Assessment & Evaluation in Higher Education

ISSN: 0260-2938 (Print) 1469-297X (Online) Journal homepage: https://www.tandfonline.com/loi/caeh20

Self-assessment in the process of self-regulated learning and its relationship with academic achievement

Zi Yan

To cite this article: Zi Yan (2020) Self-assessment in the process of self-regulated learning and its relationship with academic achievement, Assessment & Evaluation in Higher Education, 45:2, 224-238, DOI: 10.1080/02602938.2019.1629390

To link to this article: https://doi.org/10.1080/02602938.2019.1629390

Published online: 26 Jun 2019.

Submit your article to this journal

Article views: 554

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=caeh20

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
2020, VOL. 45, NO. 2, 224–238
https://doi.org/10.1080/02602938.2019.1629390

Routledge
Taylor & Francis Group

Check for updates

Self-assessment in the process of self-regulated learning and its relationship with academic achievement

Zi Yan

The Education University of Hong Kong, Hong Kong

ABSTRACT

Despite the theoretical argument and empirical evidence regarding the impact of self- assessment on academic achievement and self-regulated learning (SRL), the mechanism for this impact is understudied. The present study aimed to investigate the characteristics of self-assessment practices at different SRL phases and its relationship with academic achievement. Using a course assignment as the learning task, sixty-three students enrolled in a one-year master programme in a teacher education institute responded to an instrument assessing their self-assessment practices (including four self-assessment actions) at the SRL Preparatory, Performance and Appraisal phases of the task. Their final scores of the assignment were also collected. The results showed that self-assessment is a fundamental skill for SRL and occurs at each SRL phase with different patterns. Autoregressive relationships were found for all self-assessment actions between different SRL phases. Self-reflection at Performance phase was found to influence feedback seeking at Appraisal phase. Self-directed feedback seeking through monitoring at Performance phase was the strongest and positive predictor of academic achievement; and achievement had negative impact on all self-assessment actions at Appraisal phase. This study may assist educators and researchers to better understand the complexity of self-assessment in relation to learning process.

KEYWORDS

self-assessment; self-regulated learning; process; academic achievement

Introduction

Self-assessment has been increasingly used as a necessary learning and assessment strategy in higher education to develop students as independent and lifelong learners (Boud 1995; Andrade and Du 2007; Yan and Brown 2017; Bourke 2018). It is widely reported that self-assessment is positively related to academic achievement in review studies (e.g., Topping 2003; Brown and Harris 2013) and recent empirical studies (e.g., Kissling and O'Donnell 2015; Jay and Owen 2016). There is also a strong claim in the literature about the strong connection between self-assessment and self-regulated learning (SRL) (e.g. Andrade 2010; Panadero, Andrade and Brookhart 2018). The contribution of self-assessment to students' academic achievement was suggested to be achieved through the use of self-regulation strategies, such as clarifying the learning goals, monitoring the learning process and inspiring self-reflection (Brown and Harris 2013). However, both self-assessment and SRL are complicated processes, rather than one-off behaviours. Although the positive effect of self-assessment on academic achievement and SRL is well documented, the intertwined relationship between self-assessment and SRL processes as well as the

CONTACT Zi Yan zyan@eduhk.hk

© 2019 Informa UK Limited, trading as Taylor & Francis Group

mechanism for the effect of self-assessment are of particular importance but unfortunately understudied (Panadero, Jonsson and Botella 2017). This study contributes to the understanding of such mechanisms by depicting characteristics of self-assessment practices at different SRL phases and their relationships with students' academic achievement.

The self-assessment process and actions

Self-assessment may cover a wide range of activities, from simple self-grading/self-rating without further reflection to having students undertake comprehensive reflection on their own performance (Brown and Harris 2013). In practice, self-assessment can be used as a learning strategy to support student learning or an alternative assessment method for summative purposes (Panadero, Brown and Strijbos 2016; Yan 2016, 2018b). However, using self-assessment for summative purposes has been seriously doubted due to the concerns about the accuracy of self-assessment (Brown, Andrade and Chen 2015). The formative use of self-assessment, as a learning strategy, may have great educational merit for student learning (Boud 1999; Yan and Brown 2017). It has been increasingly agreed that self-assessment is a process that can be analyzed and learned, rather than a one-off terminal action (Boud 1995; Andrade, Du and Wang 2008; Yan 2016; Yan and Brown 2017). Self-assessment conducted in different forms may share a similar mental process. For the sake of supporting learning and teaching, an understanding of the possible universal mental processes of self-assessment may be more important than selection of any particular form of self-assessment.

Researchers have attempted to break down the self-assessment process (e.g., Ross 2006; McMillan and Hearn 2008; Sargeant et al. 2010; Fastré et al. 2012). A recent attempt was made in the Hong Kong context. Based on interviews with a sample of Hong Kong university students, Yan and Brown (2017) defined self-assessment as "a process during which students collect information about their own performance, evaluate and reflect on the quality of their learning process and outcomes according to selected criteria to identify their own strengths and weaknesses (p. 1248)". They further proposed a cyclical self-assessment process model which encompasses three common and sequenced actions (i.e., determining assessment criteria, self-directed feedback seeking and self-reflection). In the self-assessment process, the first step is to determine the assessment criteria. Students then engage in self-directed feedback seeking to collect information regarding their own performance from various sources which could be external and/or internal. External feedback could be obtained by inquiry and monitoring approaches. Inquiry approach refers to asking for feedback directly from relevant people (e.g. teachers, peers, parents); monitoring means that students obtain feedback from interactions with learning evidence (e.g. comments on assignments, past tests, reference books). Internal feedback refers to implicit information provided by "the self" (e.g. emotions, feelings, physical sensations and internal states). Based on the feedback, students then reflect on the quality of the process and product of their learning with the purpose of identifying their own strengths and weaknesses. Following this initial set of processes, a self-assessment judgement is made and this judgement is subjected to further calibration. As Yan and Brown's (2017) model explicitly identified the common actions students conduct and how they do so within the process of self-assessment, the present study abides by this definition and process model of self-assessment.

Self-assessment and self-regulated learning

Self-regulated learning (SRL) is defined as "self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals" (Zimmerman 2000, p. 14). Many SRL models have been developed since its first launch (Panadero 2017). These models present different perspectives on SRL, but, more noticeably, all of the models conceptualise SRL as a

cyclical process consisting of different phases and sub-processes. Puustinen and Pulkkinen (2001) compared five major SRL models in a theoretical review. They concluded that, even if the terminology used varied, these models shared three common phases: (a) Preparatory, which consists of activities relevant to task analysis, goal setting and planning; (b) Performance, in which the actual task is completed with the progress of performance being monitored; and (c) Appraisal, including the evaluation and reflection of outcomes. In a more recent review of SRL models, Panadero (2017) echoed this classification. As the terminology proposed by Puustinen and Pulkkinen (2001) summarized the essentials of different SRL models and is easy to understand, the present study adopted this terminology.

