Evaluating and comparing Singaporean and Taiwanese eighth graders’ conceptions of science assessment

Abstract
Researchers have indicated that assessment practices and methods ought to be tailored to support learners’ construction of meaningful understanding of knowledge. To achieve this aim, understanding students’ conceptions of science assessment would be essential since it will enable us to construct more realistic, valid and fair assessments. Understanding how learners conceptualize assessment would be imperative to serve as an essential reference to properly evaluate their learning progress. The main purpose of this study was to evaluate and compare the Singaporean and Taiwanese middle school students’ conceptions of science assessment. Within each country, gender comparisons were also explored. A total of 333 Singaporean and 424 Taiwanese grade eight students were invited to complete a questionnaire named Conceptions of Science Assessment (COSA) to capture their views on science assessment. The results indicated that, first, the COSA questionnaire was valid and reliable for measuring the Singaporean and Taiwanese eighth graders’ conceptions of science assessment, including Surface, Summative, and Formative conceptions. Second, the findings showed that the summative purpose of assessment still dominates in the Taiwanese science classroom, while both formative and summative purposes of assessment are usually perceived in Singaporean science classrooms. In contrast, the Singaporean students had a greater tendency than their Taiwanese counterparts to perceive the surface purpose of assessment as merely a way of reproducing scientific knowledge, and the formative purpose of assessment as improving learning, problem solving and critical judgment. No gender differences were found among either the Singaporean or the Taiwanese students regarding their three science assessment conceptions. The findings suggest that science educators in
both countries should provide learners with more opportunities to experience process-oriented science assessment activities and de-emphasize the usage of examination-oriented practices to achieve the sophistication of conceptions.

Introduction

Classroom assessment has been a crucial part of teaching and learning and a prime focus of researchers, educators, and policy makers (Bell, 2007). Given the multiple purposes of assessment, it can be used to, for example, improve teachers’ instructional practices or enhance the quality of students’ learning processes and outcomes (e.g., Bell & Cowie, 2001; Brown et al., 2009; Peterson & Irving, 2008). With the prevailing recognition of constructivism in science education, the purpose of assessment practices has seemingly been moving away from proving learning (i.e., assessment of learning) to improving learning (assessment for learning) (Shepard, 2000). It is argued that the current assessment practices in classrooms fail to allow learners to develop a clearer understanding of how they can improve in their learning and of how to prepare themselves to deal with the rapidly changing world (Birenbaum et al., 2006). Thus, what and how learners perceive and conceptualize classroom assessment practices may be an inevitable issue to contend with for researchers to improve learning and teaching.

Assessment researchers (e.g., Bell, 2007; Cowie, 2012) have raised the need to explore how learners view and conceptualize assessment since this issue is not well researched. Although, in recent years, researchers have explored teachers’ or students’ conceptions of assessment (e.g., Brown & Wang, 2013; Brown et al., 2011; Remesal, 2011; Wang & Brown, 2014), the studies probing students’ conceptions are insufficient and are mainly focused on the tertiary level. Understanding learners’ conceptions of assessment is important since these conceptions may influence their views on learning (Peterson & Irving, 2008; Vermetten, Vermunt, & Lodewijks, 2002),
adoption of learning strategies (Gijbels & Dochy, 2006), motivation (Brown, 2011; Brown, Irving, & Peterson, 2009) and academic achievements (Brown & Hirschfeld, 2008). In addition, learners’ conceptions of assessment may usually reflect and correspond to teachers’ instruction (Lee et al., 2013). The investigation of learners’ conceptions of assessment may provide researchers and educators with insights to improve the present classroom practices related to assessment as well as teaching and learning. Similarly, although the significance of assessment of science learning has been widely stressed in science education, still limited empirical research has been conducted to examine learners’ conceptions of science assessment (Wang, Kao, & Lin, 2010). This study thus aimed to mitigate this gap in our understanding of learners’ conceptions of assessment in science education.

In recent times, one of the trends of educational assessment has been that the practice of classroom assessment could be viewed as a socio-cultural practice of teachers and students (Broadfoot, 2002; Cowie, 2005). In other words, classroom assessment could be regarded as “a teacher and student practice embedded in political, historical, social, and cultural contexts” (Bell, 2007, p.966). Further, recent research findings have suggested that the formation of conceptions of assessment may be culturally sensitive (Hirschfeld & Brown, 2009; Otunuku, Brown, & Airini, 2013; Sergers & Tillema, 2011). It is hence reasonable to envisage that the socio-cultural as well as educational contexts may influence how assessment practices are conceived by students in classrooms. Also, Brown et al. (2011) have stressed that the cultural and policy practices of assessment in East Asian countries may be quite different from those in Western countries. This difference may be attributed to the cultural features of the Confucian system. Thus, it would be beneficial to investigate and compare students’ conceptions of assessment from those Confucian heritage countries that share common socio-cultural roots such as Singapore and Taiwan to enrich the
knowledge base of assessment conceptions in different societies and cultures.

