

## Original Research Article

# Self-assessment: how do third year medical students rate their performance during problem-based learning?

Sehlule Vuma<sup>1\*</sup>, Bidyadhar Sa<sup>2</sup>

<sup>1</sup>Department of Para-clinical Sciences, The University of The West Indies, St Augustine, Trinidad and Tobago

<sup>2</sup>Centre for Medical Sciences Education, The University of the West Indies, St Augustine, Trinidad and Tobago

**Received:** 21 April 2017

**Accepted:** 18 May 2017

### \*Correspondence:

Dr. Sehlule Vuma,

E-mail: [Sehlule.Vuma@sta.uwi.edu](mailto:Sehlule.Vuma@sta.uwi.edu)

**Copyright:** © the author(s), publisher and licensee Medip Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

### ABSTRACT

**Background:** Problem-based learning (PBL) emphasizes students' self-directed learning. This requires students to monitor their own learning. When students enter medical school however, they do not seem to have self-monitoring skills and these need to be developed. Self-assessment exercises are useful in developing these skills. This paper emphasizes the importance of training in self-assessment during PBL. Objective was to describe third-year medical students' self-assessment of their performance during PBL and correlate their scores with tutors' scores, and their scores in a written progressive disclosure questions (PDQ) examination

**Methods:** Using the same rubric used for students' evaluation by tutors, students scored their activities during PBL sessions. Their scores were compared with the tutor's scores. Their scores for cognitive skills were further compared with their scores in a PDQ examination.

**Results:** There was no statistically significant difference between the tutor and students' scores. Low-achieving students (in the PDQ) scored themselves higher than high-achieving students. Self-assessment scores did not predict performance in the PDQ.

**Conclusions:** Perhaps the high-achievers are more critical of themselves which drives them to work harder. Low-achieving students could be awarding themselves scores they wish to obtain and not what they deserve. They may also not understand the assessment criteria. This paper emphasizes the importance of self-assessment exercises, and the need to guide students in learning to assess themselves accurately so they may be better able to monitor their learning.

**Keywords:** Cognitive, Problem-based learning, Self-assessment, Self-monitoring

### INTRODUCTION

Institutes of higher learning encourage students to take responsibility for their own learning, and continued personal and professional development.<sup>1</sup> Doctors not only need to obtain important knowledge and skills, but also be able to direct themselves in what they do learn.<sup>2</sup> The medical profession demands that doctors engage in self-regulation and self-education.<sup>3</sup> Medical curricula have been revised to address this.<sup>2</sup> Problem based learning (PBL) is an effective approach to this end. PBL was

developed to help students learn by problem-solving, which is what doctors do.<sup>3</sup> It helps to promote students' critical thinking, sharpen communication skills and general professionalism,<sup>4</sup> working in groups, and time management skills:<sup>5</sup> all useful beyond medical school,<sup>6</sup> and during residency training.<sup>7</sup> In PBL tutors assess students' intellectual and cognitive competence, however students need to be assessed on the "process of learning itself":<sup>1</sup> e.g. Valle s' instrument assesses independent study, group interaction, reasoning skills and active group participation.<sup>8</sup>

Since PBL emphasizes students' self-directed learning, this does require students to monitor their own learning, which is important for doctors' continued learning too.<sup>9-10</sup> Students need to be able to monitor both their cognitive and non-cognitive skills.<sup>10</sup> For effective self-monitoring, self-assessment is important in PBL.<sup>2</sup> When students enter medical school however, they do not have skills for self-monitoring and these need development,<sup>11</sup> certainly if self-assessment is such an important competency in PBL.<sup>2</sup> Useful self-assessment exercises are valuable in developing these self-monitoring skills.<sup>12</sup> Students need to self-monitor and not just rely on "external feedback" from teachers. Students benefit from learning to accurately evaluate their own work,<sup>11</sup> and be self-aware of their strengths and weaknesses, which are important in developing their profession.<sup>13</sup>

Boud's definition of self-assessment- "...the involvement of students in identifying standards and/or criteria to apply to their work, and making judgements about the extent to which they have met these criteria and standards....".<sup>12</sup> Self-assessment is a "comparative process" where the student compares their activities to set standards or previous performance.<sup>14</sup> By involving students in their own assessment process and learning<sup>1</sup>, it helps them to engage more in curricula.<sup>12</sup>

However self-monitoring requires students' to be self-motivated, take responsibility for their own learning: thus, requires autonomy in the learning process.<sup>3</sup> It is not only an assessment tool, but also does enhance learning.<sup>1</sup>

Different self-assessment strategies have been used. Online, physiology MCQ-quiz helped students to direct and focus their studies, strengthening their weaker aspects.<sup>15</sup> Students watching videos of themselves interviewing patients improved their interviewing skills. In PBL some researchers feel that tutors are not able to observe all learning processes in students and that there are some learning activities that only students can truly assess themselves.<sup>12</sup> Self-assessment is useful for ongoing efficient self-reflection and leads to personal development which is in keeping with the PBL process.