Despite the strong theoretical relationship between self-assessment and SRL, the results of empirical studies in this field were far less than conclusive. For instance, in a meta-analytic review, Panadero et al. (2017) reported that self-assessment interventions demonstrated a positive impact on students' SRL strategies with the effect size ranging from small to medium depending on the measurements used for SRL. Brown and Harris (2013) reviewed 23 studies and concluded that the empirical support to the relationship between self-assessment and SRL was not strong yet and more research was needed to explore the benefits associate with self-assessment.

With regard to the relationship between self-assessment practices and different SRL phases, most of the models proposed also conformed to common sense, that self-assessment occurs at the last phase (i.e., when the learning task is completed). However, if self-assessment is essentially the act of monitoring and reflecting, it is expected to occur at each phase of the SRL process (Panadero and Alonso-Tapia 2013) even though the target of the monitoring and reflecting could be varied. In this sense, self-assessment is not only one aspect of SRL but also contributes to it by acting as a fundamental skill underlying SRL (Harris and Brown 2018; Panadero et al. 2018). However, this has remained a theoretical assumption and little is known about the intertwined relationship between self-assessment practices and SRL phases.

In discussing the process of SRL, Greene and Azevedo (2007) posed a fundamental question: whether all cognitive evaluations are the same when they occur at different phases of the learning process? A similar question is also applicable to self-assessment. If self-assessment occurs at all the three SRL phases, will it follow the same pattern or demonstrate different characteristics at different phases? No empirical study has touched on this topic. Theoretically, self-assessment at different phases might target different learning products and rely on different feedback sources and, therefore, assume different characteristics. At Preparatory phase, (predictive) self-assessment facilitates the early identification of personal and environmental resources and is helpful to set a reasonable learning goal and propose appropriate learning strategies; at Performance phase, (concurrent) self-assessment can monitor the learning process, promote self-correction, and ensure the learning activities are headed in the right direction; at Appraisal phase, (reflective) self-assessment can reflect on the learning outcomes and identify their own strengths and weaknesses as well as the directions for future learning.

At each SRL phase, the four self-assessment actions, as described in Yan and Brown's (2017) process model of self-assessment, do not necessarily happen simultaneously. Self-reflection is usually conducted based on the feedback obtained through the other three self-directed feedback seeking actions. Thus, *seeking external feedback through monitoring* (SEFM), *seeking external feedback through inquiry* (SEFI), and *seeking internal feedback* (SIF) are supposed to have an impact on *self-reflection* (SR). Across different SRL phases, there should be autoregressive impact for each of the four self-assessment actions. That is, each self-assessment action influences the same action at the next SRL phase. SR is expected to have an impact on the other three actions at the next phase. In addition, if the academic achievement (i.e., the assignment score) was released between the Performance and Appraisal phases, then all the four self-assessment actions at Performance phase should have an impact on academic achievement; and academic achievement should have an impact on the four self-assessment actions at Appraisal phase. The hypothesised model is presented in Figure 1.

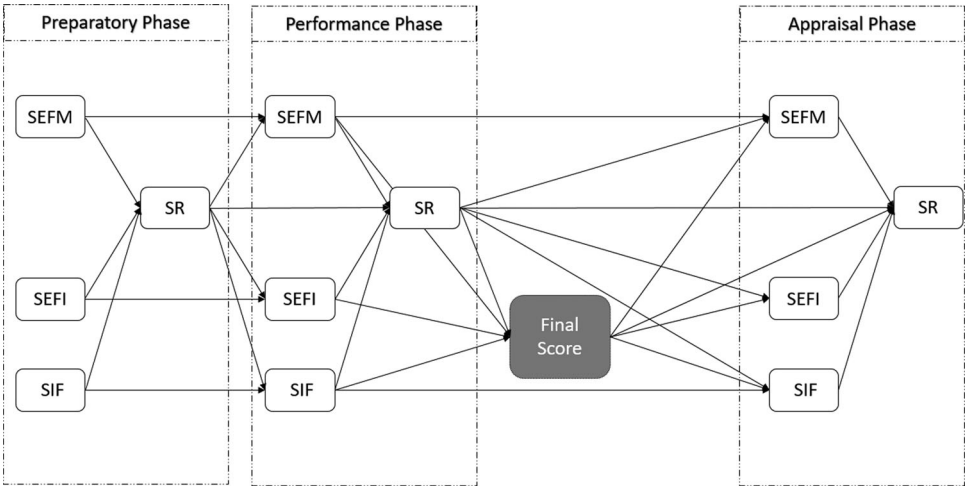

Figure 1. The hypothesised model.
Note: SEFM: seeking external feedback through monitoring, SEFI: seeking external feedback through inquiry, SIF: seeking internal feedback, SR: self-reflection

Thus, the research objective of this study was to investigate the characteristics of self-assessment practices at the three SRL phases and their relationships with academic achievement. The specific research questions are: (1) What is the pattern of self-assessment practices at each SRL phase (Preparatory, Performance and Appraisal); and (2) What are the relationships between self-assessment practices at different SRL phases and students' academic achievement?

Methods

Participants

The participants of this study consisted of 98 students who enrolled in a one-year master programme in a teacher education institute in Hong Kong. The participants were recruited from a 13-week course about assessment taught by the author. Participants attended the course in semester 1 (N = 50) or semester 2 (N = 48) in academic year 2017-2018. Participants were invited to complete three rounds of a survey. A total of 63 participants (representing a return rate of 64%) responding to two or more rounds of survey were included in the current analysis.

Design

Ethical approval was obtained from the university before undertaking the study. All participants were informed that their participation was voluntary and that they could withdraw from the study at any time. A repeated measures design was applied. In order to match responses across multiple survey administrations, students created their own 6-digit ID codes (i.e., the date and month of birth plus the last two digits of their mobile number) at the first survey which they used in each subsequent survey. Hence, participants' responses could be tracked without identifying personal information. All surveys were administered via an online method (i.e., google form).

Using the last assignment (accounts for 50 marks out of 100) of the course as the learning task, this study provided a learning scenario that consisted of three SRL phases according to Puustinen and Pulkkinen's (2001) classification. Data collection followed the procedure as presented in Figure 2. Although all assessment tasks and rubrics with which the teacher evaluated the assignments were announced to students at the beginning of the course, the detailed