Literature review

Classroom assessment in science learning

Classroom assessment is a crucial part of teaching and learning. Two types of classroom assessment are generally recognized in the assessment literature, namely the summative and formative functions of assessment (e.g., Bell, 2007; Shepard, 2000). In general, the summative purpose of assessment tends to judge students’ learning and emphasize the product of learning, whereas the formative purpose of assessment aims to improve learning and highlight the process of learning (Black, 2001; Shepard, 2000). The two distinct views regarding assessment may also reflect a transition from a subject matter-centered perspective to a learner-centered paradigm (Segers & Dochy, 2001). Researchers (e.g., Lee et al., 2013; Panizzon & Pegg, 2008) have suggested that the quantitative assessment practices (e.g., summative assessment) may involve the accretion of facts and concepts to determine a student’s intelligent ability, while qualitative practices (e.g., formative assessment) may place emphasis on improving a student’s higher-order skills and advanced knowledge. In turn, the summative purpose of assessment mainly focuses on assessing how much students have learned at the conclusion of the learning process. In contrast, the formative purpose of assessment not only emphasizes how well students have learned during the process of learning but also provides meaningful feedback to facilitate the learning progress forward.

In science education, given that promoting students’ engagement in meaningful learning is of great importance, science education researchers have paid much attention to assessment for formative purposes in the science classroom (Bell & Cowie, 2001; Coffey et al., 2011). A handful of studies (Furtak & Ruiz-Primo, 2008;
Ruiz-Primo & Furtak, 2007) have attempted to investigate the effects of teachers’ assessment practices on students’ science learning. For instance, emphasizing the notion of formative assessment practice, Furtak and Ruiz-Primo (2008) developed several formative assessment prompts to elicit students’ ideas about science concepts during their learning processes and to support various classroom activities. It has been claimed that such assessment practices for formative purposes not only assist teachers to gather and interpret information about students’ learning processes in order to adapt instruction (Bell & Cowie, 2001), but can also redirect students’ thinking to actively improve their own learning progress and achieve learning goals (Furtak & Ruiz-Primo, 2008). In addition, aligned with the prevailing notion of constructivist learning in science education, science education researchers (Atkin & Coffey, 2003) have claimed that assessment practices and methods should be tailored to support learners’ construction of meaningful understanding in scientific inquiry instead of rote learning that results in memorization of curriculum content. The alignment between students’ conceptions of assessment and teachers’ instructional practices has thus become an important issue (Wang et al., 2010). Accordingly, a comprehensive understanding of how learners conceive of classroom assessment practices may be an inevitable and fundamental issue for science educators as well as researchers to note in order to achieve the goals of the science curriculum.

**Research on Conceptions of Assessment**

A conception could be defined as individuals’ mental representations through which they understand, respond to, and interact with a particular phenomenon (Marton, Dall’Alba, & Beauty, 1993). Theoretically rooted in the line of phenomenographic research, various educational conceptions such as conceptions of learning, assessment, and teaching have been recognized by researchers (e.g., Brown,
Conceptions of assessment signify how individuals interact, respond to, or experience assessment activities implemented in learning environments (Peterson & Irving, 2008). In other words, how students conceive of assessment in educational contexts may be derived from their own learning experiences. Several researchers agree that students’ conceptions of assessment are of particular importance because assessment has a significant impact on the quality of learning outcomes (e.g., Brown & Hirschfeld, 2008; Marton et al., 1997). For example, if assessment scores are poor for a specific science topic, teachers are likely to spend more time revising the topic with students to ensure that they understand.

In this line of research, assessment researchers have begun to investigate students’ conceptions of assessment in a general sense (e.g., Brown, 2004; Brown & Hirschfeld, 2007, 2008; Brown et al., 2011; Watkins, Dahlin, & Ekholm, 2005; Remesal, 2011). Derived from the studies conducted by Brown and his colleagues regarding students’ conceptions of assessment (e.g., Brown & Hirschfeld, 2008; Brown et al., 2009; Hirschfeld & Brown, 2009), several conceptions of assessment such as improving achievement, making students accountable for their own learning, or being irrelevant were commonly identified.

More recently, a few researchers have conducted empirical studies to characterize individuals’ conceptions of assessment in the domain of science (Wang et al., 2010; Lee et al., 2013). Focusing on the various dimensions of assessment (content, processes, and attitudes) and methods of assessment, Wang et al. (2010) explored Taiwanese pre-service teachers’ conceptions of assessment of science learning and found that these conceptions were relatively consistent with their views of learning science. In a study contributed by Lee et al. (2013), an endeavor to understand Taiwanese high school students’ conceptions of science assessment was
made. Utilizing qualitative and quantitative methods, six distinct conceptions of
science assessment were identified, including science assessment as reproducing
knowledge to others (Reproducing), as rehearsing for practice for examinations
(Rehearsing), as making the students accountable for their own learning
(Accountability), as improving learning (Improving), as problem solving (Problem
Solving), and as having others critically judge one’s learning (Critical Judgment).
Besides, the result of a second-order confirmatory factor analysis of the
newly-developed “Conceptions of Science Assessment Questionnaire (COSA)”
revealed that the six categorized conceptions could reflect three main purposes of
science assessment (i.e., surface, summative, and formative). It should be noted that
the surface purpose of assessment (i.e., Reproducing conception) centers on
“…memorizing separate facts and reproducing terms, reflecting a rote-learning
situation” (Lee et al., 2013, p.258). The summative purposes of assessment refer to
the “Rehearsing” and “Accountability” conceptions, while the formative ones
correspond to the “Improving learning,” “Problem Solving,” and “Critical Judgment”
conceptions.

