Teachers' Assessment Practices in Science, Technology, Engineering and Mathematics (STEM) Related Subjects

Leong Kwan Eu', Chew Cheng Meng2 and Noraini Idris3

1Faculty of Education, University of Malaya
2School of Educational Studies, Universiti Sains Malaysia
3Office of Deputy Vice Chancellor (Research & Innovation), Universiti Pendidikan Sultan Idris

*rkleong@um.edu.my

Abstract

STEM education is the backbone of technological advancement in any country. Countries with a strong STEM workforce are able to sustain the economy by creating innovative and sustainable products. Malaysia is facing a major problem because of the decrease in the enrollment of science students at the secondary level. The main objective of this study was to elicit secondary school teachers' perceptions of assessment practices in STEM related subjects. The researchers employed a cross-sectional survey research design and the sample of this study consisted of 196 teachers who were teaching Form 4 (Grade 9) classes at the secondary level. The data were collected using a five-point Likert scale questionnaire containing 8 items. The questionnaire was developed based on the STEM Education Quality Framework. The results of this study indicated that secondary teachers generally had positive overall perceptions of assessment practices in STEM related subjects. The results also indicated that there was no statistically significant difference in overall perceptions of assessments in STEM related subjects in terms of gender, years of teaching and school category.

Keywords: assessment, secondary teachers, perception, STEM related subjects

Abstrak:

Pendidikan STEM merupakan tulang belakang kepada kemajuan teknologi dalam masyarakat negara. Negara-negara dengan tenaga kerja STEM yang kuat mampu untuk mengelola ekonomi dengan menghasilkan produk-produk inovatif dan mampu. Malaysia sedang menghadapi masalah besar kerana penurunan dalam enrolmen pelajar sains di peringkat menengah. Objektif utama kajian ini adalah untuk mencangkiri persepsi guru-guru sekolah menengah tentang analisis penilaian dalam mata pelajaran yang berkaitan STEM. Para penyelidik menggunakan kuesioner secara keraton rentan dan sampel kajian ini terdiri daripada 196 orang guru yang mengajar Tingkatan 4 (Gred 9) kelas di peringkat menengah. Dapatnya dikumpulkan menggunakan skala Likert lima mata yang mengandungi 8 item. Skal selidik ini telah dibina berdasarkan STEM Rangka Kerja Kaedah Pendidikan. Keputusan kajian ini menunjukkan bahawa guru-guru menengah secara umumnya mampu menilai persepsi yang positif terhadap analisis penilaian dalam mata pelajaran yang berkaitan STEM. Keputusan juga menunjukkan bahawa tidak terdapat perbezaan statistik yang signifikan dalam persepsi keseluruhan penilaian dalam mata pelajaran yang berkaitan STEM dari segi jantina, tahun mengajar dan kategori sekolah.

Kata kunci: pentaksiran, guru sekolah menengah, persepsi, subjek berkaitan STEM
The problem of Malaysian students lagging in the mathematics and science literacy was also highlighted in the 2009+ Program for International Student Assessment (PISA) study. In the international assessment participated by 74 countries, Malaysia ranked in the bottom third in mathematics and science literacy. The study also reported that 60% of Malaysian students failed to meet the minimum benchmark in mathematics and 43% in science (OECD, 2010). Using the scale provided, it can be inferred that 60% of Malaysian students in mathematics and 43% students in science do not meet the baseline proficiency required for students to participate effectively and productively in life. The average score of Malaysian students in mathematics was 422 compared to the international average of 463 and the score was at least 100 points lower than for countries such as Singapore, South Korea, and Japan. This indicated that Malaysian 15 year old students had three years less schooling than their peers of the same age in these countries (OECD, 2010).

**Importance of assessments**

These issues pose a serious challenge to the country and something should be done to improve the quality of STEM education. Teachers’ role in addressing this issue of providing quality assessment in the STEM fields cannot be overemphasized. To identify the current assessment methods in STEM related subjects at the secondary level would be a key step that could be undertaken to address the challenges in STEM education. Assessment is a key component in the education process (OECD, 2005). These major issues point to a serious challenge for the government to improve STEM education and thus improve the enrollment of students into the STEM fields at the secondary education level.

