Title	Implementing self-assessment in Singapore primary schools: Effects on students' perceptions of self-assessment
Author(s)	Hwei Ming WONG
Source	*Pedagogies: An International Journal, 12*(4), 391-409
Published by	Taylor & Francis (Routledge)

To cite this article: Hwei Ming Wong (2017): Implementing self-assessment in Singapore primary schools: effects on students' perceptions of self-assessment, Pedagogies: An International Journal
To link to this article: http://dx.doi.org/10.1080/1554480X.2017.1362348

Implementing self-assessment in Singapore primary schools: Effects on students' perceptions of self-assessment

Hwei Ming WONG

Centre for Research in Pedagogy and Practice, National Institute of Education, Nanyang Technological University, Singapore

1 Nanyang Walk, Singapore 637616

Tel: (65) 6790 3772, Fax: (65) 6316 4787

Email: hweiming.wong@nie.edu.sg

Notes on Contributor

Hwei Ming WONG is a Research Scientist at the Centre for Research in Pedagogy and Practice, National Institute of Education where she has worked on various research projects related to assessment for the last 13 years. Before that, she was an Educational Psychologist with the Ministry of Education, Singapore and she has also previously taught in a primary school. Her research interests are self-assessment, assessment for learning, formative assessment, assessment practices and literacies, teacher pedagogical practices, and educational psychology.

Implementing self-assessment in Singapore primary schools: Effects on students' perceptions of self-assessment

Abstract

Student academic self-assessment engages the students in deliberate reflection about what they are learning and how they are learning it. This intervention study investigated the effects of self-assessment training on students' perceptions towards self-assessment in two Singaporean primary schools. The study, which used a pretest-posttest design, involved 146 Primary 4 (aged 10) students with 75 students being taught the use of self-assessment (intervention group) and 71 students who did not receive any training (comparison group). Data was collected using a Self-Assessment Questionnaire (SAQ), wherein the students reported on the domains of Knowledge Application, Independent Learning, Communication and Motivation. The results revealed an improvement in the intervention students' perceptions towards self-assessment after the use of self-assessment and there were differences between the intervention students and the comparison students. The findings and the implications for students will be discussed in the article.

Keywords: self-assessment; primary school; students; Singapore; assessment for learning; formative assessment

Introduction

Assessment in education occurs for formative and summative purposes. It focuses on educational process outcomes, for example, review, certification, accountability, and the educational development and learning support of students (Black, 1997). For education outcomes, assessment is used for grading and ranking, evaluating teacher effectiveness and school standards (Stobart, 2008). For students, assessment is used to motivate and improve future performance, provide feedback, diagnose strengths and weaknesses, differentiate learning opportunities, direct student's choices, and for course or career selection (Black, 1997; Wiliam, 2008). Hence, assessment generates information for students, teachers, and schools which they can use to improve themselves.

This article investigates primary school students' perceptions towards self-assessment which engages the students in reflection about what they are learning and how they are learning it. This article is structured as follows: Section 1 presents the Singaporean assessment context, the definition of self-assessment, a brief literature review of formative and self-assessment, the research objectives and the study's significance. Section 2 describes the research design, participants, the Self-Assessment Questionnaire (SAQ) and self-assessment strategies used in the study, and the data collection procedures. Section 3 presents the MANOVA results and discussion. The final section reports the implications of the findings, limitations and further research and conclusion.

The assessment context in Singapore

Assessment in Singapore is traditionally centred on high-stakes national examinations such as the Primary School Leaving Examination (PSLE)[1], Singapore-Cambridge General Certificate of Education Ordinary Level and Advanced Level examinations, and they have a profound impact on teaching and learning as they formed the tried-and-

tested route that leads to successful educational outcomes (Hogan et al., 2013; Ministry of Education (MOE), 1998, 2004). Such high-stakes examinations have created a strong 'examination culture' in Singapore (Cheah, 1998) whereby teachers emphasise teaching to the test, with a narrow focus on basic knowledge and lower-order thinking skills (Koh, Lee, Tan, Wong, & Guo, 2005; Koh, Lee, Gong & Wong, 2006; Koh & Luke, 2009). For example, in primary school education, schools conduct their own mid-year and year-end examinations for Primary 2 to 6 (ages 8 to 12). At the end of Primary 6, all students are assessed on their academic abilities through the high-stakes PSLE which is a placement examination to assign the students into appropriate secondary school courses (MOE, 2012; Tan, Chow & Goh, 2008).

MOE's curriculum review (1998) recommended using multiple and alternative modes of assessment to improve Singapore's assessment practices, with similar recommendations again in 2004 and 2008. In 2009, the Primary Education Review and Implementation (PERI) committee made recommendations to improve primary education which included using holistic assessment to support learning by de-emphasising examinations in Primary 1 and 2, using more bite-sized assessments, assessment rubrics and providing qualitative feedback to students. These recommendations reflected similar changing trends in assessment worldwide with a shift from assessment of learning or summative assessment, which is used to evaluate students for purposes of grading and reporting, to assessment for learning or formative assessment, which is used to improve ongoing learning (Black & Wiliam, 1998; Shepard, 2000; Stiggins, 2002; Wiliam, 2006).

Hence, MOE is now increasingly introducing and incorporating multiple and alternative modes of assessments into Singaporean students' academic assignments and classroom assessments. The aim is to push schools and teachers to move beyond the

prevalent examination culture's focus on narrow, academic outcomes (Koh & Luke, 2009; Nichols, Glass, & Berliner, 2006). Understanding this push requires examining the nature of formative assessment and self-assessment specifically.