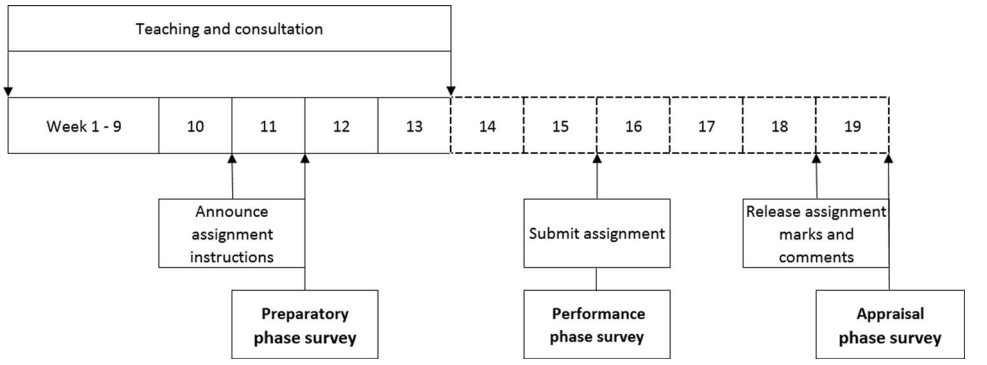

Figure 2. **The data collection procedure.**

instructions for the last assignment were provided at the end of the 10th week. Preparatory phase referred to the period when students analysed the task and set a working plan. Students were asked to make a plan (including the schedule, the method, and the resources) according to the instructions for completing this assignment. Students were encouraged to write down the plan, but it was not compulsory. The plan served as the target of self-assessment for Preparatory phase. One week after the announcement of the instructions, the first survey (for Preparatory phase) was administered. Performance phase was defined as the period when students were doing their assignment. Right after the submission, the second survey (for Performance phase) was administered. Appraisal phase refers to the period when students evaluated and reflected on the learning outcomes (the returned assignment with the final score and the teacher's written comments). One week after the release of the final scores and comments, the third survey (for Appraisal phase) was administered.

Instrument

A modified version of the Self-assessment Practice Scale (SaPS) (Yan 2018c) was used in this study. The SaPS was developed according to Yan and Brown's (2017) cyclical self-assessment process model with the aim of assessing concrete self-assessment actions. The original SaPS contains 20 items in four subscales, including *seeking external feedback through monitoring* (SEFM), *seeking external feedback through inquiry* (SEFI), *seeking internal feedback* (SIF) and *self-reflection* (SR). A six-point Likert-type scale (1 = strongly disagree, 2 = disagree, 3 = slightly disagree, 4 = slightly agree, 5 = agree, 6 = strongly agree) was applied. In this study, self-assessment targeted at a specific task, i.e., completing an assignment. Consequently, two items (including "I check whether I have mastered the course content by doing extra exercises" from the subscale SEFM, and "When I do exercise, I look at what I got wrong or did poorly on to guide me as to what I should learn next" from the subscale SR) were removed. Furthermore, the items had to be applicable for each of the three SRL phases in order to facilitate comparison of participants' self-assessment practices at different phases. Hence, two items from the subscale SR (including "I seek out the reasons for mistakes I made after getting back marked work" and "I pay attention to my assessment results in order to identify what I can do better next time") only applicable at Appraisal phase were excluded from the current analysis. The results reported here were based on 16 items (4 items for each subscale). The wording of the items was also slightly modified to fit the scenario of completing an assignment (see Table 1 for an example).

In addition to the items about self-assessment practice, participants were asked to report the final scores they actually got, as the indicator of their academic achievement, at the end of the survey for Appraisal phase. All final scores came from the same rater, i.e., the author, based on the same assessment rubrics.

Table 1. An example of item modification.

Original	Phase	Revised
Any areas I am unsure of after **finishing my work**, I go over again	Preparatory	Any areas I am unsure of after **finishing my work plan**, I go over again
	Performance	Any areas I am unsure of after **finishing my work**, I go over again
	Appraisal	Any areas I am unsure of after **getting back marked work**, I go over again

Data analysis

Before inferential statistics for addressing the main research questions, the Rasch model (Rasch 1960) was first applied to calibrate the person measures on the SaPS. The Rasch model has been widely used in this manner in previous studies (e.g., Deneen et al. 2013; Yan and Cheng 2015; Yan 2018a) due to Rasch analysis being able to transform Likert-type data, which is usually ordinal data, into interval data which is a pre-requite for conventional statistical analysis, such as analysis of variance (ANOVA) (Bond and Fox 2015). As all items in the SaPS share the same 6-point Likert type response scale, the Rating Scale Rasch Model (RSM) (Andrich 1978) was used. Given that the four subscales in the SaPS represent four inter-correlated actions in the self-assessment process, a multidimensional Rasch model was applied using ConQuest 2.0 (Wu et al. 2007). Multidimensional Rasch analysis is able to enhance the measurement precision on each subscale by taking into account the correlations between the subscales. Multiple indices including the Rasch reliability and the item fit statistics, including infit and outfit mean-square (MNSQ), were utilized to examine the psychometric properties of the modified SaPS from a Rasch measurement perspective. Missing data existed in this dataset as most of longitudinal datasets do. Among 63 participants included in the present analysis, 24 participants responded to all three rounds of the survey and 39 participants missed one of the three surveys. As the majority of missing data occurred at Appraisal phase, the missing data pattern was more like a monotone pattern rather than an arbitrary pattern. The Little's test result ($\chi^2 = 60.519$, $p = .57$) showed that the missing data are missing-completely-at-random (MCAR). Hence, full information maximum likelihood estimation, which is generally favored in structural equation modeling (SEM) under the MCAR assumption (Allison 2003), was adopted to deal with missing data.

Based on the Rasch-calibrated person measures, repeated measures multivariate analysis of variance (MANOVA) was undertaken on self-assessment actions measured by the SaPS to examine the differences of self-assessment practices at the three SRL phases. To reveal the relationship between self-assessment practices at different SRL phases and academic achievement, an autoregressive model, Markov simplex modelling (Guttman 1954) with observed variables, was applied to the Rasch-calibrated person measures using the software AMOS version 24. Markov simplex modelling assumes that Performance at time t is a function of its former performance at time t-1 plus random error. Moreover, the performance on one construct influences not only itself at the next time point, but also the other construct at the next time point.

Multiple fit indices reported to examine the model-data fit included χ^2/df, the goodness-of-fit index (GFI), the comparative fit index (CFI), the $\hat{\gamma}$, the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). In general, a χ^2/df ratio of less than 3, a GFI, CFI, and $\hat{\gamma}$ of 0.90 or greater, and a SRMR and RMSEA of 0.08 or less (Hu and Bentler 1999; McDonald and Ho 2002; Fan and Sivo 2007) indicate acceptable model-data fit.

Results

The raw dataset had three waves of data of self-assessment practices collected from the three SRL phases. Accordingly, a multidimensional Rasch model was applied to the three waves of data separately. The psychometric properties of the modified SaPS were first examined before

the main analyses. The results showed that the item fit statistics (MNSQ) of all items, except item #6 in the subscale *seeking external feedback through inquiry (SEFI),* fell into the acceptable range (0.5-1.5) as suggested by Linacre (2006). Item #6 (I ask my family members to give me advice on my work.) demonstrated underfit (MNSQ >1.5) at both performance and Appraisal phases. Hence, item #6 was removed from the scale and the analyses were conducted again. The remaining 15 items showed satisfactory fit to the Rasch model at all three phases (see Table 2 for details).