In any educational system, assessment is an important component of the learning process. It plays an important role in improving the educational system, curriculum, quality of teaching and student learning (Aydeniz et al, 2012; National Research Council [NRC], 2001, 2005). Assessment can be done through the formative or summative methods. The formative assessment (also known as assessment for learning) usually measures students' understanding and knowledge of the subject during classroom instruction. Research has shown that formative assessment improves the quality of instruction and supports student learning (Aydeniz et al, 2012; Black & William, 1998; Duschl & Gitomer, 1997; Klassen, 2006; McMillan, 2001). In particular, Black et al (1998) reviewing 250 empirical articles found that teachers who applied the classroom formative assessment practices could enhance students' learning, in particular for low-achieving students. The study also discovered that large effect sizes exist for good formative assessment practices.

Examination and tests that measure students' knowledge after a period of time is known as summative assessment. Summative assessments help monitor the effectiveness of a curriculum and evaluate the effectiveness and quality of instruction (Aydeniz et al, 2012; NRC, 2001, 2003) and enhance the efficiency of the school system (Aydeniz et al, 2012; Brickhouse, 2006).

In most cases, scores collected from assessments were used as an indicator for students' achievement in school. However, there are suggestions to widen the usage of assessment beyond the usual way (MacMillan, 2004). Assessment should emphasize the learning of STEM related subjects for understanding that is useful both for the teachers and students. This would be very beneficial for teachers as they decide their pedagogical strategies to enhance student learning. According to the National Council of Teachers of Mathematics (2000), assessment should not only be administrated to students but also for teachers to guide students' learning and understanding of the STEM related subjects.

Researchers and educators are generally interested in finding out learning practices that would be useful in improving student achievement. Classroom assessment could be a catalyst in reforming practice that support student learning. When assessment is designed for a specific learner, student learning can be facilitated in the classroom inst

**Teachers' perception**

Ayde science teachers' accountability in the learning of STEM is generally encouraged to do more, at terms of improving the tests and reform documents.

Ong (2006) be the most of the 'frequent', among the different true/false and mean. She also examinations the most popular and then reporting different forms and coursework assessments.

Ong (2006) assessment practice. English, English, science and math. In addition, Student Mathematics is a means compared to the Physical Education lot of easy questions.

In the assessment practice reported that if by the public, without the is evaluation, peer, external and more frequently gave the least feedback direct feedback on a teacher. In all the students when
a specific learning outcome, it contributes to effective classroom instruction and enhances student learning (Gronlund, 2006). In fact, placing assessment as an integral component in classroom instruction would improve student learning (NCTM, 2000).

**Teachers' perceptions of assessments**

Aydeniz et al (2012) found that from American high school and middle school science teachers' perspectives, standardized testing practices used for the purposes of accountability had a significant influence on the quality of curriculum implemented as well as the science teachers' instructional and assessment practices in ways that are counter to the learning goals promoted by science education reformists. In other words, the findings suggested that the teachers perceived that although the standardized assessments encouraged them to use effective instructional and assessment methods that are promising in terms of improving students' achievement scores on the standardized tests, they believed that the tests did not measure the type of learning goals advocated by the science education reform documents.

Ong (2010) found that the multiple-choice objective questions in Malaysia seem to be the most common choice among the school teachers with 52.7% of them using it 'frequently', and another 22.9% using them 'very frequently'. Thus it has the highest mean among the different types of written tests such as essay, fill in the blank, short response, true/false and matching questions as the least used by the school teachers with the lowest mean. She suggested that the results might be attributed to the influence of public examinations that rarely use such types of written tests. Furthermore, she found that the most popular performance assessment used by school teachers is homework with 50.5% of them reporting "often" and 23.4% reporting "very often". It has the highest mean among the different forms of performance assessments such as project, practical, portfolio, homework and coursework. Portfolio has the lowest mean among the different forms of performance assessments.

Ong (2010) also found that teachers of different subject areas do not have different assessment practices, except for written tests. Language teachers (teachers who teach English, Malay, Chinese or Tamil language) used more written tests to assess students while science and mathematics teachers seem to use less of written tests according to the analysis. In addition, Science (General Science, Physics, Chemistry or Biology) and Mathematics (Mathematics or Additional Mathematics) teachers were found to be using less fill-in-the-blanks questions, true/false questions and matching questions to assess their students as compared to teachers of other subject areas (Geography, History, Accounts, Economics, Physical Education, Arts or Integrated Living Skills). Moreover, language teachers used a lot of essay questions to assess their students compared to other teachers.