The nature of formative assessment and self-assessment

As mentioned, there are two broad purposes of assessment: summative and formative. Formative assessment or assessment for learning is defined as "encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (Black & Wiliam, 1998, pp. 7–8). It is a process of students and teachers understanding and using the evidence to take actions to improve the students' learning (Assessment Reform Group, 2002). It can be criterion-referenced where students relate their learning in terms of specific content or skills and has the potential to shift students towards intrinsic motivation and autonomous learning (Harlen & James, 1997; Hughes, 2011). For students, the success of formative assessment is dependent on their perception of the gap between their actual level and a reference level of performance, and their motivation to attend to it.

One important area related to formative assessment, which assess students' learning that are not captured by traditional paper-and-pencil testing (Fan, 2002; Koh et al., 2006), is student self-assessment. Self-assessment "involves students in thinking about the quality of their own work, rather than relying on their teacher as the sole source of evaluative judgements" (Andrade & Valtcheva, 2009, p. 13). Examples of self-assessment are learning logs, journals, checklists and charts, learning contracts, and portfolios. Self-assessment is also called "assessment as learning" (see Dann 2014; Earl, 2006, 2013) which is "a subset of assessment for learning that emphasises using assessment as a process of developing and supporting metacognition for students" (Earl,

2006, p. 7) where students are the important link between assessment and learning. In general, self-assessment is largely used for formative purpose to enhance learning and to promote the monitoring and managing of one's own learning (Andrade & Valtcheva, 2009). The feedback or information from self-assessment about students' learning can be used to "deepen their understandings and improve their performances" (Andrade, 2010, p. 91).

However, self-assessment has also been mobilised for summative purpose with undesirable consequences. For instance, studies have noted that in summative oriented self-assessment, students' self-assessment accuracy is affected resulting in over and underestimations of marks because of their desire to increase marks and the stress of assessing oneself (Leach, 2012; Tejeiro et al., 2012). Consequently, Andrade and Du (2007) and Andrade and Valtcheva (2009) argued that self-assessment must be used for formative purposes such as to identify strengths and weaknesses in one's work for revision, improve and promote one's own learning. In emphasising this point, Andrade and Du (2007) believed that "counting self-assessments toward final grades could turn students' attention away from the quality of their work and how to improve it, and toward getting a high grade, thereby compromising their honesty and their focus on learning" (p. 172). Generally, research indicated that self-assessment when used for the summative purpose of grading and reporting diverts students' attention from learning, and also reduces the accuracy of self-assessment. Therefore, self-assessment should be used primarily for formative purpose and this study emphasises using self-assessment for enhancing ongoing learning.

A number of key assumptions are embedded in self-assessment: learners are capable of evaluating themselves, they have explicit goals, and a process of self-evaluation will make them more motivated and aware of their own learning process, and

they become more engaged with their studies (Trunnel, 1992). Research documenting the potential impact of self-assessment on learning show the following: students are able to develop the ability to think for themselves, develop confidence in their ability to learn and evaluate what they learn, and continue to learn when their schooling days are over (Andrade & Du, 2007; Black & Harrison, 2001a, 2001b; Brown & Glasner, 1999; McDonald & Boud, 2003).

Research in self-assessment

In the last decade, self-assessment research have identified knowledge application (Andrade & Du, 2007; McDonald & Boud, 2003; Wilson & Johnson, 2000), independent learning (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Mok, Lung, Cheng, Cheung, & Ng, 2006; Wilson & Johnson, 2000), communication (Black & Harrison, 2001a, 2001b; Munns & Woodward, 2006) and motivation (McDonald & Boud, 2003; McMillan & Hearn, 2008) as important self-assessment outcomes.

In this study, knowledge application is defined as the ability to make connections between different knowledge and to analyse, evaluate, generate and develop ideas using different thinking skills. Independent learning is the ability to learn on their own, reflect and take appropriate actions to improve their learning. Wilson and Johnson (2000, p. 15) found that when students used self-assessment, they were able to "reflect upon their own learning processes" and demonstrated "increased self-awareness, evaluation of their own progress and understanding that reflecting on their own thinking can lead to self-improvement". Communication is the skills to communicate effectively, and to present ideas clearly and coherently to specific audience in both written and oral forms and motivation is the desire to achieve a goal and is often combined with the energy to work towards that goal. Black and Harrison (2001a) found that through self-assessment, students were "led to more thoughtful

classroom dialogues" (p. 60). McMillan and Hearn (2008) captured motivation (along with knowledge application and independent learning) succinctly as follows, "When students set goals that aid their improved understanding, and then identify criteria, self-evaluate their progress towards learning, reflect on their learning, and generate strategies for more learning, they will show improved performance with meaningful motivation" (p. 48).

These four domains are also important because MOE has indicated knowledge application, independent learning and communication to be important domains that are locally tested and used in the school subject called Project Work, which aims to generate learning experiences that enhance students' knowledge and enable them to acquire skills like communication and independent learning. The fourth domain on motivation is adapted from Soh (1993) which was locally tested and verified. Given the largely positive research findings, these four domains are chosen to provide the focus of the SAQ instrument which is described below.

In terms of school level, self-assessment research focussed mainly on secondary education (e.g., Black et al., 2004; Lasonen, 1995; McDonald & Boud, 2003; Yu, 2013) and higher education (e.g., Andrade & Du, 2007; Cassidy, 2007; Mok et al., 2006; Orsmond, Merry, & Reiling, 2000) but much less in primary education, particularly in Asia and in Singapore. Primary school students have, generally, been perceived as cognitively less matured, hence preventing them from performing self-assessment accurately (Fontana & Fernandes, 1994; Ross, 2006). Nonetheless, the few studies conducted with primary school students (e.g., Andrade, Du, & Wang, 2008; Brown, 2008; Munns & Woodward, 2006; Wong, 2016) illustrated that students, as young as 7 and 8 years could do self-assessment. These students, however, also needed training in using self-assessment and be well versed in the assessment criteria (e.g., Brown, 2008;

Ross, Hogaboam-Gray, & Rolheiser, 2002). This study therefore aims to contribute to research on younger primary school students' ability to use self-assessment.