The MNSQs of all items were within the acceptable range (0.5-1.5) with only two exceptions: the outfit MNSQ of item #7 at Performance phase (1.52) and the infit MNSQ of item #13 at Appraisal phase (1.55) were marginally higher than 1.5. The Rasch reliabilities for all subscales were higher than 0.7 at Preparatory and Performance phase, but were only around 0.57 at Appraisal phase. Perhaps one reason is that the valid responses to the subscales at Appraisal phase were less than those at the other two phases. Person fit statistics were also checked. Seven cases demonstrated marginal or moderate underfit (MNSQ > 1.5), including #46 (1.62), #47 (1.55), and #50 (1.72) at Preparatory phase; #36 (1.61) and #45 (1.59) at Performance phase; #12 (1.71), and #35 (1.70) at Appraisal phase. As none of the cases showed underfit at more than one phase, all cases were kept for the subsequent analyses.

The resultant data for each participant contained Rasch-calibrated measures on the four self-assessment actions at the three SRL phases (i.e., a total of 12 person measures) plus their academic achievement. The repeated measures MANOVA was applied to Rasch-calibrated person measures. The results showed that, in general, there were no statistically significant differences on self-assessment actions among the three phases, $F(8, 55) = 1.659$, $p = .13$. Univariate tests (see Table 3) found no significant differences with regard to SEFM, SIF, and SR across the three phases. However, students' performance on SEFI differed significantly among the three phases. The results of within-subjects contrasts following MANOVA showed that students' SEFI at Appraisal phase were significantly lower ($p < .05$) than that at Preparatory and Performance phases. Figure 3 presents the students' performance on self-assessment actions at the three phases.

The relationship between self-assessment practices at different SRL phases and academic achievement

An autoregressive model was fitted to the Rasch-calibrated person measures to investigate the relationship between self-assessment practices at different SRL phases and academic achievement. The fit between the hypothesised model and the empirical data was good: $\chi^2 = 41.505$, df = 34, $p = .18$; χ^2/df =1.221; GFI = .917; CFI = .982; $\hat{\gamma} = .98$; RMSEA = .060; SRMR = .055.

The standardized path coefficients are presented in Figure 4. For the sake of neat visualization, the correlations between the three feedback seeking actions with the same phase are not displayed; most of the correlation coefficients were moderate and statistically significant. Feedback seeking actions (i.e., SEFM, SEFI and SIF) had impact on self-reflection within the same SRL phase. At Preparatory and Performance phases, SEFM ($r = .32$ and .31, respectively) and SEFI ($r = .31$ and .37, respectively) had stronger impact on SR than SIF ($r = .11$ and .19, respectively). However, at Appraisal phase, only SIF ($r = .44$) showed significant impact on SR. There were significant autoregressive correlations between the same self-assessment action across different SRL phases with only one exception (the autoregressive correlation between SEFI at Performance and Appraisal phases). The autoregressive correlations for the four actions between Preparatory and Performance phases were significant and positive, ranging from .26 to .52. The correlations for SIF and SR between Performance and Appraisal phases were also significant and positive ($r = .27$ and .36, respectively).

However, the correlation for SEFI between Performance and Appraisal phases was non-significant and the correlation for SEFM was significant and negative ($r = -.23$). SR at Performance phase had a statistically significant impact on SEFM, SEFI, and SIF at Appraisal phase, ranging

Table 2. Rasch reliabilities and item statistics of the modified SaPS.

Scale/item	Rasch Reliability	Item measure[a]	S.E.	Infit MNSQ	Outfit MNSQ
Preparatory phase					
Seeking external feedback through monitoring (SEFM)					
Item 1	0.786	0.356	0.137	0.83	0.85
Item 2		−0.026	0.142	0.86	0.79
Item 3		0.339	0.144	1.02	0.98
Item 4		−0.669	0.244	0.92	0.83
Seeking external feedback through inquiry (SEFI)					
Item 5	0.785	−0.319	0.114	1.34	1.35
Item 6[b]					
Item 7		0.172	0.111	1.08	1.12
Item 8		0.147	0.159	1.14	1.25
Seeking internal feedback (SIF)					
Item 9	0.714	−0.23	0.116	1.1	1.17
Item 10		−0.266	0.116	1.27	1.21
Item 11		0.628	0.109	0.73	0.76
Item 12		−0.132	0.197	0.94	0.94
Self-reflection (SR)					
Item 13	0.804	0.323	0.129	1.09	0.98
Item 14		0.171	0.13	1.13	1.07
Item 15		−0.088	0.133	0.82	0.8
Item 16		−0.406	0.226	0.92	1.17
Performance phase					
Seeking external feedback through monitoring (SEFM)					
Item 1	0.762	0.055	0.15	0.81	0.87
Item 2		0.156	0.149	0.72	0.75
Item 3		0.422	0.145	0.81	0.82
Item 4		−0.634	0.257	0.89	0.9
Seeking external feedback through inquiry (SEFI)					
Item 5	0.761	−0.179	0.117	1.23	1.3
Item 6[b]					
Item 7		0.02	0.116	1.37	1.52
Item 8		0.159	0.165	1.38	1.36
Seeking internal feedback (SIF)					
Item 9	0.733	−0.208	0.12	0.96	0.93
Item 10		−0.122	0.12	1.27	1.27
Item 11		0.573	0.116	1.29	1.35
Item 12		−0.242	0.206	0.8	0.79
Self-reflection (SR)					
Item 13	0.800	0.34	0.137	1.3	1.42
Item 14		0.233	0.138	1.37	1.31
Item 15		−0.052	0.14	0.65	0.63
Item 16		−0.521	0.239	0.82	0.68
Appraisal phase					
Seeking external feedback through monitoring (SEFM)					
Item 1	0.582	−0.516	0.186	0.61	0.6
Item 2		−0.184	0.181	0.6	0.55
Item 3		1.009	0.163	1.34	1.38
Item 4		−0.309	0.306	1.1	1.1
Seeking external feedback through inquiry (SEFI)					
Item 5	0.585	−0.322	0.15	1.09	1.21
Item 6[b]					
Item 7		0.418	0.146	1.06	1.05
Item 8		−0.096	0.209	0.6	0.56
Seeking internal feedback (SIF)					
Item 9	0.564	−0.722	0.169	0.87	0.98
Item 10		0.389	0.156	0.87	0.76
Item 11		0.481	0.155	0.78	0.69
Item 12		−0.148	0.277	1.34	1.23
Self-reflection (SR)					
Item 13	0.572	0.114	0.178	1.55	1.13
Item 14		−0.07	0.181	1.03	0.86
Item 15		0.287	0.176	0.99	0.78
Item 16		−0.331	0.309	0.74	0.74

[a]All measures are in logits.
[b]Removed from the scale.

Table 3. The results of univariate tests in the repeated measures MANOVA.

	Mean person measures				Univariate tests		
Sub-scale	Preparatory	Performance	Appraisal	*N*	*df*	*F*	*p*
Seeking External Feedback through Monitoring	2.03	2.26	1.81	63	2	1.570	.212
Seeking External Feedback through Inquiry	0.41	0.56	−0.17	63	2	5.851	.004*
Seeking Internal Feedback	0.93	1.15	1.18	63	2	.774	.463
Self-Reflection	1.90	2.08	1.82	63	2	.934	.396

Note: Mean person measures are in logits..

