ABSTRACT

The prevalence of distress among medical students continues to rise. Studies have shown that stress-related to examinations – test anxiety (TA) – is the most frequently reported source of distress. Research on the relationship between TA and assessment modalities is thus critical for determining the potential ramifications of this problem. This study aimed to explore assessment modalities that aggravate TA among medical students in a Malaysian medical school. A cross-sectional study was conducted among medical students at the School of Medical Sciences (SMS) of Universiti Sains Malaysia (USM). Students rated TA for each assessment modality used in the school. Each modality was scored from 1 to 10, with 1 indicating no TA and 10 representing extreme TA. Forty-five students participated in the study. The group was divided almost equally in terms of sex. The assessment modalities that provoked the most TA were the objective structured clinical examination (OSCE), the short case, the short essay question (SEQ) and the long case, with mean scores of 7.9, 7.8, 7.7 and 7.7, respectively. The case write-up, the problem-based learning (PBL) assessment, the multiple true-false (MTF) questions and the Simplified Thematic Engagement of Professionalism Scale (STEPS) were the assessment methods that induced the least TA, with mean scores of 5.1, 5.0, 4.4 and 4.0, respectively.

This study found that the worst assessment modalities in terms of TA were the OSCE, the short and long cases, and the short essay question, while the case write-up, the PBL assessment, the MTF questions, and the STEPS induced the least TA. Most students reported that memorisation difficulties and facing examiners were the most common causes of TA. Remedial measures include examiner training on how to deal with examinees during assessments, evaluating the distribution of marks according to assessment modality and student training focused on study skills and exam preparation.

Keywords: Assessment, Formats, Modalities, Test anxiety, Medical students
INTRODUCTION

The prevalence of distress among medical students continues to rise. Studies have consistently raised concerns about the scope of the problem (1–2). Distress leads to burnout, depression, poor academic performance, poor clinical performance, impaired decision-making, poor peer interaction, interpersonal conflict, academic dishonesty and sleeping problems (3–4). Distress has also been linked to drug abuse, alcohol consumption and suicide (5–8). These negative consequences eventually affect the provision of optimal medical care to patients (9–10). While medical students have identified a variety of stressors, studies have shown that examinations are the most frequently reported causes of stress (3, 11–16).

Research into the factors that contribute to examination stress has identified a concept known as test anxiety (TA). TA refers to more than just a fear of examinations; it pertains to the physiological and psychological changes that occur when certain stress thresholds are exceeded (17). TA is defined as a set of phenomenological, physiological and behavioural responses associated with fear of negative consequences or failure on a test or other evaluative situation (18). In evaluative situations, test-anxious students have a low anxiety response threshold; they see tests/exams as personally threatening. As a result, they are more likely to perceive failure as a threat. This decreased their feelings of self-efficacy, self-derogatory cognitions and anticipatory failure attributions, as well as increased their emotional reactions and arousal levels (19). TA may manifest as organic symptoms (20) or as psychological impairments in concentration and working memory (21), both of which affect students’ academic performance (22–23). While numerous precipitating factors of TA have been identified, time constraints, the volume of learning material, falling behind on work and failing to complete work on time are the most frequently reported (13, 24).

Additionally, TA was found to be higher when the test was regarded to be difficult and/or had high stakes or consequences (25).

To aid in the improvement of medical students’ mental health and to mitigate the effects of TA, scholars have investigated the relationship between TA and assessment modalities (18–19, 26–27). This work is critical in determining the potential ramifications of the problem. By identifying assessment modalities that cause distress, educators can have solid data on which assessment modalities are strongly associated with TA and direct their efforts towards overcoming modifiable stressors of these assessment modalities, thereby alleviating TA and improving the mental health of medical students. In this vein, this study explores the assessment modalities that aggravate TA among medical students in a Malaysian medical school and identifies the reasons behind this phenomenon.

METHODS

Study Setting

We conducted a cross-sectional study among medical students at the School of Medical Sciences (SMS), Universiti Sains Malaysia (USM). The foundation of the medical curriculum is the model known as SPICES (student-oriented, problem-based, integrated, community-oriented, electives, self-learning, and systematic learning) (28). This is a five-year programme with two phases: preclinical (first and second years) and clinical (third, fourth and fifth years). Preclinical medical students gain basic and applied knowledge about the normal human being and early clinical exposure to common pathological conditions. During the clinical phase, students learn clinical sciences and skills in a real-world setting (29). The SMS uses the following assessment modalities: multiple true-false (MTF) questions, single best answer (SBA), short essay question (SEQ), scenario-based question (SBQ),
The survey was enhanced qualitatively by asking students to reflect on why they assigned a high score (eight or above) to any modalities. Before data collection, we validated the survey by a group of medical education experts (face validity) and administered to a group of medical students, who indicated that it was understandable.