Examining reliability, validity and utility of self-assessment, Ross (2006) found that reliability for self-assessment was consistent across tasks, items, and over short time periods when students were trained in self-assessment. Self-assessment was however less consistent between two time periods when involving young children, and there were variations among subjects. Validity had mixed results: "self-assessment provides information about student achievement that corresponds only in part to the information generated by teacher assessment" (Ross, 2006, p. 4). This could be due to students' overestimations, their inability to use assessment criteria, and inclusion of information known only to students. Self-assessment therefore contribute to student achievement when teachers provide direct instruction in self-assessment skills, leading to increased self-efficacy and greater intrinsic motivation. The strengths of self-assessment could be augmented through training students on how to self-assess and the weaknesses of self-assessment, such as teachers' concerns about sharing control of assessment with students, students' lack of commitment to self-assessment, lack of time, could be reduced through teacher action. The literature revealed that with proper implementation, self-assessment produces valid and reliable information about student achievement and provides formative data on students' cognition about their achievement that is otherwise not available to teachers (Ross, 2006; Ross et al., 2002; Wong, 2016).

Research objectives

While this study focuses only on primary students of a certain age group, it nevertheless hopes to critically question the perception of children's inability to self-assess through the systematic implementation and use of self-assessment. The study also examined how

self-assessment worked for the students, and the improvements and benefits as perceived by them. With the main purpose to examine the perceptions of primary school students towards self-assessment, the following research questions (RQ) were specifically examined:

RQ1: How do Primary 4 students (aged 10) change their perception in terms of their knowledge and use of self-assessment in Mathematics (using checklists, learning logs and rubrics) after a structured classroom intervention?

The hypothesis is that there would be a change in the intervention students after an intervention in the use of self-assessment, with students moving towards a more positive view of self-assessment and an increase in willingness to use self-assessment in the future.

RQ2: How do Primary 4 students (aged 10) perceive using self-assessment in relation to knowledge application, independent learning, communication and motivation in Mathematics after the structured classroom intervention?

In the literature, students generally appreciated self-assessment and showed increases in motivation, understanding, self-awareness and better knowledge application. Similar gains for the students in Singapore are expected when they use self-assessment.

The Primary 4 (P4) students were chosen as subjects because they were deemed to have adequate literacy and language skills. Primary 1, 2, and 3 students (aged 7, 8, and 9) might not have the literacy and language skills to fully comprehend the demands required of them for the study. Primary 5 and 6 students were not involved because these students had a tight curriculum schedule, especially the Primary 6 students who were preparing for the high stakes PSLE.

Significance of the study

This study aims to contribute to the empirical evidence of the potential usefulness of self-assessment for young students in the classroom context. The benefits of using self-

assessment are examined through students' self-report because students' thoughts and experiences provide a critical source of data and feedback about self-assessment and about their learning. Student perceptions of self-assessment can potentially have an impact on their use of self-assessment. This study also aims to develop and document a systematic process of training students to undertake self-assessment, hence contributing to the international literature on how a systematic introduction and training of young students in the use of self-assessment can result in changes in student beliefs about the usefulness of self-assessment.

Methodology

Participants

A purposive school selection method was used with the intention to work with average-performing 'neighbourhood' schools which represent the majority of schools. Such neighbourhood schools are located in "mostly underprivileged, working-class neighbourhoods and receive students from these areas" (Kamer-Dahl & Kwek, 2011, p. 162). Two government mixed-gender neighbourhood schools (Schools A and B) were chosen in terms of comparability of student characteristics (i.e., gender and academic achievements) and class student-teacher ratio (approximately 40 students: 1 teacher, equal number of boys and girls) to enable comparison between schools for this study.

In both Schools A and B, one P4 class was involved in an intervention that trained students to use self-assessment (classes A1 and class B1) while another P4 class served as a comparison (classes A2 and B2). The pair of P4 classes chosen in each school (A1 and A2, B1 and B2) were comparable in terms of the students' gender, academic achievements and their knowledge and use of self-assessment. All students were previously exposed to limited forms of self-assessment (checklists) while in Primary 3. .

Due to ethical considerations, random assignment of students into various experimental groups was impossible and thus, a quasi-experimental approach was used. The intervention and comparison classrooms, inclusive of the teachers, were kept intact. Both the intervention and comparison teachers were instructed not to discuss the study with one another. The comparison teachers were also instructed to continue with their usual assessment practices in their classes. All four teachers adhered strictly to the agreement of non-disclosure to their colleagues.

Information on the students involved in the study is presented in Table 1.

< insert Table 1 here>

Subject

The study selected Mathematics as it is a technical and highly procedural subject which primary students will find easier to gauge their progress from 'not understanding' to 'understanding'. Both schools used the same MOE-approved Mathematics textbooks and workbooks, reducing potential issues arising from teachers using different textbooks, content or topic. However, it should be noted that while Singapore mathematics teachers may use such textbooks for content, they can flexibly weave in additional resources depending on student abilities and progress. To clarify, this study does not attempt to improve mathematics classroom instruction but focuses on increasing students' awareness of the repertoire of learning skills used in self-assessment.

Instruments and procedures

The SAQ was adapted from (Wong, 2012) and developed further to assess students' perceptions of self-assessment. The SAQ items asked students' thoughts about their perceived achievement through self-assessment - knowledge application, independent learning, communication and motivation.

There were 10 questions under each domain, for example, "I can use the method of problem solving." (knowledge application), "I am aware of my own thinking." (independent learning), "I can clearly communicate ideas to others and provide support and details." (communication), and "I put in more effort in my work." (motivation). The students answered the items on a 5-point Likert scale ranging from "Strongly agree" to "Strongly disagree".