Even though the exploration of students’ conceptions of science assessment has
received much attention recently, empirical studies are still arguably insufficient. For
example, as suggested in the literature (e.g., Struyven, Dochy, & Janssens, 2005;
Peterson & Irving, 2008; Turner & Gibbs, 2010), students’ perceptions of assessment
practices and environments may vary in terms of gender. It remains unclear whether
such gender differences exist in middle school students’ conceptions of science
assessment. Besides, research evidence (e.g., Brotman & Moore, 2008; Hong & Lin,
2013) has reported notable discrepancies in terms of gender in students’ involvement
in science learning. It is possible that the differences between male and female
students’ conceptions of science assessment would be one of the potential contributors
to this phenomenon. Thus, a focus on the gender issue in this regard may provide useful insights to inform the science education community.

**The Socio-Cultural and Educational Contexts of Singapore and Taiwan**

Singapore and Taiwan are two Asian countries sharing a Confucian-heritage learning culture that exhibit considerable similarities and differences in how educational policies are implemented and realized. One of the renowned similarities is the emphasis on rigorous assessment. In Singapore, three main national examinations from elementary (the Primary School Leaving Examination, PSLE), secondary (the General Certificate of Education (GCE) O-level examination), to pre-university (the GCE A-level examinations) levels are carried out to track students. In Taiwan, the nationwide high-stakes examinations still remain the primary pathway to enter both senior high school and university. As a society, Taiwanese remain highly respectful of authoritative figures such as parents and teachers, and hence, students experience pressure from them to be successful in these examinations (Lin, Tan, & Tsai 2013). This situation is mirrored in Singaporean society. However, the differences between the educational systems in Singapore and Taiwan may be worth noting. For instance, the education system of Singapore is a highly centralized one with centrally planned school curricula and assessment frameworks (Kim, Tan, & Talaue, 2013; Tan, 2011). Thus, Singaporean teachers’ practices typically closely follow the pre-planned curricular scripts. In contrast, the Ministry of Education in Taiwan has tended to emphasize a “decentralization” policy regarding the curriculum to provide local schools and teachers with more flexibility and autonomy to organize learning and teaching activities as long as their practices are aligned with the curriculum guidelines (Huang, 2012).

At the turn of the century, an alternative school-based assessment reform called
the Science Practical Assessment (SPA) initiative at the secondary and pre-university levels was implemented in Singapore to partially replace the one-time high-stakes national practical examinations (Ng, 2008). This particular initiative that was integrated into the existing science curriculum aimed to highlight authentic scientific inquiry and investigations as a way of learning science. Science teachers have equipped themselves with the ability to use diverse assessment approaches (e.g., summative and formative) to meet accountability requirements and appropriately develop students’ understanding of science and practical work (Tan & Towndrow, 2009). More recently, in 2008, there was a renewed emphasis on expanding the assessment practices in Singapore. Formative assessment and alternative assessment modes were introduced and widely adopted across all subjects (including science) taught in schools. For example, in recent years, the Singapore education system has introduced alternative modes of assessment such as project-based learning or school-based assessment of science practical work skills in addition to traditional high-stakes assessments to provide more holistic feedback for pupils, teachers, and parents. Moreover, the current national science curriculum incorporated inquiry learning as the core element in 2008 (Kim et al., 2013). Although the learning content may remain the same, the new inquiry-centered curriculum emphasizes the pedagogical approaches aligned with the tenets of constructivism (Jocz, Zhai, & Tan, 2014).

Similarly, the Taiwanese curriculum guidelines in recent years embrace Western philosophy and constructivist pedagogy (Chiu & Whitebread, 2011). Although the past research has documented that science teachers in Taiwan may be familiar with the constructivist-oriented instruction, they still focus on the science content knowledge to prepare students for the examinations as the top priority (Chang, Chang, & Yang, 2009; Lee, Chang, & Tsai, 2009). Since the year of 2014, the compulsory
education system in Taiwan has been extended from 9 to 12 years (Ministry of Education, 2015). A traditional standardized examination in terms of total score for entering senior high schools was replaced with “Comprehensive Assessment Program for Junior High School Students” which emphasizes standards-based assessment and primarily grades students with only three proficiency levels, including mastery, basic, below basic (Sung, Chou, & Tseng, 2014). Students, in turn, will be placed in nearby senior high schools according to their residential school districts to reduce the intensive competitions for entering prestigious senior high schools and discard the testing-oriented teaching and learning (Gou & Chiu, 2016).