Some educators however, have questioned the accuracy, reliability and validity of self-assessment as they believe that students are lenient to themselves. There may be "over-rating or under-rating": the tendency being over-rating.<sup>19</sup> Machado showed that, self-assessment scores only correlated with peer-assessment scores, not tutor scores: concluding that students' scores are not useful for summative assessment, but should only be used for formative purposes.<sup>4</sup> In Das et al study, 75% of students, and 83% of tutors echoed that, students must be able to grade their work accurately, for their grades to be reliable.<sup>19</sup> Reports on Falchikov and Boud's study of 1989, showed correlations between self-assessment and teacher-assessments ranged from -0.05 to 0.82 (mean 0.39) and Gordon's study on health professionals, showed correlations between self-assessment and actual

factual knowledge ranged from 0.02 to 0.65.<sup>16</sup> Some writers however, have the opposing view that if marks for self-assessment are "meaningful" students will engage more. If marks count for a percentage in summative scores more students will engage more.<sup>12</sup> There are ways to improve the success and reliability of self-assessment process: indeed, key components have been identified.<sup>17</sup>

### **Research setting**

Many schools have a hybrid of PBL and didactic teaching, as is the case in our school.<sup>3</sup> Our PBL system is more "guided discovery approach", rather than "open discovery".<sup>18</sup> Teaching is systems-based and emphasizes application and integration of basic, clinical knowledge and skills, in keeping with PBL. In third-year PBL groups are composed of 9-12 students with random assignment of tutors and students. At the end of the course, students are scored on a 12point/category rubric which address: (1) Problem solving, (2) information processing and cognitive reasoning strategies, (3) Metacognitive skills, (4) Interpersonal skills and attitudes (5) communication skills (6) general social and scientific attitudes and work habits. Each of the twelve points is scored from 0 to 5; maximum score-60. The rubric also provides an overall or holistic measure of performance. Again, students are rated from 0-5. Total PBL score is 65points. This is scaled down (dividing by 13) to a final PBL score of 5. The tutor's final PBL score forms part (5%) of the continuous assessment (CA) (or in-course assessment). (CA accounts for 30% of the final course grade. The PDQ makes up the 25%). Two other sections are assessed: attendance and punctuality (though this section does not form part of the 65 marks).

Evaluation of the PBL process itself by tutors and students forms part of course evaluation. Yet there has not been any formal evaluation of students' self-monitoring/assessment activities.

This study was done to address this. How do students see their progress and learning within PBL? What do they see as their strengths and weaknesses? At the end of the course students also take an integrated progressive disclosure questions (PDQ) examination:<sup>18</sup> integrating all sub-specialties involved in PBL, (anatomical pathology, chemical pathology, haematology, immunology, microbiology, pharmacology and public health), where a clinical-scenario is presented: information revealed progressively and questions asked at each stage (free-response, short answers). (A score of <50% is a fail, 50-69% is a pass, 70-74% pass with second-class honours, 75-79%, pass with first-class honours, and >= 80% pass with distinction).

Objectives of the study was to describe third-year medical students' self-assessment of their performance during PBL and correlate their scores with tutors' scores, and their scores in a written progressive disclosure questions examination.

**METHODS**

Approval was obtained from the Ethics Committee and the Office of The Dean, Faculty of Medical Sciences. Using a twelve-point rubric developed by the Centre for Medical Sciences Education students rated their activities in PBL (at the end of a Semester-2 course).

Overall performance, attendance and punctuality were also assessed. Descriptive statistics (SPSS) and Pearson correlations (r) were performed.<sup>19</sup> Students’ self-assessment scores (on the 12point rubric) were compared and correlated with tutors’ scores. Measure of agreement was calculated (% matches, % under-rating, and % over-rating (where students scored themselves lower/higher than the tutor).

Tutors’ final PBL scores(T-FPS) were compared and correlated with the grade as assessed by the students (students’ final PBL score, st-FPS). Bland-Altman graphs were plotted to describe levels of agreement between T-FPS and st-FPS.<sup>20,21</sup> (The mean of the scores was plotted against the difference between the tutors and students’ scores.

Three lines were drawn: mean difference, upper and lower limits of agreement. The limits of agreement were defined as the mean difference ±5% of the maximum score).

PDQ questions were constructed and graded only by teachers, (each sub-specialty constructed and graded their

own questions. The final paper was reviewed by a core-committee comprised of representatives from all subspecialties). PDQ score was correlated with the T-FPS), the st-FPS and final CA grade (FCA). (FCA = T-FPS +PDQ score).

Agreements for % matches/under/over-ratings for the high-achievers (students who scored =>70% in the PDQ ie honours/distinction students), and low-achievers (scored <50% in the PDQ) were compared.

Rubric point 5 (which deals with cognitive skills/critical thinking) by both tutors and students, was correlated with PDQ scores.