The SAQ was piloted to a class of 41 P4 students from a third school. Using Cronbach's Alpha, the results showed that the SAQ's reliability values ranged from .80 to .87, indicating good reliability (see Table 2). The SAQ was also analysed using confirmatory factor analysis (CFA) based on the post-intervention SAQ data as a one-factor model and as a four-factors model. Results indicated the four-factors model had better 'fit' (see Table 3). Thus, the four-factors model was used for the study. The factor loading values for the SAQ items were mostly larger than .50 and reached significance ($p < .01$) (Table 2). However, the Tucker Lewis Index (TLI) and Comparative Fit Index (CFI) were below the recommended value of > 0.90 (Hu & Bentler, 1999). This will be discussed further under the study's limitations.

<insert Table 2 here>

<insert Table 3 here>

The SAQ was administered to the intervention and comparison students separately, before and after the intervention students' use of the self-assessment strategies.

A set of student self-assessment strategy worksheets was created for the intervention: checklists, learning logs, and rubrics. These strategies used the following criteria in which the students would assess themselves on: (1) Deep understanding – their understanding of the problem or task, (2) Strategies and reasoning – their ability to solve the problem or task through strategies and reasoning, (3) Clarity - the clarity of their solution, (4) Written communication - their communication of their solutions, and (5) Effort - the extent of effort they gave. The criteria embedded in the self-assessment strategies are derived from the SAQ domains. To elaborate, deep understanding is related to the independent learning domain as students become aware of what they do or do not understand and take appropriate action(s), e.g. seek help from teacher. Strategies and reasoning is related to the knowledge application domain as students think about and decide on the appropriate strategy to apply to problems. Clarity and written communication pertain to the communication domain as students present ideas clearly and coherently, and effort is related to the motivation domain.

The checklist worksheet required the students to read two sentences under each criterion and tick – 'Yes', 'No', 'Unsure' - to reflect their thinking of their own work. Thus, there were a total of 10 sentences for the five criteria, for example, "I understand what this Math Topic is about." and "I understand what the questions want." (understanding).

In the learning log worksheet, the students were required to write honestly on their thought about their work under the five criteria, for example, "What do I think about my understanding? Good or bad? Why?" (understanding).

The rubrics worksheet required the students to indicate the levels they were at in each of the five criteria. Under each criterion, there were descriptors for the four levels (from 1 to 4). Students read these descriptors and select the level most suitable for each of the five criteria, for example, for understanding: "I try but I don't understand what the questions want." (level 1), "I have a little understanding of what the questions want." (level 2), "I understand most of the time what the questions want." (level 3) and "I have very good understanding of what the questions want." (level 4).

Intervention

The intervention was based on the following principles adapted from Fuchs (2011) and incorporated the conditions advocated by Andrade (2010), and Andrade and Valtcheva (2009): explicit instructions and training in self-assessment skills and criteria, and opportunities for practice. The design of the intervention – training students to self-assess their work – was modified from the intervention conducted in Ross et al. (2002). The process for this intervention is: (1) creating self-assessment criteria, (2) teaching the students how to apply the criteria, and (3) giving students feedback about self-assessment. The assessment criteria were first created in collaboration with the intervention teachers, who provided input about its appropriateness for P4 students, because students were not familiar with the setting of criteria. Once the criteria were finalised with the teachers, the students were taught by the researcher how to apply the criteria. The students then made judgements on exemplars and gave evidence for their judgements. The researcher gave feedback to students on how well they had assessed the exemplars. Throughout the

intervention, it was stressed to the students that the self-assessment is used purely for formative purposes, that is, for improving their learning and not for allocating grades.

The intervention took place during two of the four terms in the school year, with a total of 20 weeks used as the intervention period. In the first week of Term 2 (late March – late May), the researcher introduced learning styles, goal-setting, and self-assessment (a condition suggested by Andrade (2010), and Andrade and Valtcheva (2009) to raise students' awareness of self-assessment), as the first of four sessions (approximately an hour), to the intervention students in the two schools. Throughout Term 2, the researcher conducted the subsequent three 20-minute sessions to teach three self-assessment strategies – checklist, learning log and rubrics, one in each session. At each 20-minute session, students were taught how to use the self-assessment strategy to assess their work. The self-assessment criteria were explicitly explained to the students and student work examples corresponding to different levels of the criteria were shown to students to illustrate how to self-assess. The intervention students had a fortnight to practice the self-assessment strategy three times before the next strategy was introduced. The researcher monitored the students each time they used the self-assessment strategy, and remained in the class for students to ask questions when using the self-assessment strategy.

In Term 3 (late June – early September), the intervention students continued to practice and use the self-assessment strategies for another four fortnights. In the first fortnight of Term 3, the researcher refreshed the students on the self-assessment strategies that they had learnt in the last term, followed by students using each strategy once. After the first fortnight, the intervention students picked two self-assessment strategies to use for the next three fortnights: one strategy for two fortnights and another for a fortnight. Again, the intervention students used the self-assessment strategy three times in each

fortnight period. The researcher continued to visit the intervention classes to monitor the use of the self-assessment strategies.

Throughout Terms 2 and 3, while the intervention students received training on the use of the self-assessment strategies and were using the self-assessment strategies on a regular basis in their classes, the comparison students were not given any self-assessment training and they continued with the conventional way of being assessed, that is, by their teachers, "the sole source of evaluative judgements" (Andrade & Valtcheva, 2009, p. 13).

Results and discussion

Self-Assessment Questionnaire (SAQ)

The data was analysed using MANOVA to determine if the intervention has an effect on the dependent variables of knowledge application, independent learning, communication and motivation between the two groups. This study proceeded with MANOVA because the sample in each cell was greater than the number of dependent variables (Hair, Black, Babin, & Anderson, 2010; Van Voorhis & Morgan, 2007).