The traditional Taiwanese culture has been influenced by the concept of filial piety that children (especially males) have the responsibility to support their elderly parents and relatives (Tsai, Yang, & Chang, 2015). Although this phenomenon may be prevalent in Taiwanese society, the goal of a recent reform aims to mitigate the gender gaps and promote gender equity and fairness since 2004 (Lin et al., 2013). A recent investigation conducted by Hong and Lin (2011) indicated that Taiwanese female students expressed higher interest in science than male students did, implying the success of reducing the image of masculine science in school science. In addition, Singapore, a country with more than 75% Chinese population, traditionally inherits similar Chinese culture as Taiwan does (Lee, Tsai, & Chai, 2012). Areepattamannil et al. (2015) also found significant gender differences in the Singaporean students’ science achievement (favoring males) by means of the Program for International Student Assessment (PISA) 2009 data, indicating that school science seems to be prone to a male domain in Singapore. However, Jocz, Zhai, and Tan (2014) found that gender was not a predictor of interest in science for Singaporean grade 4 students. This observation was also supported by qualitative classroom data gained from this study whereby there is active involvement in the science learning process by both
boys and girls. In this aspect, we argue that the classroom culture of equal opportunity for learning influenced the students’ learning more so than traditional societal stereotype of male domination in science.

**Research purpose and questions**

The primary aim of this study was to evaluate and compare the Singaporean and Taiwanese eighth graders’ conceptions of science assessment. For the current study, the specific research questions were formulated as follows:

1. Is the questionnaire adopted in this study valid and reliable for assessing the Singaporean and Taiwanese eighth graders’ conceptions of science assessment?
2. What are the Singaporean and Taiwanese eighth graders’ conceptions of science assessment?
3. What are the differences, if any, between the Singaporean and Taiwanese eighth graders’ conceptions of science assessment?
4. What are the gender differences, if any, within the Singaporean and Taiwanese eighth graders’ conceptions of science assessment?

**Methods**

**Participants**

A total of 424 Taiwanese eighth graders (217 males and 207 females) were invited from junior high schools in northern, central, and southern Taiwan to participate in this study. The average age of the Taiwanese participants was 14.07 years old. In addition, 333 Singaporean eighth graders (180 males and 153 females) were invited for a cross-country comparison of their conceptions of science assessment. These Singaporean participants were from three different secondary schools, and had an average age of 13.76 years. Data were gathered by means of the
quantitative survey method. After acquiring permission from the school administrators and class teachers, the adopted instrument was distributed to the participating students in the same school semester. All the student participants completed the given instrument within a thirty-minute time period.

**Instrumentation**

All participants completed a 30-item questionnaire that was adapted from the Conceptions of Science Assessment (COSA) questionnaire developed by Lee et al. (2013). It should be noted that the original version of COSA contained 34 items. In order to adequately compare the conceptions of science assessment between the two groups of students, several rounds of review were carried out. First, since the original version of COSA was created in Chinese, the researchers translated the Chinese version into English for the Singaporean students. This initial English version was then approved and verified by two science education experts who are familiar with this research topic. Next, a series of translation and back translation processes suggested by Brislin (1970) were initiated to ensure the content validity and comparability of the two versions. After the establishment of the English version, the two science education experts from Taiwan and Singapore also evaluated the appropriateness of the two versions of the COSA questionnaire for cross-country comparison. Four redundant items with similar statements were removed. It should be noted that the wording of each statement in Chinese version was carefully reviewed by high school students in the previous study (i.e., Lee et al., 2013). The English statements were scrutinized by a handful of Singaporean students and their teachers to ensure the readability of survey items. As a result, the original 34-item questionnaire was reduced to 30 items for both the Chinese and English versions.

A detailed description of each dimension of the COSA questionnaire with
corresponding sample items is presented below:

(1) Reproducing knowledge/5 items: Science assessment is regarded as a way of reproducing scientific knowledge. A sample item is “School science assessment is to write down formula and definitions from the textbook.”

(2) Rehearsing/5 items: Science assessment is characterized as a way of familiarizing oneself with scientific knowledge. A sample item is “School science assessment aims to help me understand scientific knowledge more clearly through repeatedly practicing and rehearsing.”

(3) Accountability/5 items: Science assessment is to make students accountable for their own learning. A sample item is “School science assessment shows me my current learning level.”

(4) Improving learning/5 items: Science assessment is seen as improving learning activity. A sample item is “School science assessment is to increase my understanding of science concepts.”

(5) Problem solving/5 items: Science assessment is viewed as problem solving that extends science knowledge to open-ended contexts. A sample item is “School science assessment is to apply the science knowledge and skills learnt in a practical manner.”

(6) Critical judgment/5 items: Science assessment is conceptualized as evaluating knowledge claims, challenging one’s understandings, and exploring the value of knowledge. A sample item is “School science assessment is to evaluate different concepts and theories in a more rational way.”

In the present study, both the Chinese and English versions of COSA were presented in a five-point Likert mode, ranging from 5, “strongly agree” to 1, “strongly disagree.” Therefore, students gaining higher scores in a certain dimension show stronger agreement with the statements in that COSA dimension.
**Data analysis and procedure**

Following established practice in the literature (e.g., Lee et al., 2013; Segers & Tillema, 2011), an exploratory factor analysis (EFA) with the method of varimax rotation was employed to clarify the structure of COSA for the two versions. This approach aimed to analyze the commonalities and differences between survey items based on the participants’ responses. In the EFA, only those items with a factor loading of at least 0.4 within their own factor were retained (Stevens, 1996), and items with many cross-loadings were omitted (Bentler, 1990). In addition, the purpose of EFA was to retain the same items within each scale in both versions for later comparison. In order to generate a shared meaningful factor structure, by means of the split file function of Statistical Product and Service Solutions (SPSS) software, the participants’ responses on both versions could be computed simultaneously to acquire the same retained items. This approach is similar to several previous cross-cultural studies (e.g., Arino de la Rubia, Lin, & Tsai, 2014; Lin et al., 2013). The reliability of each version was evaluated in terms of the Cronbach’s alpha coefficient. Accordingly, the construct validity and reliability of the two COSA versions were established.