Tutors’ and students’ scores for attendance and punctuality were compared. Tutors’ scores and students’ scores for the high-achieving and low-achieving students were further analysed.

**RESULTS**

There were 20 tutorial groups, 20 tutors, 13 males, 7 females: all medical doctors, with specialist qualifications, in different sub-specialties involved in PBL, ranging in experience in the department from 6months to >20 years. Of the 244 students in the class, 18 (7.4%) were high-achievers, with range of PDQ scores of 17.5 (70%) to 22 (88%), and 96 (39.3%) were low achievers, with range of PDQ scores of 4.83 (19.32%) to 12.45 (49.8%). Out of the 244 students 181(74.2%) completed the 12-point rubric (hence were in the study).

**Table 1: Comparison of tutor and students’ scores for 12-point rubric, and correlations(r).**

Rubric point	1	2	3	4	5	6	7	8	9	10	11	12
Min score	T-2 S-3	T-2 S-3	T-2 S-3	T-2 S-3	T-2 S-3	T-2 S-2	T-2 S-2	T-2 S-3	T-2 S-2	T-2 S-3	T-2 S-3	T-2 S-3
Max score	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5
Modal score	T-4 S-5	T-4 S-5	T-4 S-5	T-4 S-5	T-4 S-5	T-4 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-5 S-5	T-4 S-5	T-5 S-5
r between T and S	0.25	0.08	0.09	0.20	0.16	0.12	0.18	0.09	0.15	0.10	0.09	-0.03
(Sig level)	(0.01)	(NS)	(NS)	(0.01)	(0.05)	(0.12)	(0.05)	(NS)	(0.05)	(NS)	(NS)	(NS)

T-tutor, S-student, r-Pearson correlation, Sig level-significance level).

**Comparison of scores for 12-point rubric**

There was a wide distribution of tutors’ and students’ scores (2-5) (Table 1). Modal scores for tutors were 4 and 5 and for students 5. There were negligible to weak correlations between the tutors’ and students’ scores (Table 1). For rubric-point 12 (Collaborative and decision-making skills), the relationship was even a weak negative.

Agreements/Matches for the 12-point rubric matches ranged from 39.2% to 52.8%, under-ratings 10.5% to 24.9% and over-ratings 29.2% to 49.2% (Figure 1).

**Final tutors’ PBL score compared to final students’ PBL score**

Mean T-FPS was 4.370±0.6920, and mean st-FPS was 4.564±0.3510. Tutor and students’ scores were statistically significant (Table 2). t-test=0.01.

Correlation(r) between tutors' and students' scores was low,  $r=0.181$ . (Significance (2-tailed)\* =0.15) Bland-Altman plots (Figure 2): showed good level of agreement for students and tutors.

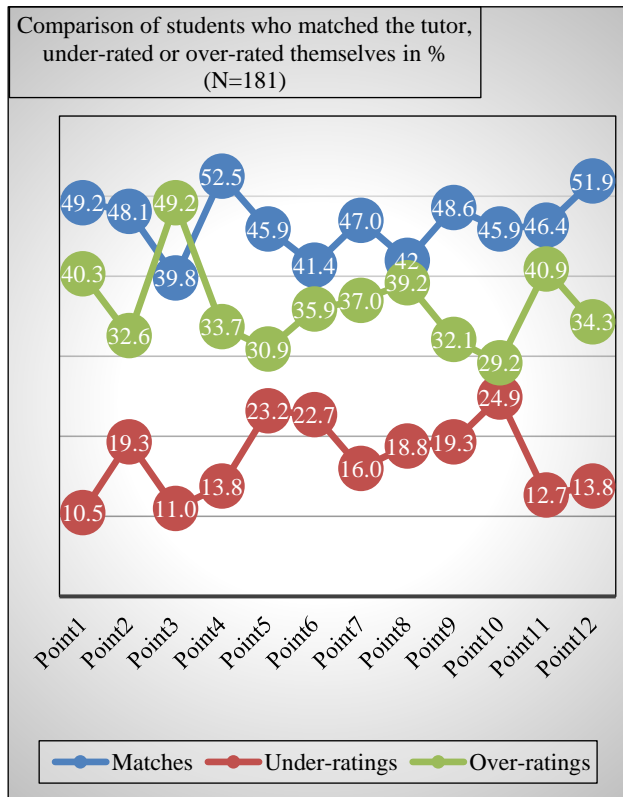


Figure 1: Comparison of matches, under-ratings and over-ratings (12-point rubric).