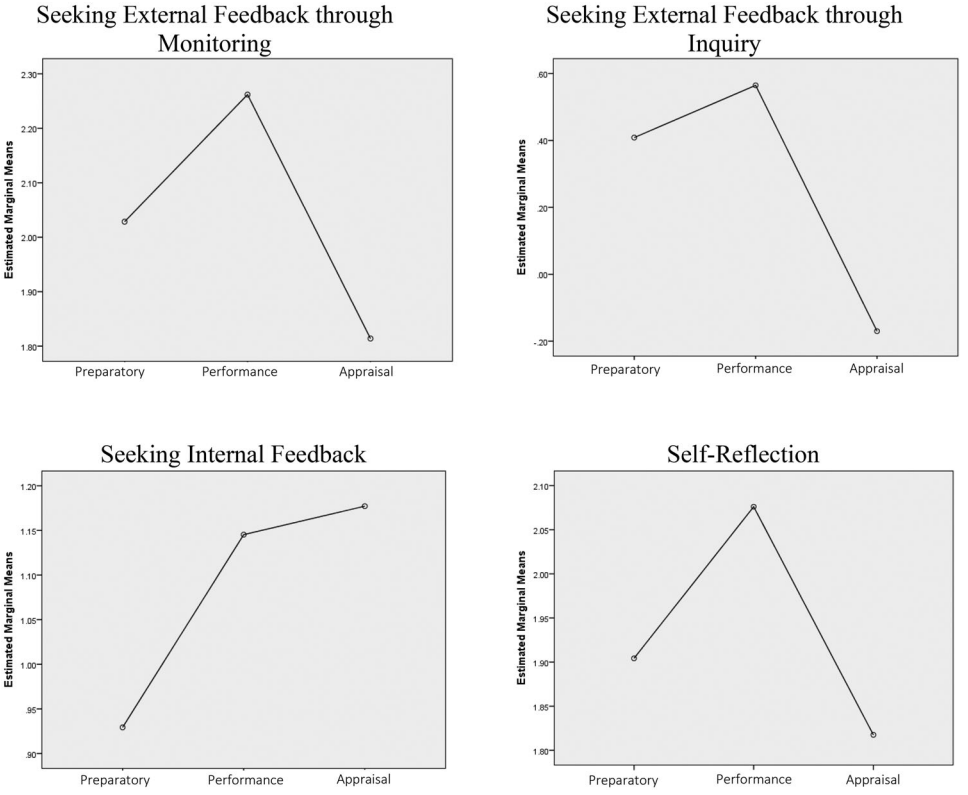

Figure 3. **Students' performance on self-assessment actions at the three SRL phases.**

from .47 to .89. Furthermore, there were significant relationships between self-assessment actions and academic achievement. SEFM at Performance phase was the strongest and positive (.61) predictor of academic achievement, while correlation between SEFI/SR at Performance phase and academic achievement was negative (-.27 and -.32 respectively). The academic achievement had negative impact on all self-assessment actions at Appraisal phase, although only the correlations with SEFI and SR were significant.

Discussion

This study hypothesized that self-assessment is a fundamental skill for SRL and plays an important role during the whole SRL process. It aimed to investigate characteristics of self-assessment practices at different SRL phases as well as its relationships with students' academic achievement.

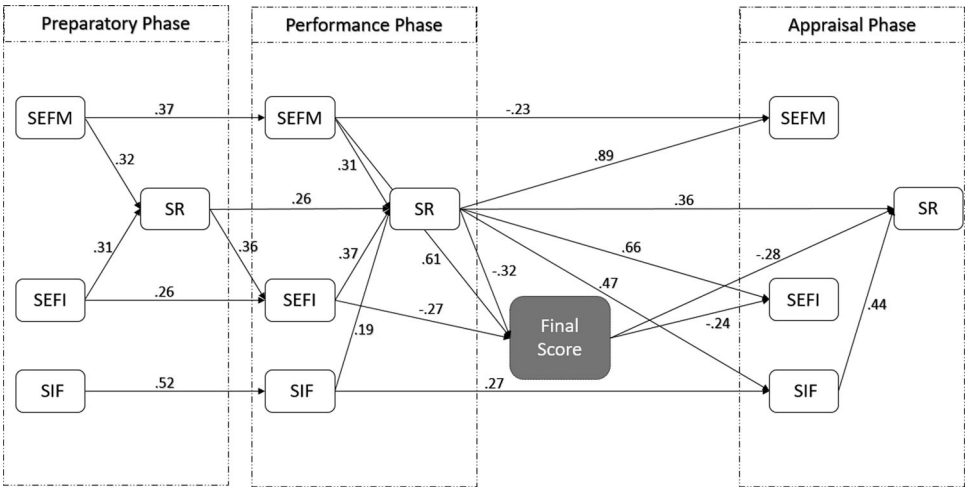

Figure 4. The schematic model of structural relations with standardized path coefficients.
Note: The correlations among SEFM, SEFI, and SIF within the same phase were not displayed; only path with significant coefficients were displayed. SEFM: seeking external feedback through monitoring, SEFI: seeking external feedback through inquiry, SIF: seeking internal feedback, SR: self-reflection

The results generally supported the hypothesis. Students engaged in self-assessment actions at each SRL phase, although the pattern varied. Self-assessment is an ongoing practice across the whole SRL process, rather than a one-off action occurring at a particular time point. In Preparatory phase, students can self-assess the appropriateness of their goals and strategies that will be used. A well calibrated learning goal and a set of well justified strategies would ensure a higher chance of success in subsequent phases. In Performance phase, the function of self-assessment is to continuously monitor the learning progress and ensure it is progressing towards the learning goal as planned. In Appraisal phase, self-assessment enables students to look back into their learning journey and evaluate their strengths and weaknesses demonstrated in the learning process and outcomes. This finding corroborates the argument that self-assessment should be treated as a useful learning strategy, rather than an assessment method (Brown and Harris 2013; Yan and Brown 2017). Pedagogically, it implies that it is better to teach students to prepare for self-assessment as early as possible when engaging in a learning task, rather than hold self-assessment until the task is completed (Topping 2003; Andrade and Valtcheva 2009).