Next, a series of paired comparisons (i.e., paired $t$-tests) of the students’ means of COSA between the Taiwanese and Singaporean students were conducted for understanding their COSA. Furthermore, independent $t$-tests were performed to examine cross-cultural differences in the Taiwanese and Singaporean students’ responses to the COSA questionnaire. Then, in order to explore the gender differences within each country, independent $t$-tests were subsequently utilized. It should be noted that the Cohen’s effect size index $d$ was computed for either paired $t$-tests or independent $t$-tests to illustrate the practical differences. As indicated by Cohen (1988), Cohen’s $d$ values of 0.20, 0.50, 0.80, and 1.0 are interpreted as small, medium, large, and very large, respectively.
Results

The factor analysis of the COSA questionnaire

In order to clarify the factor structure of the COSA questionnaire in each version, an EFA with the principal component method was conducted for each country. Thus, as depicted in Table 1 and Table 2, a total of 18 items were retained according to the Singaporean and Taiwanese participants’ responses. In both versions, the retained items were grouped into the three identical meaningful factors based on the three purposes of science assessment as indicated by Lee et al. (2013), including Surface, Summative and Formative COSA. All items in terms of their factor loadings were weighted greater than 0.4 on the relevant factors, and the eigenvalues of the three factors were all larger than 1. The total variance explained for the Chinese-version/English-version COSA was 69.11% and 55.10%, respectively. Table 1 and Table 2 also illustrate the reliability in terms of Cronbach’s alpha coefficients of the two COSA versions. That is, the three COSA factors of the English-version/Chinese-version were 0.85/0.72, 0.93/0.84, and 0.93/0.88, and the overall alpha values were 0.93/0.89, respectively. In sum, the EFA results indicated that the three factors that emerged had sufficient reliability and construct validity for assessing the Singaporean and Taiwanese eighth graders’ COSA.

The paired comparison for the obtained factors of COSA within each country

Table 3 delineates the 333 Singaporean students’ factor means, the standard deviations, and the results of the paired comparisons of the three COSA factors. As shown, both Summative COSA (Mean = 3.62, SD = 0.68) and Formative COSA (Mean = 3.61, SD = 0.76) were the most often recognized, while Surface COSA (Mean = 3.18, SD = 0.68) was the least perceived among these Singaporean students.
For their Taiwanese counterparts, Summative COSA was the feature most frequently recognized (Mean = 3.51, SD = 0.87), followed by Formative COSA (Mean = 3.32, SD = 0.89) and Surface COSA (Mean = 3.05, SD = 0.84) as shown in Table 4. In sum, it seems that the summative purpose of assessment still dominates in the Taiwanese science classroom, while both formative and summative purposes of assessment are usually perceived in the Singaporean science classroom. Besides, the surface purpose of science assessment is the least perceived by both the Singaporean and Taiwanese eighth graders in this study.

Cross-cultural comparisons

To understand the cross-cultural differences in the Singaporean and Taiwanese eighth graders’ conceptions of science assessment, a series of independent t-tests were performed. The results of the t-tests and the Cohen’s effect size index d are presented in Table 5. In general, the differences could be attributed to the Surface and Formative COSA between the two student groups. More specifically, the Singaporean students had significantly higher scores on the Surface COSA factor ($t = 2.45, p < 0.05$) and Formative COSA factor ($t = 4.74, p < 0.001$) than the Taiwanese students did. In other words, the Singaporean eighth graders had a greater tendency than the Taiwanese students to perceive the surface purpose of assessment as merely a way of reproducing scientific knowledge, and the formative purpose of assessment as improving learning, problem solving and critical judgment. The Cohen’s $d$ values also indicated close to small to close to medium practical effects of the difference between the two countries in the three dimensions of the COSA ($d = 0.14$-$0.35$).

Gender differences within the two countries

In addition to the cross-country comparisons, this study further examined the
gender differences in each country by means of independent \( t \)-tests. As summarized in Tables 6 and 7, neither the Singaporean nor the Taiwanese male and female students showed any significant differences \((t = -0.63~1.95, p > 0.05)\), suggesting that, both in Taiwan and in Singapore, the male and female students perceived science classroom assessment and characterized their conceptions of science assessment similarly, without any salient differences.

**Discussion and implications**

This is one of the few studies that have attempted to understand middle school students’ conceptions of science assessment. Moreover, this study further attempted to compare the students’ conceptions of science assessment from two Confucian heritage countries that share similar socio-cultural roots (i.e., Singapore and Taiwan). To answer the research questions for this study, a quantitative instrument developed by Lee et al. (2013) was adopted to investigate and compare the participants’ conceptions of science assessment in the present study.