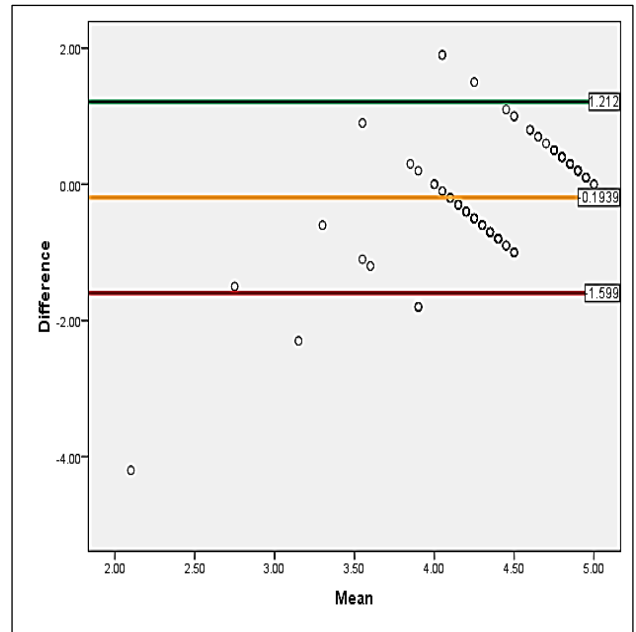


Figure 2: Bland-altman graphs of students' and tutors' final PBL scores (Std deviation=0.717).

**PDQ Analysis and correlations with Final PBL score and Final CA**

The mean PDQ score was 13.341 (53.4%)±3.1165. The mean FCA was 17.586 (58.6%)±3.2896.

There was a high correlation between the PDQ and FCA. Correlations between PDQ and T-FPS, and PDQ and st-FPS were low (Table 2).

Table 2: Comparison between tutors' and student's final PBL score, PDQ, and final CA grade.

Partial correlations between students' and tutor's final PBL score, PDQ and final CA				Partial correlations between tutors, and students' final PBL score, PDQ and final CA							
Control variables		T-FPS	PDQ	FCA	Control variables		st-FPS	PDQ	FCA		
st-FPS	T-FPS	Correlation- r	1.000	0.100	0.264	st-FPS	Correlation- r	1.000	0.119	0.129	
		Significance (2 tailed)		0.181	0.000		Significance (2 tailed)		0.112	0.084	
		Df	0	178	178		Df	0	178	178	
	PDQ	Correlation- r	0.100	1.000	0.979	T-FP	PDQ	Correlation- r	0.119	1.000	0.993
		Significance (2 tailed)	0.181		0.000		Significance (2 tailed)	0.112		0.000	
		Df	178	0	178		Df	178	0	178	
FCA	Correlation- r	0.264	0.979	1.000	FCA	Correlation- r	0.129	0.993	1.000		
	Significance (2 tailed)	0.000	0.000			Significance (2 tailed)	0.084	0.000			
	Df	178	178	0		Df	178	178	0		

Df-degrees of freedom, FCA-final continuous assessment, PDQ-progressive disclosure questions, st-FPS-students' final pbl score, t-FPS, tutor final pbl score.

High and low achievers of the 181 students who completed the 12-point rubric, 69 (38.1%) were low-achievers, 15 (8.3%) were high-achievers. For this group of low-achievers, on the 12-point rubric, matches ranged from 39.1%-60.0%, under-ratings from 8.7%-26.1%,

over-ratings from 29.0%-46.4%. (Table 3) For these 15 high-achievers matches ranged from 33.3%-66.7%, under-ratings from 6.6%-26.7%, over-ratings, 13.3%-46.7% (Table 3).

**Table 3: Comparison of matches, under and over-ratings (12-point rubric) for low and high-achievers.**

Rubric point	Low achievers (PDQ <50%) (N=69 students)			High achievers (PDQ >=70%) (N=15 students)		
	Matches	Under-ratings	Over-ratings	Matches	Under-ratings	Over-ratings
1	40 (60.0%)	8 (11.6%)	21 (30.4%)	7 (46.7%)	1 (6.6%)	7 (46.7%)
2	36 (52.2%)	13 (18.8%)	20 (29.0%)	6 (40%)	4 (26.7%)	5 (33.3%)
3	31 (44.9%)	6 (8.7%)	32 (46.4%)	8 (53.4%)	2 (13.3%)	5 (33.3%)
4	35 (50.7%)	10 (14.5%)	24 (34.8%)	7 (46.7%)	2 (13.3%)	6 (40.0%)
5	31 (44.9%)	16 (23.2%)	22 (31.9%)	7 (46.7%)	4 (26.7%)	4 (26.7%)
6	33 (47.8%)	10 (14.5%)	26 (37.7%)	5 (33.3%)	4 (26.7%)	6 (40.0%)
7	34 (49.3%)	11 (15.9%)	24 (34.8%)	6 (40.0%)	3 (20.0%)	6 (40.0%)
8	27 (39.1%)	13 (18.9%)	29 (42.0%)	9 (60.0%)	2 (13.3%)	4 (26.7%)
9	34 (49.3%)	13 (18.8%)	22 (31.9%)	10 (66.7%)	1 (6.6%)	4 (26.7%)
10	29 (42.0%)	18 (26.1%)	22 (31.9%)	9 (60.0%)	4 (26.7%)	2 (13.3%)
11	32 (46.4%)	7 (10.1%)	30 (43.5%)	6 (40.0%)	2 (13.3%)	7 (46.7%)
12	32 (46.4%)	11 (15.9%)	26 (37.7%)	9 (60.0%)	4 (26.7%)	2 (13.3%)

(Table 4) shows descriptive statistics and correlations between low and high-achievers. There was a strong

negative relationship between st-FPS and PDQ scores among high-achievers, and weak positive relationship between the st-FPS and PDQ among low-achievers.

