (SLP)		Providing Feedback (PFB)		Instructional Adjustments (IAD)		Growth Mindset		Formative Assessment Correlation with Reading Achievement			Formative Assessment's Correlation with Growth Mindset			Growth Mindset's correlation with Reading Achievement
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	SLP	PFB	IAD	SLP	PFB	IAD	
A. English-speaking Countries/economies																		
Australia	14,273	502.63 (1.63)	108.66 (0.9)	2.89 (0.01)	0.76 (0.01)	2.61 (0.01)	0.84 (0)	2.7 (0.01)	0.78 (0)	2.87 (0.01)	0.87 (0)	−.06*** (.01)	.05*** (.01)	.11*** (0.01)	.02 (0.01)	.05*** (0.01)	.06*** (0.01)	.23*** (0.01)
United Kingdom	13,818	503.93 (2.58)	100.21 (1.27)	2.91 (0.01)	0.73 (0.01)	2.74 (0.02)	0.81 (0.01)	2.71 (0.01)	0.79 (0.01)	2.9 (0.01)	0.89 (0.01)	−.07*** (.02)	.05*** (.02)	.13*** (0.02)	.08*** (0.01)	.10*** (0.01)	.12*** (0.01)	.18*** (0.01)
Ireland	5577	518.08 (2.24)	90.7 (1.04)	2.73 (0.02)	0.74 (0.01)	2.56 (0.02)	0.82 (0.01)	2.52 (0.01)	0.75 (0.01)	2.97 (0.01)	0.83 (0.01)	−.10*** (.02)	.04** (0.02)	.09*** (0.02)	.09*** (0.01)	.08*** (0.01)	.11*** (0.01)	.17*** (0.02)
New Zealand	6173	505.73 (2.04)	106.3 (1.34)	2.86 (0.01)	0.73 (0.01)	2.72 (0.01)	0.81 (0.01)	2.72 (0.01)	0.77 (0.01)	2.84 (0.01)	0.88 (0.01)	−.13*** (.02)	.00 (0.02)	.12*** (0.01)	−.01 (0.01)	.04** (0.01)	.08*** (0.01)	.31*** (0.01)
United States	4838	505.35 (3.57)	107.89 (1.57)	2.99 (0.01)	0.72 (0.01)	2.57 (0.02)	0.88 (0.01)	2.59 (0.01)	0.78 (0.01)	2.9 (0.01)	0.94 (0.01)	−.06*** (0.02)	.05** (0.02)	.10*** (0.02)	.01 (0.02)	.04* (0.02)	.05*** (0.02)	.26*** (0.02)
B. Confucian Countries/economies																		
Hong Kong SAR	6037	524.28 (2.73)	99.48 (1.49)	2.85 (0.02)	0.76 (0.01)	2.39 (0.02)	0.76 (0.01)	2.55 (0.01)	0.75 (0.01)	2.39 (0.02)	0.85 (0.01)	.02 (0.02)	.02 (0.02)	.08*** (0.02)	.07*** (0.01)	.07*** (0.02)	.06*** (.01)	−0.03 (.02)
Japan	6109	503.86 (2.67)	97.12 (1.68)	3.03 (0.02)	0.71 (0.01)	2.05 (0.02)	0.81 (0.01)	2.48 (0.02)	0.75 (0.01)	2.76 (0.01)	0.84 (0.01)	.05*** (0.02)	−.03* (0.02)	.02 (0.01)	.07*** (.01)	.07*** (0.01)	.06*** (0.01)	.14*** (0.02)
Korea	6650	514.05 (2.94)	102 (1.65)	3.21 (0.01)	0.7 (0.01)	2.46 (0.02)	0.98 (0.01)	2.87 (0.02)	0.8 (0.01)	2.59 (0.01)	0.86 (0.01)	.10*** (0.02)	−.04*** (0.02)	.10*** (0.02)	.10*** (0.01)	.10*** (0.01)	.12*** (0.01)	.08*** (0.01)
Macau SAR	3775	525.12 (1.23)	92.11 (1.07)	2.91 (0.01)	0.68 (0.01)	2.16 (0.01)	0.71 (0.01)	2.36 (0.01)	0.7 (0.01)	2.48 (0.01)	0.88 (0.01)	.04* (0.02)	−.03* (0.02)	0.01 (0.02)	.10*** (0.02)	.09*** (0.02)	.09*** (0.02)	.07*** (0.01)
Mainland China	12,058	555.24 (2.75)	87.23 (1.72)	3.24 (0.02)	0.69 (0.01)	2.52 (0.02)	0.88 (0.01)	2.82 (0.02)	0.81 (0.01)	2.65 (0.01)	0.89 (0.01)	−.02 (0.02)	.01* (0.01)	.09*** (0.02)	.10*** (0.01)	.10*** (0.02)	.07*** (0.02)	−.11*** (0.02)
Chinese Taipei	7243	502.6 (2.84)	101.73 (1.46)	2.95 (0.01)	0.73 (0.01)	2.58 (0.01)	0.82 (0.01)	2.62 (0.01)	0.78 (0.01)	2.67 (0.01)	0.85 (0.01)	.01 (0.01)	.07*** (0.01)	.05*** (0.01)	.08*** (0.01)	.09*** (0.01)	.07*** (0.01)	.07*** (0.01)

Note: Standard Errors are enclosed in parentheses. *** $p < .001$; **$p < .01$; *$p < .05$; B-S-Z-J: Beijing-Shanghai-Zhejiang-Jiangsu