After ensuring the factor structure by means of exploratory factor analysis (EFA), three meaningful factors, including Surface, Summative, and Formative COSA were found in both the Singaporean and Taiwanese samples. The three purposes of COSA reflect the Singaporean and Taiwanese students’ experiences of science classroom assessment activities and form such fundamental conceptions. Since this study was explorative and preliminary in nature, EFA approach was adopted to generate a shared meaningful factor structure for comparison and provide useful insights for future research. Researchers are encouraged to adopt more advanced statistical methods such as confirmatory factor analysis and structural equation modeling technique with a larger sample to conduct relevant studies to reassure the results derived from this study. In addition, in-depth interview could be utilized as an alternative and
complementary approach to understand how students interpret the survey items with respect to their assessment experiences.

It should be noted that these middle school students did not differentiate the six distinct conceptions of assessment as proposed by Lee et al. (2013). Conceptions of assessment are formed and aligned with learners’ educational experiences (Brown, 2011; Segers & Tillema, 2011). The results of this study may suggest that the participants did not have a clearly differentiated view regarding the purpose of science assessment. As indicated by Peterson and Irving (2008), secondary school students may be situated in a transitional stage of developing their understanding of assessment practices while encountering different assessment purposes and procedures. It is possible that the current assessment practices in both countries did not provide sufficient opportunities for the students to engage in, experience, or even differentiate such activities.

Another explanation is related to whether teachers make explicit references to the purposes of assessment in the classroom. Studies (e.g., Brown & Harris, 2012; Peterson & Irving, 2008) have suggested that teachers should explicitly inform students what, how and why science assessment is implemented to develop in uniform ways towards a shared understanding. Otherwise, students may easily misinterpret the underlying the purpose and nature of assessment tasks. Although the purpose of assessment has been clearly documented in Singapore and Taiwan national curriculum guidelines, it is reasonable to infer that teachers in both countries may infrequently to achieve this goal based on the results. Since the six assessment conceptions are supported by and derived from the three main purposes of science assessment, practitioners could cultivate and guide students to develop corresponding conceptions.

This study further reveals that, in general, the Singaporean participants tended to view both summative and formative purposes of science assessment as equally
important, whereas their Taiwanese counterparts still emphasized the importance of summative assessment. As previously mentioned, the Science Practical Assessment initiative as well as the recent emphasis on formative assessment system-wide in Singapore seems to have deepened students’ conceptual understanding and problem-solving skills, and may have transformed their views about science assessment from merely involving product-oriented assessment work to more of a focus on learning. Therefore, the Singaporean participants in this study may not only value the feedback of science assessment as proving how much they have learned (i.e., summative) but also perceive it as a viable way to improve their science learning and competency during the assessment process (i.e., formative). This finding is not unexpected in the strong national emphasis on improvement and continuous learning. Singapore students are enculturated into the mindset that assessment of various forms are necessary and is an essential part of learning.

On the contrary, in Taiwan, the cultural climate and educational context still shape the Taiwanese students’ conceptions of the purpose of science assessment as proving learning (i.e., summative). Due to the examination-oriented assessment system, Taiwanese students may frequently and mainly perceive the assessment work as a certain kind of product-oriented assessment that stresses recalling information or practicing science formulae and concepts in order to demonstrate their accountability and get into highly reputable senior high schools (Hong & Lin, 2013; Lee et al., 2013). In turn, this result may suggest that science educators in Taiwan should de-emphasize summative assessment types (e.g., paper-and-pencil/multiple-choice tests) that may influence students’ summative assessment purpose conceptions such as reproducing knowledge, rehearsing, or accountability. At the same time, science teachers should provide students with more opportunities to engage in authentic assessment activities such as knowledge building, socio-scientific argumentation, or peer assessment tasks.
that highlight higher-order thinking and problem solving skills in authentic contexts to foster their process-oriented conceptions of science assessment.

Additionally, in comparison, the results from this study also indicated that, by and large, the Singaporean students tended to score higher on the surface purpose of assessment and the formative purpose of assessment than their Taiwanese counterparts. This presents a potential dilemma to the Singaporean students. While they may appreciate the value of formative assessment in helping their learning, they still regard science assessment as merely reproducing knowledge (i.e., Surface COSA) more often than the Taiwanese students did. This result may be related to the assessment practice of the current education system in Singapore. On the one hand, it seems that, as previously mentioned, while the school-based schemes of assessment as well as the emphasis on formative assessment practices in Singapore are pervasively used in science classrooms, the looming national examinations and annual school examinations are also very familiar experiences for every student. The formative and summative assessment practices are dominant in different time cycles of the school year, and hence students in Singapore are well aware of the potential tensions of the various forms of assessment they experience.