In the case of teachers' viewpoints on assessment, Ong (2010) found that teachers' assessment practices are clearly influenced by the public examinations as 46.8% of them reported that they were "often" influenced and another 15.5% were "very often" influenced by the public examinations. Only 2.7% of the teachers selected their assessment methods without the influence of the public examinations. In terms of providing feedback of evaluation, primary school teachers reported giving more feedback compared with secondary and pre-university teachers. The primary school teachers were giving feedback more frequently to parents compared with teachers from other levels. Pre-university teachers gave the least feedback of evaluation results to parents but they were giving a lot of feedback directly to their students. Primary school teachers were also giving a lot of feedback on assessment results to school administrators compared with pre-university teachers. In addition, primary school teachers tend to use other strategies to evaluate students whereas pre-university teachers seldomly use "other strategies" to evaluate
students. Primary school teachers were found to use more of observation technique compared with other teachers. Pre-university teachers rarely used observation technique to assess students. Primary school teachers also used more of oral questioning techniques and interviews to evaluate students compared with other teachers.

Chien, Wu and Hsu (2014) explored teachers’ beliefs about technology-based assessments (TBAs) and investigated the possible interplay between their beliefs and their usage of TBAs in classrooms. They found that although 85% of the teachers perceived TBAs as useful tools and identified a variety of usefulness, nearly 40% indicated the difficulties in using TBAs and their beliefs of easy usage were mainly negative. In addition, the teachers’ controlled beliefs about TBA focused on the social and external components such as time, supporting personnel, and infrastructure rather than the personal factors. The normative beliefs of the teachers tended to view school policies and parents’ opinions as constraints, though they also realized the benefits of using TBAs for learning.

Ogan-Bekirgozlu (2009) examined pre-service physics teachers’ attitudes towards assessment and found two factors affecting their attitudes. Firstly, the difficulties that pre-service teachers experienced relating to assessment and secondly the teachers' self-efficacy regarding their ability to assess. These were two types of difficulties, namely internal and external. Internal difficulties represented the difficulties that pre-service teachers encountered during preparation and evaluation of assessment methods and depended on their assessment skills and subject matter knowledge. External difficulties depended on the external factors like school policy and facilities that might affect their implementation of classroom assessment.

Wang, Wang, and Huang (2008) proposed the "Practicing, Reflecting and Revising with WATA system: (P2R-WATA) Assessment Literacy Development Model" for improving pre-service teacher assessment literacy. The advantages of WATA system were personalized learning resources and opportunities for pre-service teachers to assemble tests and administer them to students on-line, facilitated performance of test analysis and item analysis, and enabled them to review statistical information from the test and item analyses to revise test items. They found that the P2R-WATA Assessment Literacy Development Model helped the pre-service teachers to improve their assessment knowledge and assessment perspectives.

Statement of the Problem

However, not much is known about secondary school teachers' perceptions of assessments in STEM related subjects in Malaysia. As such, it would be very pertinent to elicit teachers' perceptions of their classroom assessments at the secondary level to understand and further enhance assessment practices in STEM related subjects.

Does gender difference or gender bias occur in teachers' assessment practice? Gender difference could also happen with the style and type of assessments teachers used. A study on teaching mathematics with technology found different general tendencies among male and female teachers (Trigueros & Lozano, 2011). Male teachers concentrated on the learning of the mathematical concepts and procedures and how the classroom assessment could be implemented while female teachers have a more holistic approach to student learning in mathematics that includes emotional and cognitive aspects. However, only limited studies focus on gender differences in teachers' perceptions of assessments in STEM education in the literature. Thus, it is also important to explore gender differences in teachers' perceptions of the current assessments in STEM related subjects at the secondary school level so that help can be provided to enhance student learning in STEM related subjects.
Teachers’ experience is usually reflected through the number of years they teach in schools. Many believe that experienced teachers generally have better classroom management than novice teachers. It would be interesting to explore the assessment practices of STEM related subjects by looking at the teacher’s number of years of teaching in the classroom. According to Mulloin et al. (2012), the older the teachers the more mature they perceived the student’s education competence. In addition, the study discovered that teachers’ ratings on students’ attitude and education competence varied systematically by their gender and age. This bias may have an effect on assessment practices that measure students’ performance in STEM related subjects.

Objectives of the Study

The main objective of this study was to elicit secondary school teachers’ perceptions of assessments in STEM related subjects. The secondary objective focused on whether there were any significant differences in their perceptions of assessments in STEM related subjects in terms of gender, years of teaching and school category. Specifically, this study aimed at answering the following research questions:

1) What were secondary school teachers’ perceptions of assessment practices in STEM related subjects?
2) Was there any significant difference in the teachers’ perceptions of assessment practices in STEM related subjects in terms of gender?
3) Was there any significant difference in the teachers’ perceptions of school assessment practices in STEM related subjects in terms of years of teaching?
4) Was there any significant difference in the teachers’ perceptions of assessment practices in STEM related subjects in terms of school category?