Relationships between self-assessment actions at different SRL phases

The repeated measures MANOVA results showed that self-assessment actions were in general stable across different SRL phases except for SEFI (see Table 3). A closer examination revealed a shared pattern that three self-assessment actions (i.e., SEFM, SEFI, and SR) were less performed at Appraisal phase than the first two phases, although only the difference on SEFI reached statistical significance. This finding contradicts a conventional understanding of the role of self-assessment in SRL. According to most SRL models, self-assessment is usually specified at the last SRL phase and it is supposed that self-assessment at the last phase is more intensive than that at earlier phases. One possible explanation is that students are using self-assessment instrumentally to get a high score in their assignments. They might realize, through previous experiences, that self-assessment at the first two phases can influence the final scores they will get. Consequently, students tend to have higher motivations to self-assess before completing and submitting the assignments. When the assignments have been marked and grades been finalized, they feel no need to do self-assessment anymore. This is similar to what Brown, Peterson and Yao (2016)

found about students' conception about the timing of feedback. Summative feedback provided together with final grades at the end of study has been perceived not to be as powerful as formative feedback provided during study in terms of diagnosing problems and informing improvement.

However, caution should be taken in interpreting this result. First, the nature and purposes of self-assessment at different phases varied. At Preparatory and Performance phases, the targets of self-assessment were the working plan or parts of the assignment; and the main purpose was to improve the quality of the assignment. In contrast, at Appraisal phase, the target of self-assessment was the whole assignment and the purpose was providing feedback for future learning. Second, the survey for Appraisal phase in this study was conducted after provision of the final scores and comments. The time period between submission of assignment and provision of final scores could be an important part of Appraisal phase but was not covered. Although practical experience suggests that students are less likely to conduct self-assessment during that time period, there is no data collected in this study to verify this speculation.

Self-assessment actions (i.e., SIF) were least conducted at Preparatory phase compared to the other two phases, although the difference did not reach statistical significance. This finding, together with the pattern of the other two feedback seeking actions (i.e., SEFM and SEFI), indicates that self-assessment at different SRL phases seems to rely on different sources of feedback. SEFM appeared to be the most utilized feedback seeking action for all the three SRL phases, indicating that evidence regarding their performance collected through monitoring or interaction with the learning task/materials was preferred by students. SEFI was the least performed feedback seeking action for each SRL phase. Previous studies on self-assessment in Hong Kong (e.g. Yan, 2016, 2018b) reported that Hong Kong primary and secondary students performed quite well on self-reflection based on available feedback, but were less inclined to take the initiative to seek feedback regarding their own performance. This study further revealed that Hong Kong students, in universities in this case, were particularly reluctant to seek feedback on their performance via direct inquiry with other people.

As Morrison, Chen and Salgado (2004) suggested, cultural variables must be considered in investigations of feedback seeking behaviour. For example, De Luque and Sommer (2000) found that, in seeking feedback, collectivist cultures tend to prefer indirect inquiry or self-monitoring, but not direct-inquiry behavior, since it might bring too much undesired individual attention to a person. With regard to seeking feedback from internal sources (e.g., emotions, feelings, physical sensations and internal states), students were more inclined to do so at Performance and Appraisal phases, probably because a concrete learning product, such as the half-completed assignment or the marked assignment, might easily elicit internal reactions. However, at Preparatory phase, the working plan (some students might just have an idea in mind rather than a real written plan) appeared to be a less strong stimuli to generate internal reactions regarding the quality of their work.

The results of SEM on an autoregressive model demonstrated the relationships between self-assessment actions within and across SRL phases, as well as the relationships between self-assessment actions (at different SRL phases) and academic achievement. With regard to the impact of feedback seeking actions on self-reflection within the same SRL phase, different patterns emerged. At Preparatory and Performance phases, SEFM and SEFI had stronger impact on SR than SIF. At Appraisal phase, SIF showed much stronger impact than SEFM and SEFI. It implies that self-assessment at Preparatory and Performance phases are "rational" activities as it is based on "objective" external feedback information collected through interaction with learning materials or resource persons. While self-assessment at Appraisal phase is more "emotional" as the feedback information are mainly from internal feelings and physical sensations that are relatively "subjective".

The results regarding the autoregressive correlations between the same self-assessment action across different SRL phases revealed different patterns. The predictive relations of SIF and SR

across SRL phases were relatively stable. The more students conducted SIF or SR at the previous phase, the more they would conduct the same action at the next phase. For SEFI and SEFM, the positive correlation held between Preparatory and Performance phases, but not between Performance and Appraisal phases. Students' SEFI at Performance phase could not predict their SEFI at Appraisal phase at all. Students' SEFM at Performance phase had a negative impact on their SEFM at Appraisal phase. In other words, students who seek more feedback through monitoring at Performance phase would be less likely to do so at Appraisal phase.

This is probably because feedback seeking within self-assessment process serves different purposes at different phases. At Preparatory and Performance phases, students seek feedback for improving the quality, and the possible final grade, of their work. While at Appraisal phase when the final grade has been confirmed, there seems no point for students to seek feedback if they adopt a performance-approach goal. Only those with a mastery goal probably continue self-assessment even though they know the final grade. Yan (2018b) investigated the relationship between self-assessment and goal orientation using a cross-sectional data set. He found that mastery goal orientation was positively related to self-assessment practices, while performance goal orientation was negatively related to self-assessment. With the longitudinal data, this study further revealed the complexity of self-assessment processes. It suggests that students may attach different purposes to self-assessment at different SRL phases and hence differ in terms of self-assessment practices, at least for the feedback seeking actions.

SR at Performance phase had a statistically significant impact on SEFM, SEFI and SIF at Appraisal phase. This means that the more students engaged in self-reflection at performance phase, the more external and internal feedback regarding their performance they would seek at Appraisal phase. This finding corroborates the cyclical nature of self-assessment process. That is, the products of the self-assessment at earlier phases could influence the process and products of self-assessment at the next phase (Yan and Brown 2017).

Relationships between self-assessment actions and academic achievement

SEFM, SEFI, and SR at Performance phase had statistically significant impact on the academic achievement, which had statistically significant impact on SEFI and SR at Appraisal phase. The correlation between SEFM at Performance phase and academic achievement was strongest and positive (.61), while correlation between SEFI/SR at Performance phase and academic achievement was negative (-.27 and -.32 respectively). As SEFM refers to interactions with learning materials and processes, engaging more in SEFM indicates students invest more time and effort on learning, leading to better academic achievement. Furthermore, students who perform better on SEFM at Performance phase are probably those with higher academic ability because they mainly examine their performance by their own, which is an indication of confidence about the way they learn. According, they tend to have better academic achievement. In contrast, too much SEFI, i.e., seeking more feedback from others, is probably an indication of lack of confidence and certainty on their own learning. Brown et al. (2016) found that students who trusted tutors' feedback and seek more peer feedback were likely to have weaker academic performance. They speculated that reliance on feedback from others rather than oneself results in surface learning, but not deeper cognitive engagement, which contributes to lower performance. The result that more self-reflection at Performance phase led to lower academic achievement at first seems at odds. Perhaps this is because self-reflection could be counter-productive under some circumstances. Counter-productive self-reflection, or rumination, has been widely studied in counselling and clinical psychology (Smith and Alloy 2009). Rumination as a form as self-reflection is negative. It does not help problem solving but often results in anxiety related problems. To make productive and learning-oriented self-reflection, student should avoid indulgence in self-doubt. Instead, they

should engage with meaningful learning activities and try to identify feasible way for further improvement.