There was a statistically significant difference between intervention and comparison group on knowledge application, independent learning, communication and motivation over time, Pillai's Trace = .286, F (4, 141) = 14.141, $p < .001$, partial eta squared (η_p^2) = .29 with a large effect size. Generally, the η_p^2 values of >.05, >.1, and >.2 represents 'small', 'medium', and 'large' effect sizes respectively (Cohen, 1969). Table 4 summaries the MANOVA results. The results showed that there were significant differences between the intervention group's pre-intervention and post-intervention scores and between the intervention and comparison group for knowledge application, independent learning, communication and motivation. The knowledge application results echo that of Black and Harrison (2001a, 2001b), McDonald and

Boud (2003), and Mok et al., (2006). The independent learning findings reflect studies by Black et al., (2004), McDonald and Boud (2003), and Mok et al., (2006). The communication results are similar to Black and Harrison (2001a), and Munns and Woodward (2006) and the motivation findings are likewise observed in Andrade and Du (2007), Brown (2008), McMillan and Hearn (2008), and Ross (2006). Table 5 summarises the mean scores and standard deviations of the groups for knowledge application, independent learning, communication and motivation.

< insert Table 4 here>

< insert Table 5 here>

The intervention students changed their perceptions of self-assessment after the intervention as compared to the comparison students. Within the intervention students, their perceptions of self-assessment also became more positive after having learnt to use self-assessment.

While small to medium effects are usually found in educational research, the medium to large effects found in this study could be due to the frequency and intensity of the intervention. The researcher was present whenever the self-assessment was used in the intervention classes on an average of one-two times a week for at least 15-20 minutes. The students received help from the researcher who clarified their doubts when using self-assessment and reinforced their understanding and use of self-assessment. The intervention teachers also encouraged the students to use self-assessment.

Andrade (2010), Andrade and Valtcheva (2009), and Cassidy (2007) advocated the importance of self-assessment criteria to be clear and be understood by students, and

guided opportunities for students to self-assess must be explicitly presented. Thus, the significant differences between the pre-intervention and post-intervention scores for the intervention students for the four domains suggested that when self-assessment is explicitly introduced in a structured and organised way to students, and when the students understood the criteria to base their work upon (see Andrade, 2010; Andrade & Valtcheva, 2009), they saw the benefits of using self-assessment and perceived self-assessment to improve their knowledge application, independence in learning, ability to communicate and motivation. This is heartening as it indicates that young students have the capability to use self-assessment as well as older students. The significant differences between the intervention and comparison students indicated the importance of providing opportunities to students to use self-assessment and the need for students to engage in self-assessment on a regular basis in order for the benefits of self-assessment to be perceived by the students.

Students' responses to self-assessment questions

At the post-intervention SAQ administration, there were several questions at the end of SAQ for the students to answer. Apart from the four common questions, the intervention students answered eight additional questions and the comparison students answered two additional questions. Tables 6 and 7 showed the students' responses to self-assessment questions.

The intervention students' responses to the self-assessment questions after they have been through the use of self-assessment are shown in Table 6. Majority of the intervention students preferred using the rubrics (69.3%) for their self-assessment, reflecting similar preferences of students in studies by Andrade and Du (2007) and Orsmond et al., (2000). The intervention students felt that they have learnt most about independent learning (41.3%) through self-assessment, 44% indicated that they would

use self-assessment weekly, and 53.3% indicated that they would sometimes use self-assessment without teacher prompts. Majority of the intervention students (88%) felt that they needed more practice in assessing their own work and they were willing to use self-assessment for other subjects (92%).

<insert Table 6 here>

The intervention students were also asked to rank the self-assessment domains in order of importance as they perceived it. Most students perceived and ranked independent learning as the most important, followed by knowledge application (2nd), communication (3rd), and motivation as least important.

Table 7 showed the responses from the intervention and comparison students Majority of the intervention students (98.7%) and the comparison students (83.1%) believed that they should participate in assessing their own work. A large number of intervention students (74.7%) and comparison students (69%) also felt that self-assessment would give information to their teachers and parents and that their teachers and parents should be informed about their self-assessment (84% and 73.2% for intervention and comparison students respectively). This further confirms the findings by Ross (2006) that self-assessment provides valid and formative information on students' cognition about their achievement that is otherwise not accessible to teachers.

<insert Table 7 here>

The comparison students were asked, "Have you used self-assessment before?" and "Would you learn how to use self-assessment?". Majority of the comparison

students indicated that they had not used self-assessment before (76.1%) and 81.7% were willing to learn to use self-assessment.

Overall, for RQ1, after using self-assessment, the intervention students found it beneficial, were willing to use it frequently and for other subjects, and they were using it voluntarily. They found rubrics most useful and felt that self-assessment would provide feedback to teachers and parents about their learning. For RQ2, after using self-assessment, the intervention students rated their learning on knowledge application, independent learning, communication, and motivation much more favourably and positively. They indicated they learnt most about independent learning and ranked it as most important. This suggested that intervention students were aware of the need for, and of the importance of, independent learning and self-reliance as opposed to relying only on teachers (Andrade & Valtcheva, 2009).

Implications

In this study, the intervention teachers provided the time and space for their students to use self-assessment on a regular basis. This reinforces the idea that self-assessment could be an important part of the classroom learning culture where young students are encouraged to use self-assessment (Nichols & Berliner, 2008).

The intervention design required that students understood the purpose of self-assessment and were instructed on the self-assessment criteria with exemplars provided. The documentation of the training procedures for student self-assessment in this study can facilitate replication studies and allow practitioners to attempt self-assessment in their own classrooms. It is also important to communicate to students that self-assessment is used purely for formative purposes to improve their learning rather than for allocating grades (Andrade, 2010; Andrade & Du, 2007). Classroom time is needed to teach students how to use self-assessment, such as explaining the self-assessment

criteria and standards and ensuring these are understood, and the modelling the use of self-assessment (McDonald & Boud, 2003; McMillan & Hearn, 2008).