Methodology

This was a quantitative study. The data were gathered through a self-reported questionnaire. This study utilized a cross-sectional survey research design as it could provide relevant information from the population of secondary school teachers’ perceptions of assessments in STEM related subjects (Gay, Miles, & Airasian, 2011).

Sample

The sample of this study consisted of 196 teachers who teach Form 4 (Grade 9) classes at the secondary school. These teachers comprised of 30 males and 166 females who were randomly selected from three categories of schools. The teachers taught STEM related subjects in the Malaysian secondary curriculum such as Physics, Chemistry, Biology, Science, Mathematics, Additional Mathematics, Information Technology and Engineering Drawing. The three categories of schools involved were classified as daily normal (ND), normal high performing (NHP) and full boarding high performance (FBHP). Schools using the term ‘high performing’ indicated that the students achieved outstanding academic and co-curricular activities and had fulfilled the criteria set by the education authorities (MOE, 2012). However, only 138 teachers completed the survey questionnaire of this study. Table 1 summarizes the demographics of the samples for this study.
Table 1. Number of teachers by school category and gender

<table>
<thead>
<tr>
<th>School Category</th>
<th>Gender</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Daily</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Performing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Boarding</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>Performing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily Normal</td>
<td>39</td>
<td>102</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>146</td>
</tr>
</tbody>
</table>

Instrument

The instrument in this study was a survey questionnaire titled.... The first section of the questionnaire asked for demographic details while the second section featured eight items related to rating of STEM assessment. Items of this instrument were developed by the researchers using the STEM Education Quality framework as a reference (STEM, 2011). The framework was created by the Dayton (Ohio) Regional STEM Center with the assistance of Dr. Rowley from the University of Dayton. All the instrument items used the five point Likert type scale options from strongly disagree, disagree, neutral, agree and strongly agree; the ratings of each item were given weighted values from 1 (strongly disagree) to 5 (strongly agree).

To ensure instrument reliability, the Cronbach’s alpha was determined. In this study, the Cronbach’s alpha value obtained was .84, indicating that all the items in this instrument have a high degree of internal consistency. Five experienced lecturers in assessment procedures and instrumentation validated the items on the instrument. This ensured that the content validity of the instrument covers all the main concepts in assessment. Thus, this ascertains the usability of the instrument to measure the expected teachers’ perceptions of STEM assessment appropriately.

Results

The instrument comprising of eight items was subjected to the exploratory factor analysis. The analysis generated two factors with an eigenvalue of greater than one which explained 68% of the total variance of the data. This means that the eight items belong to two factors. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy test is more than .6 indicating that the variables are highly factorizable. The result of the Bartlett’s Test of Sphericity is significant (p < .05). This indicates that the variables are related and therefore the factors exist. Using the principal component analysis with varimax rotation method, items 4, 6, 7, 8 were shown to belong to Factor 1 as the values are larger than .3. Items 1, 2, 3, 5 have values larger than .3 and this showed that these four items are related to Factor 2. Factor 1 refers to items on acquiring knowledge while factor 2 consists of items on acquiring skills.
Several descriptive statistics were calculated such as mean and standard deviation. With these values, the researchers were able to report about the teachers' perceptions of assessment practices in STEM-related subjects. Based on the responses in the questionnaire, the overall perception of the teachers was generally high (overall mean = 3.84). The mean of the overall perception of the teachers showed that generally teachers were moderately satisfied with the current assessment of STEM-related subjects. The mean values of the items under the question are as follows:

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.84</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3.70</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.77</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.68</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.53</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3.42</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.36</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3.27</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Descriptive statistics of school assessment practices in STEM-related subjects.

Table 2: Paired component matrix.