**Table 4: Descriptive statistics between low and high-achievers (Students' self-score and PDQ scores).**

	Low achievers (<50% in PDQ)			High achievers (>=70% in PDQ)		
	Mean	Std deviation	Correlation (r) between PBL self-score st-FPS vs PDQ	Mean	Std deviation	Correlation (r) between PBL self-score st-FPS vs PDQ
Students' PBL self-score st-FPS	4.544	0.363	0.27 (sig: 0.05)	4.64	0.30	-0.68 (sig: 0.01)
PDQ score	10.283	1.524		18.79	1.56	

**Table 5: Comparison of tutor/students' scores on "Cognitive reasoning/critical thinking skills" and PDQ.**

	Tutor's score	Students' score
Max score	5	5
Mean	4.354	4.444
Std deviation	0.6724	0.6000
Pearson correlation (r) with PDQ score	0.10 (Low correlation)	0.11 (Low correlation)
Pearson correlation (r) between tutor's score and students' score	0.11 (Low correlation)	

**Comparison of rubric point 5 (Cognitive reasoning/critical thinking) with PDQ scores**

low. Correlation between the tutors' score and the students score was also low.

(Table 5) Correlations between tutors' score and the PDQ, and between students' scores and PDQ were both

*Attendance and punctuality* 146 (80.7%) of the 181 students completed the section on attendance and 148

(81.8%) on punctuality. 12 (80%) of the 15 high achievers completed both the attendance and punctuality sections, 52 (75.4%) low-achievers (out of the 69) completed the attendance section, and 54 (78.3%)

completed the section on punctuality. There were discrepancies between tutors' scores and students' scores (Table 6).

**Table 6: Comparison of attendance/punctuality between students and tutors' scores for whole group, high and low-achievers.**

		Matches	Under-ratings	Over-ratings	Total responders	Non-responders
Whole group	Attendance	120 (82.2%)	10 (6.8%)	16 (11%)	146 (100%)	35 (19.34%) out of 181
	Punctuality	140 (94.6%)	4 (2.7%)	4 (2.7%)	148 (100%)	33 (18.23%) out of 181
High achievers >=70%	Attendance	11 (91.7%)	-	1 (8.3%)	12 (100%)	3 (20%) out of 15
	Punctuality	12 (100%)	-	-	12 (100%)	3 (20%) out of 15
Low achievers <50%	Attendance	38 (73.1%)	4 (7.7%)	10 (19.2%)	52 (100%)	17 (24.64%) out of 69
	Punctuality	49 (90.7%)	4 (7.4%)	1 (1.9%)	54 (100%)	15 (21.74%) out of 69

## DISCUSSION

### *Correlations/agreements 12-point rubric*

Of the 181 students in the study, there were weak correlations between students' and tutors' scores in the 12-point rubric. A considerable number of students over-rated themselves: like what has been reported in the literature. Agreement between tutor and students' scores is difficult to achieve.<sup>1</sup>

There were also students who under-rated themselves. The high self-scores may be scores that students would like to get, not what they deserve, or self-confidence, as in Das et al's study.<sup>9</sup> Among the under-raters, could it be that these students don't want to appear boastful? as did students in Sadler's study, who were concerned about appearing boastful. In Donnon's study, students scored themselves lower than their peers, and in Papinczak's study students under-scored themselves: thought to be due to confidence, self-awareness, humility and objectivity.<sup>13,21,22</sup> Luciano even suggested a lack of transparency in assessments, or possibly unfamiliarity with the PBL process.<sup>4</sup>

However, unlike in Luciano where the analysis included students from 1<sup>st</sup>/2<sup>nd</sup> years, this study, only involved third-year students who had done PBL for first year.<sup>4</sup> Students can be taught to assess themselves. Kelberlau-Berks showed that with regular reflection and guidance by teachers, even seventh-graders were accurate in self-assessment.<sup>23</sup> Indeed teaching them early helps students to take ownership of their learning. With time students even completed assessments with little help from teachers. Self-assessment saves teachers time and effort.<sup>24</sup> Students trained in self-assessment out-performed those who were not.<sup>24</sup> Through self-assessment they were able to identify and address their weaknesses early. Students

who are "more active" during PBL should be better able to assess their performance.<sup>25</sup>