Study Analysis

Excel as well as Statistical Package for Social Sciences (SPSS) were used to analyse data. The mean scores of each assessment modality were calculated and compared statistically. Because data were not normally distributed, non-parametric test were used. To compare means between two groups (males and females), Mann-Whitney test was used and Kruskal–Wallis test was used to compare means between three groups (academic levels). Graphical distribution of the mean scores was prepared. The results were sorted in ascending order. A comparison in term of gender and academic level were run.

Content analysis was used to analyse the content of students’ responses to the qualitative component of the survey and explore students’ perceptions of why they gave those assessment modalities higher scale scores (32). Students’ statements were coded with the key terms that they wrote. Categorising similar codes was done to create themes. The frequency of the emerging themes was tabulated. Unedited quotations of students’ responses were presented concerning each theme.

RESULTS

Of 50 students recruited, 45 participated in the study. Twenty-three participants were male (51%) and 22 were female (49%). Twenty-one students (47%) were from the fourth year (Table 1).
Among the different assessment modalities, the OSCE, the short case, the SEQ and the long case provoked the most TA, with mean scores of 7.9, 7.8, 7.7 and 7.7, respectively. The case write-up, the PBL assessment, the MTF questions and the STEPS induces the least TA, with mean scores of 5.1, 5.0, 4.4 and 4.0, respectively (Figure 1).

The analysis revealed no statistically significant difference in mean scores for assessment modalities between male and female students. Female students, on the other hand, scored higher on some assessment modalities, such as the SMR viva (Figure 2). According to Figure 3, all academic years (which use the same assessment modalities) perceived OSCE, SEQ and SBQ to be the most TA-provoking modalities. While both short and long cases exams were perceived as increasing TA in the third and fourth years. This analysis showed no statistical differences ($p$-value > 0.05).

**Table 1**: Participants’ demographic characteristics

<table>
<thead>
<tr>
<th>Participants</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>23</td>
<td>51</td>
</tr>
<tr>
<td>Female</td>
<td>22</td>
<td>49</td>
</tr>
<tr>
<td>Academic level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second year</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>Third year</td>
<td>16</td>
<td>36</td>
</tr>
<tr>
<td>Fourth year</td>
<td>21</td>
<td>47</td>
</tr>
</tbody>
</table>

**Figure 1**: Overall average anxiety scores of different assessment modalities.

**Figure 2**: Comparison of average anxiety scores of different assessment modalities based on gender. Mann-Whitney test was use and showed statistical insignificant ($p$-value > 0.05).
causes of stress were the weightage of marks in specific assessment modalities (e.g., the SEQ), the time limitation during an examination, and the requirement to think critically and apply knowledge to answer the assessment tasks. Table 2 shows the content analysis of participants’ open-ended responses, including relevant quotations.

Table 2: Content analysis of students’ responses to open-ended questions

<table>
<thead>
<tr>
<th>Theme</th>
<th>Frequency</th>
<th>%</th>
<th>Example of students’ quotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorisation</td>
<td>29</td>
<td>60</td>
<td>The SEQ needs a lot of memorisation for each detail.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>It requires a lot of memorising and takes up a lot of marks.</td>
</tr>
<tr>
<td>Facing examiners</td>
<td>21</td>
<td>44</td>
<td>The anxiety of facing the doctors makes me nervous, so I forget the answers, although I have revised them very well.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>It is more stressful when you have to meet the examiners.</td>
</tr>
<tr>
<td>More marks</td>
<td>9</td>
<td>19</td>
<td>It carries the highest mark.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Due to the high marks that are involved in it.</td>
</tr>
<tr>
<td>Time restriction</td>
<td>6</td>
<td>13</td>
<td>The long case is stressful because of the fear of not taking the whole history in time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The time allowed is not sufficient.</td>
</tr>
<tr>
<td>Deep thinking</td>
<td>6</td>
<td>13</td>
<td>The SBQ is stressful because it is hard to relate the symptoms to the disease.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>For the long and short cases, you need to be able to pick up points (signs and symptoms) to reach the correct diagnosis.</td>
</tr>
</tbody>
</table>
DISCUSSION

This study found that the worst assessment modalities in terms of TA were the OSCE, the short and long cases, and the SEQ. The case write-up, the PBL assessment, the MTF questions and the STEPS induced the least TA. Most students reported memorisation difficulties and facing examiners as the most common causes of TA.