On the other hand, Kim et al. (2013) revealed that many Singaporean elementary science teachers do not hold science-related degrees, thus limiting their ability to effectively implement scientific inquiry activities. Researchers (e.g., Lee et al., 2013; Wang et al., 2010) have recognized the value of scientific inquiry as a form of formative assessment that enables learners to critically evaluate knowledge claims or challenge one’s understandings. Thus, the effectiveness of inquiry teaching may have an impact on how Singaporean learners perceive the process of inquiry and form their conceptions of assessment. The findings also indicate that Singaporean teachers may easily express their concerns about the tension between how to appropriately
implement process-oriented assessment and simultaneously meet the requirements of the national examinations. The dominant social discourse in Singapore is one that places the importance of academic achievements in helping learners secure a good future. This is likely to be rooted historically to the emphasis that Singapore placed on nurturing her human resource through education. The forefathers of this young nation impressed upon the citizens that in the absence of any natural resources, development of human talent is the only way to ensure economic survival and success of the country. As such, in the 50 years of Singapore independence, educational wastage is not tolerated and each phase of educational change ensured that we maximize the educational outputs (Poon, 2014). In such circumstances, Singaporean students are reminded, both in school and at home of the importance of doing well in assessments. As such, they may feel bewildered, and this in turn, resulted in their inadequate conceptions of science assessment as merely reproducing science content and knowledge. This phenomenon may endanger the adequate maturation of students’ COSA. The findings in this study certainly provide useful insights for science education researchers and practitioners in Singapore to narrow the potential gap between the recent assessment innovation and assessment processes carried out in schools.

Another issue investigated in the present study is the gender difference regarding the participants’ COSA. Although a number of past studies (e.g., Peterson & Irving, 2008; Turner & Gibbs, 2010) have suggested potential gender differences in students’ perceptions of assessment tasks, this study did not find salient gender differences in the three grouped COSA, including Surface, Summative and Formative conceptions within each country. One of the interpretations may be related to the extent of congruence between instructional objectives and assessment tasks (e.g., Dhindsa, Omar, & Waldrip, 2007). On the one hand, for the Singaporean students, the required
practical assessment practices are rigorously implemented to ensure that every student undergoes and perceives similar assessment activities (Tan, 2011; Towndrow et al., 2010). And as highlighted by Jocz, Zhai and Tan (2014), both boys and girls are given similar and equal opportunities to engage in learning science in the classroom. In other words, this is very likely to create a gender-blind culture in Singapore school science. On the other hand, for the Taiwanese students, the recent curriculum reform has been striving to reduce the stereotypes of gender differences and the image of “masculine” science, and has embraced a constructivist stance that can formatively assess each student’s learning progression (e.g., Lee et al., 2013; Lin et al., 2013). Taken together, it is reasonable to imagine that each Singaporean and Taiwanese male or female student may experience similar assessment practices in the classroom. Based on the gained findings and provided interpretations here, further investigations are encouraged and warranted.

In sum, some implications derived from the abovementioned discussions may be informative. First, science educators in both countries should provide learners with more opportunities to experience process-oriented science assessment activities and de-emphasize the usage of examination-oriented practices as merely reproducing knowledge to achieve the sophistication of conceptions. It is also imperative to equip in-service teachers with adequate understanding of the principles related to formative assessment by means of professional development activities. Moreover, the various assessment conceptions derived from the three main purposes of science assessment may be useful to provide a conceptual framework for the critical review of teacher education programs and curriculum guidelines. Science teachers can also apply it as a pedagogical framework to design appropriate assessment practices in both countries.
References

‘Otunuku, M., Brown, G. T. L., & Airini (2013). Togan secondary students’ conceptions of schooling in New Zealand relative to their academic achievement. Asia Pacific Education Review, 14, 345-357.


<table>
<thead>
<tr>
<th></th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Surface COSA</td>
<td>Summative COSA</td>
<td>Formative COSA</td>
</tr>
<tr>
<td>Factor 1: Surface COSA, alpha = 0.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface 1</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface 2</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface 3</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface 4</td>
<td>0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface 5</td>
<td>0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor 2: Summative COSA, alpha = 0.93</td>
<td></td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Summative 1</td>
<td></td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Summative 2</td>
<td></td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Summative 3</td>
<td></td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Summative 4</td>
<td></td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>Summative 5</td>
<td></td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>Summative 6</td>
<td></td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Summative 7</td>
<td></td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Factor 3: Formative COSA, alpha = 0.93</td>
<td></td>
<td></td>
<td>0.80</td>
</tr>
<tr>
<td>Formative 1</td>
<td></td>
<td></td>
<td>0.80</td>
</tr>
<tr>
<td>Formative 2</td>
<td></td>
<td></td>
<td>0.81</td>
</tr>
<tr>
<td>Formative 3</td>
<td></td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>Formative 4</td>
<td></td>
<td></td>
<td>0.72</td>
</tr>
<tr>
<td>Formative 5</td>
<td></td>
<td></td>
<td>0.77</td>
</tr>
<tr>
<td>Formative 6</td>
<td></td>
<td></td>
<td>0.69</td>
</tr>
</tbody>
</table>

Overall alpha: 0.93
Total variance explained: 69.11%
Table 2. Rotated factor loadings and Cronbach’s alpha values for the English version of the COSA questionnaire