### *PBL scores compared to PDQ scores for low and high-achievers*

There was a strong negative correlation between high-achievers' PDQ and PBL scores. This may mean that high-achievers are more critical of themselves, which drives them to work harder. This too was like studies in literature,<sup>1,22</sup> where poorer students scored themselves more generously, compared to academically stronger students,<sup>1,22,26</sup> who tended to score themselves more harshly.<sup>26</sup> This may also suggest the possibility of poor understanding of self-assessment criteria. Indeed Tan, stated that this lower academic ability was related to low self-assessment ability<sup>1</sup>, or inability to assess their own quality of work,<sup>26</sup> not understanding the rubric itself by low-achievers, which in-turn translates to "lack of learning".<sup>22</sup> Hence self-assessment is not only a skill to be learned, but also in itself, is an "important component of one's competence".<sup>25</sup> Some authors have suggested even involving students in formulating the actual self-assessment tool, and criteria.<sup>1,12,17</sup> Interestingly, in Whitfield's study,<sup>27</sup> tutors' scores were higher than students' examination scores. This was seen more in the low 25% than the high 25% of the class. Thus, the tutors' assessment of PBL knowledge was said to be not useful. Some tutors "may not be willing to fail students".<sup>27</sup> There were no significant differences in self-scores compared to peer-scores in high-achievers in the study by Donnon.<sup>13</sup> At the same time, a question may be raised about factors that may affect results of written examinations too-including preparation and anxiety.<sup>28</sup>

Clearly self-scoring, doesn't always discriminate low and high-achievers.<sup>4</sup> Students' judgement of their performance, may be unreliable if there are consequences, (e.g. scores that affect final grade). The

use of self-assessment scores as final summative scores, could promote more generous self-scoring,<sup>4</sup> affecting productivity of PBL. Perhaps it would be useful for our students, to have this formative exercise mid-course, not end of course. Identifying their deficiencies and weaknesses early-on could help direct their learning and better preparation for PDQ. They do need help though: improving their self-assessment skills, helps students to evaluate or appraise their performance as an ongoing process, of personal development planning (PDP) as they reflect on themselves.<sup>9</sup> Students should reflect on their own learning, their goals, performance and achievements. This improves by reviewing, planning and essentially taking responsibility for their learning.<sup>29</sup> Thus, this PDP helps them with life-long learning in keeping with PBL process. PDP will help them in a holistic manner, academically, personally, and professionally. It must however be honest self-reflection and analysis, noting strengths and weaknesses, which will help<sup>30</sup> them get valuable learning experiences. Discussion of score discrepancies, mid-course, would be useful in sharpening their habits to self-evaluate, and would lead to better understanding of course objectives.<sup>9</sup>

These discussions would also be useful in improving communication between students and tutors.<sup>9</sup> Course coordinator could facilitate. Self-assessment should be introduced early in first-year as suggested by others.<sup>12</sup> Students can be taught early, proper self-assessment skills. Sadler,<sup>22</sup> indeed showed that even middle-school students were capable of grading tests accurately and learned from the grading process. 7<sup>th</sup> graders were accurate in their self-ratings.<sup>23</sup> Regular self-assessments help tutors to recognize poorer students.<sup>23</sup> However, as suggested by others, emphasis should be more on self-assessment “ability” not just seeing the “activity” of self-assessment.<sup>1</sup> Indeed, suggestions have been made to even reward (with extra points), accurate self-grading, using the teacher’s grade as the ruler.<sup>22</sup> Reporting on Boud, Sadler reported that students tended to be more accurate when reporting on “effort” as opposed to when assessing their “achievement”.<sup>22</sup> Indeed self-assessment by students in a learning context is different from one for summative purposes.<sup>2</sup> Tutor based assessment is useful in that it is done over a prolonged period of time<sup>25</sup> where students’ performance is observed over prolonged interactions.

#### ***Rubric point 5 (cognitive/critical thinking) and PDQ scores***

PBL scores did not predict the performance in the summative PDQ, also similarly shown previously.<sup>12</sup> Ward,<sup>16</sup> showed correlations between self-assessment and actual factual knowledge of as low as 0.02 (in an oral examination). However, it could be said that tutor assessment of cognitive skills in PBL, requires oral communication skills, whereas the PDQ requires written skills. Furthermore, the point of English as a second language was raised by Das.<sup>9</sup> However, in our setting English is the first language for most students. Whitfield

suggested that tutors’ assessment of PBL knowledge was not useful.<sup>27</sup> A question of tutors’ content knowledge was also raised,<sup>27</sup> where non-subject specialists assessed PBL: like our setting: however even though the PDQ is integrated, PDQ-question construction/grading is specialty-based. Our tutors’ PBL scores are subjective despite the rubric. Hence, the PBL tutors’ scores account for only 5% of the final course grade. In Joseph et al’s study, the correlation between tutor and students’ scores for the brainstorming part of PBL were significant, and they suggested that tutors could thus focus on assessing the presentation part of PBL.<sup>5</sup>

#### ***Attendance and punctuality***

Interestingly the study shows some disagreements between students and tutors on attendance and punctuality. Again, could it be the scores that the students wish to get? On attendance, it may also be that the tutor may take the register at the start of the session and students who come in late may not necessarily be marked present. Regarding punctuality there may be differences in interpretation of what “late-comer” means. This emphasizes the need for proper training in the assessment process, and agreement on assessment criteria. Also interesting is that not all students who completed the 12-point rubric (181 students) completed the section on attendance and punctuality.