It was interesting to find that the academic achievement was a negative predictor for all self-assessment actions at Appraisal phase (ranging from -.08 to -.28). In other words, the higher-achieving students were less likely to conduct self-assessment after getting marked assignments back. This indicates that, although in general self-assessment was less conducted at Appraisal phase, students with lower grades had higher motive to self-assess probably because they saw the potential of self-assessment in helping them improve for the next course. On the other hand, students with good grades would think "the task is done" and self-assessment, as an instrument for the task, has no value anymore. Considering that SRL is one of the important goals of higher education, it is important to make students aware that learning has a recursive nature. Self-assessment at Appraisal phase, i.e., when the task is done, could also serve a formative purpose and hence have effects on the success of similar learning tasks in the future.

Limitations and future directions

This study was limited by its reliance on a small sample consisting solely of master students from a teacher education institute. This limitation adds caution regarding the interpretation and generalisability of the findings. Notwithstanding the relatively limited sample, the present study serves well as an attempt to empirically explore the intertwined relationship between self-assessment practices, SRL phases and academic achievement. The findings of this study also provide insights for the direction of future research.

First, the survey for Appraisal phase was conducted after final scores were given and, therefore, did not cover the time period between submission of assignment and provision of final scores. Future investigations could examine whether there is any difference regarding self-assessment practices at Appraisal phase before and after knowing the final scores of the assignments. Second, even though self-assessment incorporates similar mental processes at each SRL phase, it may rely on different resources, serve different purposes and generate different products. Given that the present study revealed different patterns of self-assessment practice at different SRL phases, it would be interesting to validate the speculative explanations proposed, or to further explore the reasons underlying such differences. Only if the micro mechanism of interaction between self-assessment and SRL phases is uncovered, can appropriate instruction and intervention programmes be designed to promote productive self-assessment and SRL.

Third, this study reported the potentially different impact of self-assessment at each SRL phase upon academic achievement using self-reported data. However, without a controlled experimental design, the impact of self-assessment was possibly intertwined with students' characteristics (e.g., academic ability and self-confidence). For example, a student with high ability who does not self-assess might still have good academic achievement, but this does not mean that self-assessment is useless because that student's achievement might be further improved if s/he engaged more in self-assessment. Thus, more research using controlled experiments is needed to add further credibility about the causal relationships between various self-assessment actions at each SRL phase and students' learning outcomes.

Conclusion

This study set out to provide empirical evidence to verify the conceptual argument, made in previous work (e.g., Panadero and Alonso-Tapia 2013; Harris and Brown 2018; Panadero et al. 2018), that self-assessment occurs at each SRL phase. The results of this study indicate that self-assessment is not merely a subprocess that exists in the last SRL phase but acts as the hub of regulation within each phase and continuously monitors the whole process of SRL. The findings

demonstrate that self-assessment at one SRL phase can influence self-assessment at subsequent phases. By describing characteristics of the four self-assessment actions at different SRL phases as well as its relationship with academic achievement, this study introduces a new angle to understand the complexity of self-assessment in relation to learning process. The findings reported here contribute to our understanding of good timing and adaptive pattern for self-assessment relative to the process of SRL and shed new light on the design of instruction or intervention programmes aiming at promoting meaningful self-assessment.

Acknowledgements

This work was supported by a General Research Fund (GRF) (Project No.: EDUHK 18605715) from the Research Grants Council of Hong Kong. I would like to thank Prof. Gavin Brown for his insightful comments on an early draft of this paper, as well as two anonymous reviewers for their constructive feedback.

References

Allison, P. D. 2003. "Missing Data Techniques for Structural Equation Modeling." *Journal of Abnormal Psychology* 112(4):545–557. doi:10.1037/0021-843X.112.4.545.

Andrade, H. L. 2010. "Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning." In *Handbook of Formative Assessment*., edited by H. L. Andrade, and G. J. Cizek, 90–105. New York: Routledge.

Andrade, H., and Y. Du. 2007. "Student Responses to Criteria-Referenced Self-Assessment." *Assessment & Evaluation in Higher Education* 32(2):159–181. doi:10.1080/02602930600801928.

Andrade, H. L., Y. Du, and X. Wang. 2008. "Putting Rubrics to the Test: The Effect of a Model, Criteria Generation, and Rubric-Referenced Self-Assessment on Elementary School Students' Writing." *Educational Measurement: Issues and Practice* 27(2):3–13. doi:10.1111/j.1745-3992.2008.00118.x.

Andrade, H., and A. Valtcheva. 2009. "Promoting Learning and Achievement through Self-Assessment." *Theory into Practice* 48(1):12–19. doi:10.1080/00405840802577544.

Andrich, D. 1978. "A Rating Scale Formulation for Ordered Response Categories." *Psychometrika* 43(4):561–573. doi:10.1007/BF02293814.

Bond, T. G., and C. M. Fox. 2015. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences.* (3rd ed.). New York: Routledge.

Boud, D. 1995. *Enhancing Learning through Self-Assessment*. London: Kogan Page.

Boud, D. 1999. "Avoiding the Traps: Seeking Good Practice in the Use of Self Assessment and Reflection in Professional Courses." *Social Work Education* 18(2):121–132. doi:10.1080/02615479911220131.

Bourke, R. 2018. "Self-Assessment to Incite Learning in Higher Education: developing Ontological Awareness." *Assessment & Evaluation in Higher Education* 43(5): 827-839. doi:10.1080/02602938.2017.1411881.

Brown, G. T. L., H. Andrade, and F. Chen. 2015. "Accuracy in Student Self-Assessment: directions and Cautions for Research." *Assessment in Education: Principles, Policy & Practice* 22(4):444–457. doi:10.1080/0969594X.2014.996523.

Brown, G. T. L., and L. R. Harris. 2013. "Student self-assessment." In *The SAGE Handbook of Research on Classroom Assessment*., edited by J. H. McMillan, 367–393. Thousand Oaks: Sage.

Brown, G. T. L., E. R. Peterson, and E. S. Yao. 2016. "Student Conceptions of Feedback: Impact on Self-Regulation, Self-Efficacy, and Academic Achievement." *British Journal of Educational Psychology* 86(4):606–629. doi:10.1111/bjep.12126.

De Luque, M. F. S., and S. M. Sommer. 2000. "The Impact of Culture on Feedback-Seeking Behavior: An Integrated Model and Propositions." *Academy of Management Review* 25(4):829–849. doi:10.2307/259209.

Deneen, C. C., G. T. L. Brown, T. G. Bond, and R. H. Shroff. 2013. "Telling the Difference: A First Evaluation of an Outcome-Based Learning Innovation in Teacher Education." *Asia-Pacific Journal of Teacher Education* 41(4): 441–456. doi:10.1080/1359866X.2013.787392.

Fan, X., and S. A. Sivo. 2007. "Sensitivity of Fit Indices to Model Misspecification and Model Types." *Multivariate Behavioral Research* 42(3):509–529. doi:10.1080/00273170701382864.