Numerous studies found that clinical examinations such as OSCE, short case and long cases increase TA (23, 26, 33–37). The current finding corroborates these studies (Figures 1, 2 and 3). While the statistical analyses were insignificant, these add another dimension to this study by demonstrating that clinical assessment modalities are rated higher than other modalities by all participants, regardless of gender or academic level. There are several factors contributing to TA during clinical examinations. These include the presence of examiners, the attending patients and sometimes the anxiety of being recorded so that others may observe you at a later stage (38). Zeidner (18) speculated that evaluation in a social context aggravates TA. Direct or indirect observation increases anxiety levels and interferes with working memory, leading to difficulty in memorisation and consequently poor performance during examination (39). The content analysis of the current finding supports these speculations. Students frequently report having difficulty with memorisation during the clinical examination (Table 2). This situation is worsened when examiners are more stringent and have higher expectations of examinees (40). To address these issues, academic leaders in medical schools should facilitate comprehensive clinical examination training for both examiners and examinees. Several studies showed that training of examiners on how to interact with examinees and how to evaluate them objectively using a checklist or rubric minimise the subjectivity and prejudice and hence reduce TA (40–42). On the other hand, examinees should be instructed on how to perform during clinical examinations and given advice on how to maximise their performance while minimising their TA level (43). Additionally, they can be exposed to a mock clinical examination (e.g., OSCE) to familiarise them with the exam’s format (34).

The SEQ requires students to memorise factual knowledge and recall it within a limited time. For this reason, students considered this assessment modality as TA-inducing. Many studies have reported that the inability to recall facts during examinations are a significant factor of TA (37, 44–45). Moreover, the time constraint factor is also associated with TA (37). For this reason, effective learning skills training should be provided to students to manage TA arising from this type of format (46–48), including practising answering questions (formative assessment) (49). It is thus highly recommended that health professions schools should have student development programmes dealing with these skills. In their meta-analyses, Soares and Woods (50), and Huntley et al. (51) found that such programmes significantly reduce TA.

One of the reported factors that aggravate TA is the weightage of marks for specific assessment modalities. Although increasing marks for some assessment formats reflects their importance in professional life and signals to students the need to prepare more for these formats, this will create what is called “score gainer” students (52). Students concentrate their efforts on gaining marks and improving their grades. As a result, students develop into superficial learners with limited capacity for critical thinking and reasoning. Concerning this problem, several studies (53–56) discussed using a pass-fail grading system rather than a numerical or letter grade to decrease TA. They concluded that shift to pass-fail grading is associated with improve mental health (53–55) and learning process (56). While this approach will shift students’
mindsets from a grade orientation to achievement behaviour (56–57), it will be ineffective for certain certifications that are based on ranking. Another solution is to distribute marks fairly across assessment modalities and to communicate this distribution to students. This will improve students’ perceptions of assessment and result in a reduction in TA (42).

The study findings establish a foundation for various changes to the assessment system, in the medical schools, aimed at reducing TA. These include assessor training programmes, evaluations of scoring and grading systems and self-development programmes for students focused on academic skills and learning strategies.

While the current study’s findings were similar with earlier studies, it has limitations. The limited sample size reduces generalisability. Another constraint is the restriction to a single school with a single health professions education speciality, medical students. Further studies are required that involve multiple centres and use mixed-mode designs to extensively explore assessment modalities and TA.

**CONCLUSION**

This study found that the assessment modalities that provoked the most TA were, in majority, the examiner-based modalities such as OSCE and clinical exams. Most students reported that memorisation difficulties and facing examiners were the most common causes of TA. The findings of this study are important as they identified assessment modalities associated with TA, so that medical school leaders could consider these modalities when working to reduce TA. Remedial measures include examiner training on how to deal with examinees during assessments, evaluating the distribution of marks according to assessment modality, and student training focused on study skills and exam preparation.

**ACKNOWLEDGEMENTS**

This research is part of a larger project supported by the Fundamental Research Grant Scheme (FRGS/1/2018/SSI09/USM/02/2), Ministry of Education, Malaysia.

**ETHICAL APPROVAL**

We obtained ethical approval from the USM Human Research Ethics Committee (JEPeM USM Code: USM/JEpE/18060286). Students who participated in the study received a small token of appreciation.
APPENDIX A

Anxiety Score of Types of Assessment

Gender: □ Male □ Female
Year of study: □ Year 1 □ Year 2 □ Year 3 □ Year 4 □ Year 5

Based on your experience, please rate the level of anxiety you feel performing these types of assessment:

<table>
<thead>
<tr>
<th>Type of assessment</th>
<th>No stress</th>
<th>Extremely stressful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple True False (MTF)</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>Single Best Answer (SBA)</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>Short Essay Questions (SEQ)</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>Scenario-Based Question (SBQ)</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>OSCE</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>Short case</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>Long case</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>Case write-up</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>Simplified Thematic Engagement of Professionalism Scale (STEPS)</td>
<td>0 1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
<tr>
<td>360-degree assessment</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>PBL assessment</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
<tr>
<td>Student Medical Record (SMR) viva</td>
<td>0 1 2 3 4</td>
<td>5 6 7 8 9 10</td>
</tr>
</tbody>
</table>

For those types of assessment rated 8–10, why do you think they are extremely stressful?

---------------------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------------------
REFERENCES