<table>
<thead>
<tr>
<th>Component 1</th>
<th>Component 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 3</td>
<td>0.43</td>
</tr>
<tr>
<td>Item 4</td>
<td>0.45</td>
</tr>
<tr>
<td>Item 5</td>
<td>0.46</td>
</tr>
<tr>
<td>Item 6</td>
<td>0.44</td>
</tr>
<tr>
<td>Item 7</td>
<td>0.47</td>
</tr>
<tr>
<td>Item 8</td>
<td>0.43</td>
</tr>
</tbody>
</table>
(Item 1, $M=3.87$; item 2, $M=3.96$; item 3, $M=3.87$; item 5, $M=4.00$) were higher than the mean values of the items under the acquiring skills factor (item 4, $M=3.84$; item 6, $M=3.87$; item 7, $M=3.47$; item 8, $M=3.70$). All the items in this questionnaire reported a mean value of above 3.5 suggesting that generally teachers have moderate perceptions of the assessment practices in STEM related subjects. Item 5 recorded the highest mean value ($M=4.00$) indicating that teachers are very satisfied with the current STEM assessment on evaluating the levels of student thinking skills. In contrast, the lowest mean value ($M=3.47$) in item 7 regarding the assessment of STEM related subjects based on students' understanding and skills in engineering. This happened because the curriculum does not explicitly include understanding and engineering skills in the syllabus.

(2) Difference in teachers' perceptions of assessments in STEM related subjects

As shown in Table 4, the perceptions of assessment practices in STEM were gathered from samples of 45 males and 123 females, with a female sample mean of 3.81 ($SD=0.74$) and a male sample mean of 3.84 ($SD=0.73$). To determine whether there is any significant difference in the teachers' perceptions of assessment practices in STEM related subjects in terms of gender, an independent samples t-test was conducted. Table 4 shows the results of the independent samples t-test for each item in the questionnaire. The mean score of the male teachers was higher for items under the acquiring skills factor (item 4, $M=4.00$; item 6, $M=4.00$; item 7, $M=3.54$) compared to those of the female teachers. Conversely, female teachers recorded a higher mean for items of acquiring knowledge factors (item 1, $M=3.89$; item 3, $M=3.85$; item 5, $M=4.02$) than the male teachers.

Table 4. Results of the Independent-Samples T-Tests

<table>
<thead>
<tr>
<th>Item</th>
<th>Male ($N=45$)</th>
<th>Female ($N=143$)</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment practices in STEM related subjects are</td>
<td>$M$</td>
<td>SD</td>
<td>$M$</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>1 based on students learning experience</td>
<td>3.76</td>
<td>.88</td>
<td>3.89</td>
<td>.79</td>
<td>-8.63</td>
</tr>
<tr>
<td>2 designed to help students relate the knowledge and skills of the subject.</td>
<td>4.02</td>
<td>.79</td>
<td>3.94</td>
<td>.68</td>
<td>.665</td>
</tr>
<tr>
<td>3 based on the integrated manner among the subjects</td>
<td>3.89</td>
<td>.73</td>
<td>3.85</td>
<td>.58</td>
<td>.397</td>
</tr>
<tr>
<td>4 takes into account students ability in solving problems and project management skills</td>
<td>4.00</td>
<td>.66</td>
<td>3.79</td>
<td>.79</td>
<td>1.514</td>
</tr>
</tbody>
</table>

*p = .01; * * * = .001

items is more significant between male and female teachers.
<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Mean</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Takes into account the level of students thinking skills including the 'why' and 'how' questions</td>
<td>3.92</td>
<td>.88</td>
<td>4.02</td>
<td>.56</td>
<td>-.859</td>
<td>169</td>
<td>.391</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Takes into account the skill of students working together in groups</td>
<td>4.00</td>
<td>.41</td>
<td>3.83</td>
<td>.78</td>
<td>1.246</td>
<td>167</td>
<td>.214</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Based on the students understanding and skills in engineering design.</td>
<td>3.54</td>
<td>.84</td>
<td>3.46</td>
<td>.96</td>
<td>.473</td>
<td>153</td>
<td>.637</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Takes into account the usage of few available resources from numerous technologies</td>
<td>3.65</td>
<td>.66</td>
<td>3.70</td>
<td>.79</td>
<td>-.355</td>
<td>155</td>
<td>.723</td>
<td></td>
</tr>
</tbody>
</table>

Overall perceptions  
3.84  
.73  
3.81  
.74  
-69  
164  
.46  

*Significant at p < 0.05

The independent t-test indicated that the differences in the mean scores for all the items between the male and female teachers, however, were not significant ($r = .69, df = 164, p = .46$) indicating that no differences exist in the perceptions of assessment in STEM related subjects between teacher's gender.

Furthermore, the mean score of the male teachers' overall perceptions were slightly higher than that for the female teachers. But the difference in the mean was not significant between the male and female teachers at the $p$-value of .05, suggesting that the overall perceptions of assessment in STEM related subjects were not significantly different and quite similar in terms of teachers' gender.
(3) Difference in teachers’ perceptions of assessments in STEM related subjects in terms of years of teaching

Table 5 displays the results of the one-way ANOVA for all the items in the questionnaire. A one-way between subjects ANOVA was conducted to compare the teachers’ perception of assessment practices in terms of years of teaching experience in less than five years, five to ten years and more than 10 years.