This section does not constitute part of the 65 points that constitute the final PBL grade. Again, this may be because students recognize this, and they won’t do something that isn’t worth a grade.

Limitation of this study was the study did not seek the views of students and tutors. Kelberlau-Berks reported that students felt the positive experience achieved its goals.<sup>23</sup> In Das’s study, students said the useful exercise clarified course objectives, roles and requirements during PBL.<sup>9</sup> Tutors saw improved confidence among students. However, in Bollela’s study, some students felt they were not trained enough nor confident to perform self-assessment, the exercise was “tough”, and that PBL itself “forces” students to talk even if they may not like talking! Students believe that good grades are related to how well they do in oral presentation.<sup>31</sup> Tutors thought that students were not “mature enough” to perform assessments.<sup>31</sup>

The study was done at one point only. Das did self-evaluation repeated over a period which showed no variation over time.<sup>9</sup> Many evaluations are time consuming.<sup>9</sup> However, they could help tutors identify students who tend to “pad” their scores, hyper-inflating their scores undeservably.<sup>22</sup>

On the other hand, tutors and students’ scores might correlate better over time, as students gain more experience with the regular feedback.<sup>25</sup>

Gender and age differences were not analysed. Most students were female, majority of tutor's male. In Das's study, male students gave higher self-evaluations than female students.<sup>9</sup> In Donnon's study, female students had higher ratings than male students concerning professional and personal attributes.<sup>13</sup> Majority of students were about 21 years old, with no previous university degree. Papinczak, showed that students with previous arts degrees scored themselves higher than students with previous science degrees.<sup>21</sup> Only 4 students were repeating the course in the study).

The results were not analysed for years of tutor experience, nor per subject specialist. Is there tutor-bias, subjectivity or unfairness in non-subject specialists judging knowledge in PBL?<sup>27</sup> Indeed, in Das's analysis<sup>9</sup>, both students and tutors reported that some tutors were stricter than others. Jaiprakash showed that the correlation coefficient between students' written MCQ test scores (based on the PBL case discussed) and tutor's scores was higher if tutors that had been trained in the PBL process: hence highlighting the importance of tutor training. Jaiprakash's study also highlighted the need to use standardised tests like MCQs for PBL and not just relying on tutors' scores only.<sup>32</sup>

The curriculum is a hybrid of PBL and didactic lectures, requiring different learning strategies. Perhaps the students concentrate on one strategy and ignore others, which may be a limiting factor in the impact of the results, and hence may not be generalizable. However as reported by others, even in non-PBL curricula, in lab/clinical skills, the poor correlations were also shown between tutor/students' scores and examinations.<sup>21,33-35</sup> Accuracy of self-assessment was more closely related to how familiar the task was.<sup>33</sup> Students were also inaccurate in predicting their own scores in written examinations but overestimated their examination performance significantly.<sup>28</sup>

Self-assessment is a skill that can be trained. Students should be motivated to participate in this learning exercise, even if there is no grade, without fear of victimization. Future efforts can aim to have the exercise mid-course, providing useful feedback, earlier, as advantages have been shown in the literature. It is an important part of formative learning, a useful tool for ongoing learning.<sup>10,15</sup> Learning improves when students engage in the learning process.<sup>15</sup> Lifelong learning demands students to self-reflect, assess their strengths and weaknesses, and make decisions to self-adjust.<sup>3</sup> Taking responsibility for their own learning helps students learn more, instead of just concentrating on final grades.<sup>3</sup> Self-assessment has the potential to ensure transparency in grading students, and students would get better understanding of PBL requirements and assessments.<sup>1</sup> Students trained in self-assessment outperformed those not trained.<sup>24</sup> Clinical reasoning improves when students focus on deeper learning approaches, which engage students, and involve students

in the learning process.<sup>37</sup> Koiriyah developed a self-assessment scale for active learning and critical thinking, which may guide students to achieve the PBL objectives and stimulate better self-regulated learning.<sup>14</sup>

## ACKNOWLEDGEMENTS

Authors would like to thank the former Dean, (professor Samuel Ramsewak)- the head and staff of the department, Para-clinical sciences, and Mrs. Louise Green, former staff of centre for medical sciences education, faculty of medical sciences, university of the West Indies, St Augustine, Trinidad and Tobago.