Given that the intervention students would like to use self-assessment independently, the roles of students need to shift from being passive learner to active learner, and from being passively assessed by teachers to becoming an active assessor of their own learning (Andrade et al., 2008). Students are therefore afforded the opportunities to take on more responsibilities for their own learning, and make decisions about their own learning. It is also imperative for students to believe that through their own efforts, they are able to improve their work and take control of their own learning. The findings of the study suggested that it might be fruitful for researchers in Singapore and other Asian countries to consider investigating self-assessment at the primary school level to see if comparable results can be obtained.

Limitations and further research

As the study used only two schools and four classes, the seemingly positive effects of self-assessment could be limited to the participants of the two schools. Therefore, the results could not be assumed to be representative of all Singapore schools and at best, they suggested that it was probable that similar results might be obtained if the study was conducted in other schools and at similar age levels after controlling for intervention fidelity.

Furthermore, to run CFA, a sample size of 200 is recommended if the proposed model is not overly complex and if a maximum likelihood estimation is used (Hair, Anderson, Tatham, & Black, 1998). The values obtained for TLI and CFI which were 0.8 and 0.81 respectively were below the recommended value of > 0.90. School, class numbers, differing school levels and intervention duration could be increased in future

studies to test the generalisability of the findings and investigate how sustainable is the students' use of self-assessment across the school years.

The intervention students were taught by the researcher in the use of self-assessment in order to standardise the training and also as an attempt to minimise the intervention teacher effects on students' use of self-assessment. However, the study's quasi-experimental approach is unable to control for teacher variations in teaching effectiveness. It is possible that the study's results might be partly due to the way teachers taught Mathematics, with potential indirect effects on how students' perceptions of self-assessment might be affected especially if they did not see themselves improving (possibly due to the lack of quality instruction, for example). Increasing sample sizes should help to ameliorate this limitation.

Conclusion

This study can contribute to international educational research on how a systematic introduction, and training, of students as young as 9-10 years old in the use of self-assessment have resulted in them being more favourable in the use of self-assessment and could potentially incorporate it into their learning repertoire. Following Andrade (2010), this study also suggests that cognitive maturity does not seem to be a necessary condition for self-assessment as a younger group of students are able to use self-assessment. While students felt that using self-assessment has changed their learning and increased their interest in learning, they also expressed some lack of confidence in assessing themselves accurately, and a need for more self-assessment practice (Cassidy, 2007). One possible reason, at least in Singapore's context, is that teachers have traditionally been the ones assessing students, be it using conventional or alternative assessment. Students therefore need time to get used to assessing themselves through a formative-oriented self-assessment and disrupt the belief that assessment is always

summative in nature. Indeed, by indicating that they needed more time and practice in order to self-assess accurately, the students in this study demonstrated that they had the awareness that they could improve in their judgement of their work (Brown, 2008; Ross, 2006).

The favourable results revealed in this study suggest the potential for using self-assessment in a wider context and with a larger audience. In a strong examination-oriented learning culture like in Singapore, the formative nature of self-assessment can complement summative classroom assessments currently in place. Ultimately, self-assessment can add to the Singaporean classroom assessment repertoire not only because of the validity of the information it provides to the students about their own learning but because self-assessment is "feedback for oneself from oneself" (Andrade & Du, 2007, p. 160).

Notes

1. The Primary School Leaving Examination (PSLE) is a national examination conducted in Singapore annually at the end of the final year of students' six years of primary school education.

References

Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. L. Andrade and G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 90-105). New York: Routledge.

Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education, 32*(2), 159-181.

Andrade, H., Du, Y., & Wang, X. L. (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.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice, 48*(1), 12-19.

Assessment Reform Group (2002). *Assessment for Learning: 10 principles research-based principles to guide classroom practice*. Assessment Reform Group, London, United Kingdom. http://www.qca.org.uk/libraryAssets/media/4031_afl_principles.pdf

Black, P. (1997). *Testing: Friend or Foe?: Theory and Practice of Assessment and Testing (Master Classes in Education)*. Falmer Press.

Black, P., & Harrison, C. (2001a). Feedback in questioning and marking: The science teacher's role in formative assessment. *School Science Review, 82*(301), 39-46.

Black, P., & Harrison, C. (2001b). Self- and peer-assessment and taking responsibility: The science student's role in formative assessment. *School Science Review, 83*(302), 43-49.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 8-21.

Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: School of Education, King's College. See also *Phi Delta Kappan, 80*(2), 139-148.

Brown, W. (2008). Young children assess their learning: The power of the quick check strategy. *Young Children, 63*(6), 14-20.

Brown, S. & Glasner, A. (Eds.). (1999). *Assessment matters in higher education: Choosing and using diverse approaches*. Buckingham, UK: SRHE and Open University Press.

Cassidy, S. (2007). Assessing 'inexperienced' students' ability to self-assess: Exploring links with learning style and academic personal control. *Assessment & Evaluation in Higher Education, 32*(3), 313-330.

Cheah, Y. M. (1998). The examination culture and its impact on literacy innovations: The case of Singapore. *Language and Education, 12*, 192–209.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Dann, R. (2014). Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice, 21*(2), 149-166.

Earl, L. (2006). Assessment – A powerful lever for learning. *Brock Education, 16*(1), 1-15.

Earl, L. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2nd ed.). San Francisco: Corwin Press.

Fan, L. H. (2002). In-service training in alternative assessment with Singapore mathematics teachers. *The Mathematics Educator, 6*(2), 77-94.

Fontana, D., & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology, 64*(3), 407-417.