<table>
<thead>
<tr>
<th>Factor 1: Surface COSA, alpha = 0.72</th>
<th>Factor 2: Summative COSA, alpha = 0.84</th>
<th>Factor 3: Formative COSA, alpha = 0.88</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface 1</td>
<td>0.75</td>
<td>Formative 1</td>
</tr>
<tr>
<td>Surface 2</td>
<td>0.60</td>
<td>Formative 2</td>
</tr>
<tr>
<td>Surface 3</td>
<td>0.67</td>
<td>Formative 3</td>
</tr>
<tr>
<td>Surface 4</td>
<td>0.68</td>
<td>Formative 4</td>
</tr>
<tr>
<td>Surface 5</td>
<td>0.66</td>
<td>Formative 5</td>
</tr>
<tr>
<td>Summative 1</td>
<td>0.47</td>
<td>Formative 6</td>
</tr>
<tr>
<td>Summative 2</td>
<td>0.48</td>
<td>Overall alpha: 0.89</td>
</tr>
<tr>
<td>Summative 3</td>
<td>0.70</td>
<td>Total variance explained: 55.10%</td>
</tr>
<tr>
<td>Summative 4</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Summative 5</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Summative 6</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Summative 7</td>
<td>0.60</td>
<td></td>
</tr>
</tbody>
</table>

Overall alpha: 0.89
Total variance explained: 55.10%
Table 3. The paired comparison for the Singaporean eighth graders’ COSA

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean (SD)</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summative COSA</td>
<td>3.62 (0.68)</td>
<td>10.22***</td>
<td>0.65</td>
</tr>
<tr>
<td>Surface COSA</td>
<td>3.18 (0.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summative COSA</td>
<td>3.62 (0.68)</td>
<td>0.35</td>
<td>0.01</td>
</tr>
<tr>
<td>Formative COSA</td>
<td>3.61 (0.76)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formative COSA</td>
<td>3.61 (0.76)</td>
<td>8.78***</td>
<td>0.65</td>
</tr>
<tr>
<td>Surface COSA</td>
<td>3.18 (0.68)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note 1: *** p < 0.001  
Note 2: Effect size (Cohen’s d): small (0.2 to 0.5), medium (0.5 to 0.8) and large (0.8 and higher)

Table 4. The paired comparison for the Taiwanese eighth graders’ COSA

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean (SD)</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summative COSA</td>
<td>3.51 (0.87)</td>
<td>9.64***</td>
<td>0.54</td>
</tr>
<tr>
<td>Surface COSA</td>
<td>3.05 (0.84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summative COSA</td>
<td>3.51 (0.87)</td>
<td>6.64***</td>
<td>0.22</td>
</tr>
<tr>
<td>Formative COSA</td>
<td>3.32 (0.89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formative COSA</td>
<td>3.32 (0.89)</td>
<td>5.31***</td>
<td>0.31</td>
</tr>
<tr>
<td>Surface COSA</td>
<td>3.05 (0.84)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note 1: *** p < 0.001  
Note 2: Effect size (Cohen’s d): small (0.2 to 0.5), medium (0.5 to 0.8) and large (0.8 and higher)

Table 5. Country differences between the Singaporean and Taiwanese eighth graders

<table>
<thead>
<tr>
<th>Factor</th>
<th>Singapore (N = 333)</th>
<th>Mean (SD)</th>
<th>Taiwan (N = 424)</th>
<th>Mean (SD)</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface COSA</td>
<td>3.18 (0.68)</td>
<td></td>
<td>3.05 (0.84)</td>
<td></td>
<td>2.45*</td>
<td>0.17</td>
</tr>
<tr>
<td>Summative COSA</td>
<td>3.62 (0.68)</td>
<td></td>
<td>3.51 (0.87)</td>
<td></td>
<td>1.85</td>
<td>0.14</td>
</tr>
<tr>
<td>Formative COSA</td>
<td>3.61 (0.76)</td>
<td></td>
<td>3.32 (0.89)</td>
<td></td>
<td>4.74***</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Note 1: * p < 0.05, *** p < 0.001  
Note 2: Effect size (Cohen’s d): small (0.2 to 0.5), medium (0.5 to 0.8) and large (0.8 and higher)

Table 6. Gender differences between the Singaporean male and female eighth graders

<table>
<thead>
<tr>
<th>Factor</th>
<th>Gender (n)</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface COSA</td>
<td>Male (180)</td>
<td>3.21</td>
<td>0.67</td>
<td>0.79</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Female (153)</td>
<td>3.15</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summative COSA</td>
<td>Male (180)</td>
<td>3.61</td>
<td>0.71</td>
<td>-0.33</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>Female (153)</td>
<td>3.63</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formative COSA</td>
<td>Male (180)</td>
<td>3.65</td>
<td>0.75</td>
<td>1.01</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Female (153)</td>
<td>3.56</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7. Gender differences between the Taiwanese male and female eighth graders

<table>
<thead>
<tr>
<th>Factor</th>
<th>Gender (n)</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface COSA</td>
<td>Male (217)</td>
<td>3.02</td>
<td>0.90</td>
<td>-0.63</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>Female (207)</td>
<td>3.07</td>
<td>0.78</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Summative COSA</td>
<td>Male (217)</td>
<td>3.52</td>
<td>0.87</td>
<td>0.21</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Female (207)</td>
<td>3.50</td>
<td>0.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formative COSA</td>
<td>Male (217)</td>
<td>3.40</td>
<td>0.89</td>
<td>1.95</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Female (207)</td>
<td>3.23</td>
<td>0.90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* Effect size (Cohen’s $d$): small (0.2 to 0.5), medium (0.5 to 0.8) and large (0.8 and higher)