Fastré, G. M. J., M. R. van der Klink, D. Sluijsmans, and J. J. G. van Merriënboer. 2012. "Drawing Students' Attention to Relevant Assessment Criteria: Effects on Self-Assessment Skills and Performance." *Journal of Vocational Education & Training* 64(2):185–198. doi:10.1080/13636820.2011.630537.

Greene, J. A., and R. Azevedo. 2007. "A Theoretical Review of Winne and Hadwin's Model of Self-Regulated Learning: New Perspectives and Directions." *Review of Educational Research* 77(3):334–372. doi:10.3102/003465430303953.

Guttman, L. 1954. "A new approach to factor analysis: the rodex." In *Mathematical Thinking in the Social Sciences*., edited by P. F. Lazarsfeld, 258–349. Free Press, Glencoe, Ill.

Harris, L. R., and G. T. L. Brown. 2018. *Using Self-Assessment to Improve Student Learning*. New York: Routledge.

Hu, L., and P. M. Bentler. 1999. "Cutoff Criteria for fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling* 6(1):1–55. doi:10.1080/10705519909540118.

Jay, J., and A. Owen. 2016. "Providing Opportunities for Student Self-Assessment: The Impact on the Acquisition of Psychomotor Skills in Occupational Therapy Students." *Assessment & Evaluation in Higher Education* 41(8): 1176–1192. doi:10.1080/02602938.2015.1071317.

Kissling, E. M., and M. E. O'Donnell. 2015. "Increasing Language Awareness and Self-Efficacy of FL Students Using Self-Assessment and the ACTFL Proficiency Guidelines." *Language Awareness* 24(4):283–302. doi:10.1080/09658416.2015.1099659.

Linacre, J. M. 2006. *A User's Guide to WINSTEPS/MINISTEP: Rasch-Model Computer Programs*. Chicago, IL: Winsteps.com.

McDonald, R. P., and M. R. Ho. 2002. "Principles and Practice in Reporting Structural Equation Analyses." *Psychological Methods* 7(1):64–82. doi:10.1037//1082-989X.7.1.64.

McMillan, J. H., and J. Hearn. 2008. "Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement." *Educational Horizons* 87(1):40–49.

Morrison, E. W., Y. R. Chen, and S. R. Salgado. 2004. "Cultural Differences in Newcomer Feedback Seeking: A Comparison of the United States and Hong Kong." *Applied Psychology* 53(1):1–22. doi:10.1111/j.1464-0597.2004.00158.x.

Panadero, E. 2017. "A Review of Self-Regulated Learning: Six Models and Four Directions for Research." *Frontiers in Psychology* 8 (422): 1-28. doi:10.3389/fpsyg.2017.00422.

Panadero, E., and J. Alonso-Tapia. 2013. "Self-Assessment: Theoretical and Practical Connotations. When It Happens, How Is It Acquired and What to Do to Develop It in Our Students." *Electronic Journal of Research in Educational Psychology* 11(2):551–576.

Panadero, E., H. Andrade, and S. Brookhart. 2018. "Fusing Self-Regulated Learning and Formative Assessment: A Roadmap of Where we Are, How we Got Here, and Where we Are Going." *The Australian Educational Researcher* 45(1):13–31. doi:10.1007/s13384-018-0258-y.

Panadero, E., G. T. L. Brown, and J.-W. Strijbos. 2016. "The Future of Student Self-Assessment: A Review of Known Unknowns and Potential Directions." *Educational Psychology Review* 28(4):803–830. doi:10.1007/s10648-015-9350-2.

Panadero, E., A. Jonsson, and J. Botella. 2017. "Effects of Self-Assessment on Self-Regulated Learning and Self-Efficacy: Four Meta-Analyses." *Educational Research Review* 22:74–98. doi:10.1016/j.edurev.2017.08.004.

Puustinen, M., and L. Pulkkinen. 2001. "Models of Self-Regulated Learning: A Review." *Scandinavian Journal of Educational Research* 45(3):269–286. doi:10.1080/00313830120074206.

Rasch, G. 1960. *Probabilistic Models for Some Intelligence and Achievement Test.* Copenhagan: Danish Institute for Educational Research. Expanded ed. (1980). Chicago: The University of Chicago Press.

Ross, J. A. 2006. "The Reliability, Validity, and Utility of Self-Assessment." *Practical Assessment* 11(10):1–13.

Sargeant, J., H. Armson, B. Chesluk, T. Dornan, K. Eva, E. Holmboe, J. Lockyer, E. Loney, K. Mann, and C. van der Vleuten. 2010. "The Processes and Dimensions of Informed Self-Assessment: A Conceptual Model." *Academic Medicine* 85(7):1212–1220. doi:10.1097/ACM.0b013e3181d85a4e.

Smith, J. M., and L. B. Alloy. 2009. "A Roadmap to Rumination: A Review of the Definition, Assessment, and Conceptualization of This Multifaceted Construct." *Clinical Psychology Review* 29(2):116–128. doi:10.1016/j.cpr.2008.10.003.

Topping, K. J. 2003. "Self and peer assessment in school and university: Reliability, validity and utility." In *Optimizing New Modes of Assessment: In Search of Qualities and Standards*., edited by M. S. R. Segers, F. J. R. C. Dochy, and E. C. Cascallar, 55–87. Dordrecht, Netherlands: Kluwer Academic Publishers.

Wu, M. L., R. J. Adams, M. R. Wilson, and A. A. Haldane. 2007. *ACER ConQuest [Computer Program and Manual].* Victoria, Australia: ACER Press.

Yan, Z. 2016. "The Self-Assessment Practices of Hong Kong Secondary Students: Findings with a New Instrument." *Journal of Applied Measurement* 17(3):1–19.

Yan, Z. 2018a. "How Teachers' Beliefs and Demographic Variables Impact on Self-Regulated Learning Instruction." *Educational Studies* 44(5):564–577. doi:10.1080/03055698.2017.1382331.

Yan, Z. 2018b. "Student Self-Assessment Practices: The Role of Gender, Year Level, and Goal Orientation." *Assessment in Education: Principles, Policy & Practice* 25(2):183–199. doi:10.1080/0969594X.2016.1218324.

Yan, Z. 2018c. "The Self-Assessment Practice Scale (SaPS) for Students: Development and Psychometric Studies." *The Asia-Pacific Education Researcher* 27(2):123–135. doi:10.1007/s40299-018-0371-8.

Yan, Z., and G. T. L. Brown. 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. doi:10.1080/02602938.2016.1260091.

Yan, Z., and E. C. K. Cheng. 2015. "Primary Teachers' Attitudes, Intentions and Practices regarding Formative Assessment." *Teaching and Teacher Education* 45:128–136.

Zimmerman, B. J. 2000. "Attaining self-regulation: A social-cognitive perspective." In *Handbook of Self-Regulation*., edited by M. Boekaerts, P. Pintrich, and M. Zeidner, 13–39. San Diego, CA: Academic.