Fuchs, L.S. (2011). Mathematics intervention at the secondary prevention level of a multitier prevention system: Six key principles. Retrieved from: http://www.rtinetwork.org/essential/tieredinsstruction/tier2/mathintervention.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998) .*Multivariate Data Analysis* **(**5th ed.). Upper Saddle River, NJ , Prentice-Hall Inc.

Hair, J., Black, B., Babin, B., & Anderson, R. (2010). *Multivariate Data Analysis* (7th ed.). Upper Saddle River, NJ , Prentice-Hall Inc.

Harlen, W. & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice, 4*(3), 365 — 379.

Hogan, D., Chan, M., Rahim, R., Kwek, D., Khin, M. A., Loo, S. C., Sheng, Y. Z., & Luo, W. S. (2013). Assessment and the logic of instructional practice in Secondary 3 English and mathematics classrooms in Singapore. *Review of Education, 1*(1), 57-106.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Hughes, G. (2011). Towards a personal best: A case for introducing ipsative assessment in higher education. *Studies in Higher Education, 36*(3), 353–367.

Kramer-Dahl, A. & Kwek, D. (2011). Reading the home and reading in school: Framing deficit constructions as learning difficulties in Singapore English classrooms. In C. Wyatt-Smith, J. Elkins, & S. Gunn (Eds.), *Multiple perspectives on difficuluties in learning literacy and numeracy* (pp. 159-178). Dordrecht: Springer.

Koh, K. H., Lee, A. N., Tan, W., Wong, H. M., & Guo, L. B. (2005, May). *Looking collaboratively at the quality of teachers' assessment tasks and student work in Singapore schools.* Paper presented at the International Conference on Education – Redesigning Pedagogy: Research, Policy, Practice. National Institute of Education, Nanyang Technological University, Singapore.

Koh, K., Lee, A. N., Gong, W. G., & Wong, H. M. (2006). *Development of the Singapore prototype classroom assessment tasks: Innovative tools for improving student learning and performance.* Paper presented at the 32nd Annual Conference International Association for Educational Assessment, Singapore.

Koh, K. & Luke, A. (2009). Authentic and conventional assessment in Singapore schools: an empirical study of teacher assignments and student work. *Assessment in Education: Principles, Policy and Practice, 16*(3), 291-318.

Lasonen, J. (1995). A case study of student self-assessment in upper secondary education. *Contemporary Issues of Occupational Education in Finland,* 199-215. (ERIC Document Reproduction Service No. ED386561)

Leach, L. (2012). Optional self-assessment: Some tensions and dilemmas. *Assessment & Evaluation in Higher Education, 37*(2), 137-147.

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education, 10*(2), 209-220.

McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *educational HORIZONS, 87*(1), 40-49.

Ministry of Education. (1998). *Learning, creating, communicating: A curriculum review*. Retrieved July 11, 2000, from http://www1.moe.edu.sg/speeches/Curry%20Revue%20Report.htm

Ministry of Education. (2004). *To light a fire – Enabling teachers, nurturing students. Ministry of Education Work Plan Seminar 2004 Planning Document*.

Ministry of Education. (2008). Speech by Dr Ng Eng Hen, Minister for Education and Second Minister for Defence, at the MOE Work Plan Seminar 2008, on Thursday, 25 September 2008 at 9.30am at the Ngee Ann Polytechnic Convention Centre. http://www.moe.gov.sg/media/speeches/2008/09/25/speech-by-dr-ng-eng-hen-at-the-moe-work-plan-seminar-2008.php

Ministry of Education. (2009). *Report of the primary education review and implementation committee*. http://www.moe.gov.sg/initiatives/peri/files/peri-report.pdf

Ministry of Education. (2012). Education in Singapore: MOE corporate brochure. http://www.moe.edu.sg/about/files/moe-corporate-brochure.pdf

Mok, M. M. C., Lung, C. L., Cheng, D. P. W., Cheung, R. H. P., & Ng, M. L. (2006). Self-assessment in higher education: Experience in using a metacognitive approach in five case studies. *Assessment & Evaluation in Higher Education, 31*(4), 415-433.

Munns, G., & Woodward, H. (2006). Student engagement and student self-assessment: The REAL framework. *Assessment in Education, 13*(2), 193-213.

Nichols, S. L., & Berliner, D. C. (2008). Testing the joy out of learning. *Educational Leadership, 65*(6), 14-18

Nichols, S. L., Glass, G. V, & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives, 14*(1). http://epaa.asu.edu/epaa/v14n1/

Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education, 25*(1), 23-38.

Ross, J. A., Hogaboam-Gray, A., & Rolheiser, C. (2002). Student self-evaluation in Grade 5-6 mathematics effects on problem-solving achievement. *Educational Assessment, 8*(1), 43-59.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research & Evaluation, 11*(10). http://pareonline.net/getvn.asp?v=11&n=10 http://pareonline.net/pdf/v11n10.pdf

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Soh, K. C. (1993). Measuring motivation to learn Chinese and English through self-reported feelings and behaviours. Singapore Journal of Education, 13(1), 44-50

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan, 83*(10), 758-765.

Stobart, G. (2008). *Testing Times: The uses and abuses of assessment*. Abingdon: Routledge

Tan, Y. K., Chow, H. H., & Goh, C (2008). *Examinations in Singapore: Change and continuity (1891-2007).* Singapore; Hackensack, NJ: World Scientific.

Tejeiro, R. A., Gómez-Vallecillo, J. L., Romero, A. F., Pelegrina, M., Wallace, A., & Emberley, E. (2012). Summative self-assessment in higher education: Implications of its counting towards the final mark. *Electronic Journal of Research in Educational Psychology, 10*(2), 789-812.