<table>
<thead>
<tr>
<th>Item</th>
<th>&lt;5 years</th>
<th>5 to 10 years</th>
<th>&gt;10 years</th>
<th>F</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 85)</td>
<td>(N = 62)</td>
<td>(N = 41)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Results of the one-way ANOVA between subjects years of teaching experience

Assessment practices in STEM related subjects are

1. based on students' learning experience
   - Mean: 4.03
   - SD: .78
   - T-test: 1.01
   - Sig: .37

2. designed to help students relate the knowledge and skills of the subject
   - Mean: 4.03
   - SD: .74
   - T-test: 2.87
   - Sig: .06

3. based on the integrated manner among the subjects
   - Mean: 3.97
   - SD: .54
   - T-test: 2.55
   - Sig: .08

4. takes into account students' ability in solving problems and project management skills
   - Mean: 3.94
   - SD: .56
   - T-test: 2.85
   - Sig: .05

5. takes into account the level of students thinking skills including the 'why' and 'how' questions
   - Mean: 4.13
   - SD: .66
   - T-test: 4.02
   - Sig: .16

6. takes into account the skill of students working together in groups
   - Mean: 3.91
   - SD: .69
   - T-test: 3.89
   - Sig: .58

Overall perception value for SD = .65 mean a slightly at the statistice terms of any sig related to the perceptions. Perform 6 please overall.

In comparison, the recorded significant was 2 where there is three cases among the mean there were between the means. For item 2 very mean scores.
<table>
<thead>
<tr>
<th></th>
<th>3.77</th>
<th>.67</th>
<th>3.40</th>
<th>1.07</th>
<th>3.38</th>
<th>.90</th>
<th>2.27</th>
<th>(2,182)</th>
<th>.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 based on the students understanding and skills in engineering design.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 takes into account the usage of few available resources from numerous technologies</td>
<td>3.77</td>
<td>.56</td>
<td>3.62</td>
<td>.84</td>
<td>3.69</td>
<td>.81</td>
<td>.35</td>
<td>(2,182)</td>
<td>.71</td>
</tr>
<tr>
<td>Overall perceptions</td>
<td>3.94</td>
<td>.65</td>
<td>3.71</td>
<td>.80</td>
<td>3.84</td>
<td>.72</td>
<td>1.51</td>
<td>(2,182)</td>
<td>.32</td>
</tr>
</tbody>
</table>

The group with less than 5 years of teaching experience recorded the highest mean value for the overall perceptions of assessment practices in STEM related subjects ($\bar{X}=3.94, SD=.65$) followed by the group with more than 10 years ($\bar{X}=3.84, SD=.72$) and the lowest mean score was the 5 to 10 years group ($\bar{X}=3.71, SD=.80$). There was not a significant difference in teachers’ perception of assessment practices on years of teaching experience at the $p<.05$ level for the three conditions ($F(2,182) =1.51, p =.22$). Thus, there was no statistical significance in the overall perceptions of assessments in STEM related subjects in terms of teaching experiences.

(4) Difference in teachers’ perceptions of assessments in STEM related subjects in terms of school category

A one-way analysis of variance (ANOVA) was done to determine whether there is any significant difference in the teachers’ perceptions of assessment practices in STEM related subjects in terms of school category. The three school categories were Daily High Performing (DHF), Full Boarding High Performing (FBHP) and Daily Normal (DM). Table 6 presents the results of the one-way ANOVA for each item in the questionnaire and the overall perception of assessments in STEM related subjects.

The DHF schools obtained the highest mean value for Items 1, 2, 4, 5, 6, 7 and 8. In comparison, FBHP schools had the lowest mean scores for Items 1, 2 and 3. DM schools recorded the highest mean score for Item 3. However, the differences in mean score were not significant among the three school categories for most of the items in the questionnaire. Item 2 was statistically significant at the .05 level ($F(2,182) =4.11, p =.02$). This showed that there was a difference in the perception of assessments in STEM related subjects among the three categories of schools for Item 2 under the acquiring knowledge factor.

A Tukey post hoc test was conducted to understand whether the mean differences among the three school categories were statistically significant as shown in Table 6. Only the mean value of Item 2 showed statistical significance at the $p$-value of .05, indicating that there were differences in the perceptions of assessments in STEM regarding the relationship between the knowledge and skills of the subject among the DHF, FBHP and DM schools. For Item 2, the mean difference was not statistically significant (i) between the DHP and DN teachers and (ii) between FBHP and DN teachers. However, the difference in mean score for Item 2 was statistically significant between DHP and FBNP with the DHP achieving a better mean score.
The mean score of the FBHP teachers was the lowest among the three categories of school for the overall perception of assessment practices in STEM related subjects. However, the mean score difference for the overall perception was not statistically significant\[F(2,182) = 1.51, p = .22\], suggesting that the perceptions of assessment practices in STEM related subjects were not different among the three school categories.

### Table 6. Results of the one-way ANOVA between school categories

<table>
<thead>
<tr>
<th>Item</th>
<th>Daily Performing (N = 75)</th>
<th>Full Boarding Performing (N = 52)</th>
<th>Daily Normal (N = 43)</th>
<th>(F)</th>
<th>(df)</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M)</td>
<td>(SD)</td>
<td>(M)</td>
<td>(SD)</td>
<td>(M)</td>
<td>(SD)</td>
<td></td>
</tr>
<tr>
<td>Assessment practices in STEM related subjects are</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 based on students learning experience</td>
<td>4.05</td>
<td>.52</td>
<td>3.79</td>
<td>.69</td>
<td>.84</td>
<td>.69 (2,182)</td>
</tr>
<tr>
<td>2 designed to help students relate the knowledge and skills of the subject</td>
<td>4.32</td>
<td>.58</td>
<td>3.76</td>
<td>.56</td>
<td>.71</td>
<td>4.11 (2,182)</td>
</tr>
<tr>
<td>3 based on the integrated manner among the subjects</td>
<td>3.79</td>
<td>.63</td>
<td>3.67</td>
<td>.69</td>
<td>.57</td>
<td>2.71 (2,182)</td>
</tr>
<tr>
<td>4 takes into account students ability in solving problems and project management skills</td>
<td>4.05</td>
<td>.62</td>
<td>4.03</td>
<td>.59</td>
<td>.79</td>
<td>2.49 (2,182)</td>
</tr>
<tr>
<td>5 takes into account the level of students thinking skills including the 'why' and 'how' questions</td>
<td>4.26</td>
<td>.65</td>
<td>4.00</td>
<td>.71</td>
<td>.62</td>
<td>1.84 (2,182)</td>
</tr>
<tr>
<td>6 takes into account the skill of students working together in groups</td>
<td>4.05</td>
<td>.62</td>
<td>4.00</td>
<td>.61</td>
<td>.74</td>
<td>1.66 (2,182)</td>
</tr>
<tr>
<td>7 Based on the students understanding and skills in engineering design.</td>
<td>3.72</td>
<td>.75</td>
<td>3.72</td>
<td>.73</td>
<td>.96</td>
<td>2.61 (2,182)</td>
</tr>
</tbody>
</table>

**Discussion**

Table 6 shows the results of the one-way ANOVA between school categories for assessment practices in STEM related subjects. The table indicates that there are no significant differences in the perceptions of assessment practices among the three school categories. This suggests that the perception of STEM related subjects is not affected by the type of school. The results are consistent with the overall perception indicating similar levels of assessment practices across all school types.

In the process of improving assessment practices in STEM related subjects, ongoing professional development and collaborative design among teachers can be implemented. This, along with the design of effective assessment practices, can enhance students' understanding and skills in engineering design. The results support the need for continuous improvement in STEM education to better prepare students for the challenges of the 21st century.


<table>
<thead>
<tr>
<th>Item</th>
<th>School Category (I)</th>
<th>School Category (II)</th>
<th>Mean Difference (I-J)</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment practices in STEM related subjects are designed to help students relate the knowledge and skills of the subject.</td>
<td>DHP</td>
<td>FBHP</td>
<td>.56</td>
<td>.01**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DN</td>
<td>.36</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FBHP</td>
<td>.56</td>
<td>.01**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DN</td>
<td>.20</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DHP</td>
<td>.36</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FBHP</td>
<td>.20</td>
<td>.28</td>
</tr>
</tbody>
</table>

Discussion

The results of this study showed that generally teachers were moderately satisfied with the current assessment of STEM related subjects using the mean score of the overall perception of the teachers as an indicator. Item 5 recorded the highest mean value of 4.00 indicating that teachers are satisfied with the current STEM assessment in evaluating the levels of student thinking skills. In contrast, the lowest mean value was 3.47 (Item 7) regarding the assessments of STEM related subjects based on students' understanding and skills in engineering, implying that teachers were less satisfied about this assessment section.