Trunnel, P. (1992). Self-assessment. *Adventures in Assessment, Vol. 2*. Boston: SABES. http://www.sabes.org/resources/adventures/vol2/2trunnell.htm

Van Voorhis, C.R.W. & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology, 3*(2), 43-50.

Wiliam, D. (2006). Formative assessment: Getting the focus right. *Educational Assessment, 11*(3 & 4), 283-289.

Wiliam, D. (2008). Quality in assessment. In S. Swaffield (ed.), *Unlocking assessment: Understanding for reflection and application* (pp. 123-137). Abingdon: Routledge.

Wilson, J., & Johnson, P. (2000). Students thinking about their learning: Assessment to improve learning. *Educational Research Quarterly, 24*(2), 10-20.

Wong, H. M. (2012). *The perceptions of Singaporean teachers and students toward academic self-assessments* (Unpublished doctoral dissertation). National Institute of Education, Nanyang Technological University, Singapore.

Wong, H. M. (2016). I can assess myself: Singaporean primary students' and teachers' perceptions of students' self-assessment ability. *Education 3-13: International Journal of Primary, Elementary and Early Years Education, 44*(4), 442-457.

Yu, G. (2013). The case of St Margaret's Girls' College: How SLOA promotes self-assessment and peer assessment to enhance secondary school student English learning. In M.M.C. Mok (ed.), *Self-directed learning oriented assessments in the Asia-Pacific* (pp. 393-411). Dordrecht: Springer.

Table 1. Summary of Student Information

Class	A1	A2	B1	B2	Total
No. of P4 Students	37	36	38	35	146
Girls	19	15	16	19	69
Boys	18	21	22	16	77
Ave. Age of P4 Students	9 years 10 months	9 years 9 months	9 years 10 months	9 years 9 months	

Table 2. Factor Analysis and Reliability of SAQ

Category	CFA factor loadings	Cronbach's Alpha
	.48	
	.53	
	.52	
	.50	
	.55	
	.61	
	.53	
	.68	
	.58	
	.43	
	.53	
	.58	
	.57	
	.66	
	.66	
	.68	
	.60	
	.63	
	.71	
	.64	
	.75	
	.62	
	.66	
	.54	
	.62	
	.66	
	.54	
	.62	
	.58	
	.59	
	.43	
	.63	
	.57	
	.54	
	.40	
	.43	
	.59	
	.64	
	.59	
	.52	

Table 3. Confirmatory factor analyses: Comparison of models

		1 factor model	4 factors model
Fit measure	Macro	Default model	Default model
Discrepancy/df	CMIN/DF	2.01	1.705
Tucker-Lewis index	TLI	.715	.800
Comparative fit index	CFI	.730	.810
Parsimony comparative fit index	PCFI	.693	.762
RMSEA	RMSEA	.083	.070

Table 4. MANOVA Results

Variable(s)	Pillai's Trace	*F*	*df*	Error *df*	Sig	η_p^2	Effect Size
Knowledge Application	.255	49.401	1	144	.000	.255	Large
Knowledge Application * Group	.170	29.553	1	144	.000	.170	Medium
Independent Learning	.315	66.189	1	144	.000	.315	Large
Independent Learning * Group	.264	57.552	1	144	.000	.264	Large
Communication	.235	44.267	1	144	.000	.235	Large
Communication * Group	.177	31.048	1	144	.000	.177	Medium
Motivation	.310	64.781	1	144	.000	.310	Large
Motivation * Group	.130	21.577	1	144	.000	.130	Medium

Table 5. Mean scores and standard deviations of knowledge application, independent learning, communication and motivation

Domain	Group	Pre-intervention Mean (SD)	Post-intervention Mean (SD)	N
	Intervention	3.47 (.47)	4.04 (.46)	75
	Comparison	3.46 (.62)	3.53 (.69)	71
	Intervention	3.50 (.44)	4.13 (.44)	75
	Comparison	3.47 (.66)	3.51 (.73)	71
	Intervention	3.38 (.51)	3.94 (.53)	75
	Comparison	3.38 (.69)	3.43 (.75)	71
	Intervention	3.79 (.52)	4.40 (.46)	75
	Comparison	3.75 (.46)	3.92 (.63)	71

Table 6. Intervention Students' Responses to Self-Assessment Questions

Self-Assessment Questions for Intervention Students	Frequency	Percent
1. Self-assessment strategy that students liked best		
Checklist	22	29.3
Learning Log	1	1.3
Rubrics	52	69.3
2. What you have learnt most through self-assessment		
Knowledge Application	13	17.3
Independent Learning	31	41.3
Communication	12	16.0
Motivation	19	25.3
3. Frequency of willingness to use SA by students		
Every week	33	44.0
Every 2 weeks	16	21.3
Every month	10	13.3
Every term	16	21.3
4. Students' use of self-assessment on their own		
Told by teacher	27	36.0
Sometimes	40	53.3
Yes	8	10.7
5. Self-assessment to be considered as part of test marks		
No	42	56.0
Yes	33	44.0
6. Need more practice to be confident to assess own work		
No	8	10.7
Yes	66	88.0
Missing/ NA	1	1.3
7. Usage of self-assessment for other subjects		
No	4	5.3
Yes	69	92.0
Missing/ NA	2	2.7

Table 7. Individual Groups' Responses to Self-Assessment Questions

Self-Assessment Questions for Individual Groups	Group	No Number (%) of students	Yes Number (%) of students	N
	Intervention	1 (1.3)	74 (98.7)	75
	Comparison	12 (16.9)	59 (83.1)	71
	Intervention	4 (5.3)	71 (94.7)	75
	Comparison	6 (8.5)	65 (91.5)	71
	Intervention	19 (25.3)	56 (74.7)	75
	Comparison	22 (31.0)	49 (69.0)	71
	Intervention	12 (16.0)	63 (84.0)	75
	Comparison	19 (26.8)	52 (73.2)	71