In the STEM Education Quality Framework (2011), engineering design refers to the process involved in building a product such as brainstorming, creating, researching, improving and testing the product. However, this pertinent process of engineering design has not been given sufficient emphasis in assessment and instruction of STEM related subjects at the secondary level in Malaysia. This might be due to the missing engineering design component in the STEM related subjects at the secondary level. Thus, the findings indicated that engineering design component should be increased in the syllabus of the STEM related subjects and inculcated in the teaching and learning process of those subjects. When assessment is designed properly and used appropriately, this improves student learning by having effective instruction (Gronlund, 2006; NRC, 2001, 2005).

The mean values of the items under the acquiring knowledge factor were higher than the mean values of the items under the acquiring skills factor. This is due to importance placed by teachers in acquiring knowledge that influences their self-efficacy and assessment practices. The results of this study is consistent with the findings by Ogan-Bekiroglu (2009) on attitude of pre-service teachers on assessment.

Similarly, the differences in the mean scores for each of the eight items between the male and female teachers were not significant. This shows that no differences exist in the perceptions of assessments in STEM related subjects between male and female teachers. In
addition, the overall perceptions of assessments in STEM related subjects were not statistically different and quite similar in terms of teachers' gender (the mean differences were not significant between the male and female teachers at the p-value of .05). The findings indicated that generally male and female teachers should be exposed to different assessment techniques in STEM related subjects to enhance student learning and achievement. These results on gender differences are not consistent with the findings by Trigoso et al (2011) which found that male and female teachers have different general tendencies in teaching mathematics. In general, the results differ from the results of TIMSS 2011 which showed that there were some gender differences in science and mathematics achievement and Freeman (2004)'s study which described a small significant difference between boys' and girls' mathematics scores in the National Assessment of Educational Progress.

This study also examined whether there is any significant difference in the teachers' perceptions of assessments in STEM related subjects in terms of years of teaching. The findings indicated that the differences in the mean score for the overall perceptions of assessments in STEM related subjects was not significant at the p-value of .05 for the three groups of teachers in terms of years of teaching. Thus, there was no statistical significance in the overall perceptions of school assessment in STEM related subjects in terms of teaching experience. This implies that experienced teachers and novice teachers have similarities in their assessment practices in STEM related subjects that is of concern since Malaysia fared poorly in the TIMSS assessment results. The findings of this section contradicts the results obtained by Mullola et al. (2012) which stated that experienced teachers are more matured and would be able to perceive students' academic competence. It would be useful for experienced or novice teachers to improve their assessment practices in STEM related subjects to enhance learning of those subjects.

A one-way ANOVA was employed to examine whether there is any significant difference in the teachers' perceptions of assessments in STEM related subjects in terms of school category. The three school categories were Daily High Performing (DHP), Full Boarding High Performing (FBHP) and Daily Normal (DN). The mean score difference for the overall perception in STEM assessment was not statistically significant suggesting that the overall perceptions on assessment in STEM were not different between the DHP, FBHP and DN schools. Only one item showed statistical significance for the difference in the mean indicating that there were differences in the perception of assessment in STEM regarding the relationship between the knowledge and skills of the subject among the DHP, FBHP and DN schools. The mean difference was not statistically significant (i) between the DHP and DN school teachers and (ii) between FBHP and DN teachers. However, the difference in mean score for item 2 was statistically significant between DHP and FBHP with the results favoring DHP schools. This result implies that FBHP and DN teachers should provide more assessment practices in STEM related subjects that are designed to help students relate the knowledge and skills of the subjects.
Conclusion

In conclusion, the findings of this study indicated that secondary teachers generally were moderately satisfied with the overall perceptions of assessments in STEM related subjects. This implies that secondary teachers should further improve their assessment practices in STEM related subjects in order to address the issues of the decreasing enrollment of science students at the secondary level and the lagging mathematics and science achievement in international assessment studies. As initiatives of STEM education have been gradually implemented in Malaysia and not much is known about teachers' perceptions of assessments in STEM education per se, more studies should be carried out to examine teachers' perceptions of assessments in STEM education to address the issues of the decreasing enrollment of science students at the secondary level and the lagging mathematics and science achievement in international assessment studies. Nonetheless, the researchers acknowledged that the small sample of this study limits us from making any generalizations from the findings obtained.
References


http://dx.doi.org/10.1787/9789264091450-en