*Funding: No funding sources*

*Conflict of interest: None declared*

*Ethical approval: The study was approved by the Institutional Ethics Committee*

## REFERENCES

1. Tan K, Keat LH. Self and Peer Assessment as an Assessment Tool in Problem-based Learning. 2005:162-75. Available at [http://www.tp.edu.sg/staticfiles/TP/files/centres/pbl/pbl\\_kelvin\\_and\\_ho\\_keat.pdf](http://www.tp.edu.sg/staticfiles/TP/files/centres/pbl/pbl_kelvin_and_ho_keat.pdf).
2. Tousignant M, DesMarchais JE. Accuracy of student self-assessment ability compared to their own performance in a problem-based learning medical program: a correlation study. *Adv Health Sci Edu.* 2002;7(1):19-27.
3. White CB. Smoothing out transitions: How pedagogy influences medical students' achievement of self-regulated learning goals. *Adv Health Sci Edu.* 2007;12(3):279-97.
4. Machado JL, Machado VM, Grec W, Bollela VR, Vieira JE. Self-and peer assessment may not be an accurate measure of PBL tutorial process. *BMC Med Edu.* 2008;8(1):55.
5. Joseph N, Rai S, Madi D, Bhat K, Kotian SM, Kantharaju S. Problem-based learning as an effective learning tool in community medicine: Initiative in a Private Medical College of a developing Country. *Indian J Community Med.* 2016;41(2):133-40.
6. Hmelo CE, Gotterer GS, Bransford JD. A theory-driven approach to assessing the cognitive effects of PBL. *Instructional Science.* 1997;25(6):387-408.
7. Okubo Y, Matsushita S, Takakuwa Y, Yoshioka T, Nitta K. Longitudinal PBL in undergraduate medical education develops lifelong-learning habits and clinical competencies in social aspects. *Tohoku J Exp Med.* 2016;238(1):65-74.
8. Valle R, Petra L, Martínez-Gonzalez A, Rojas-Ramirez JA, Morales-Lopez S, Pina-Garza B. Assessment of student performance in problem-based learning tutorial sessions. *Med Edu.* 1999;33(11):818-22.
9. Das M, Mpofu D, Dunn E, Lanphear JH. Self and tutor evaluations in problem based learning tutorials: is there a relationship? *Med Edu.* 1998;32(4):411-8.
10. Hanley K, Zabar S, Charap J, Nicholson J, Disney L, Kalet A, et al. Self-assessment and goal-setting is



- associated with an improvement in interviewing skills. *Med Edu Online.* 2014;19.
11. Northern Illinois University. Self and peer assessment. spectrum newsletter. fall 2009. faculty development and instructional design center. Available at <http://www.niu.edu/spectrum/2009/fall/selfpeer.shtml> Accessed June 17, 2016.
  12. Nulty D. A Guide to Peer and Self-Assessment: Approaches and Practice Strategies for Academics. Griffith institute for higher education. Griffith university. 2008. Available at [https://www.griffith.edu.au/\\_\\_data/assets/pdf\\_file/0016/142108/GuidePeerSelfAssessment-Long.pdf](https://www.griffith.edu.au/__data/assets/pdf_file/0016/142108/GuidePeerSelfAssessment-Long.pdf).
  13. Donnon T, McIlwrick J, Woloschuk W. Investigating the reliability and validity of self and peer assessment to measure medical students' professional competencies. *Creative Educ.* 2013;4(06):23-8.
  14. Khoiriyah U, Roberts C, Jorm C, Van der Vleuten CPM. Enhancing students' learning in problem based learning: validation of a self-assessment scale for active learning and critical thinking, *BMC Med Edu.* 2015;15:140.
  15. Ramirez BU. Effect of self-assessment on test scores: student perceptions. *Adv Physiol Educ.* 2010;34(3):134-6.
  16. Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. *Adv Health Sci Edu.* 2002;7(1):63-80.
  17. Boud D. Enhancing learning through self-assessment. London: Kogan; 1995. Available at [http://onlinelibrary.wiley.com/doi/10.1016/0307-4412\(96\)82523-4/pdf](http://onlinelibrary.wiley.com/doi/10.1016/0307-4412(96)82523-4/pdf).
  18. Vuma S, Sa B. Evaluation of the effectiveness of progressive disclosure questions as an assessment tool for knowledge and skills in a problem based learning setting among third year medical students at the university of the West Indies, Trinidad and Tobago. *BMC Res Notes.* 2015;8(1):1.
  19. Pearson's r Correlation Available at <http://faculty.quinnipiac.edu/libarts/posci/Statistics.html>.
  20. Bland MJ, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;327(8476):307-10.
  21. Papinczak T, Young L, Grooves M, Haynes M. An analysis of peer, self and tutor assessment in problem-based learning tutorials. *Med Teach.* 2007;29(5): e122-32.
  22. Sadler PM, Good E. The impact of self-and peer-grading on student learning. *Edu Assessment.* 2006;11(1):1-31.
  23. Kelberlau-Berks DR. The effects of self-assessment on student learning. 2006. Lincoln, Nebraska. Available at <http://scimath.unl.edu/MIM/files/research/BerksD.pdf>.
  24. Belski I. The Impact of self-assessment and reflection on student learning outcomes. In proceedings of the 21st annual conference for the Australasian association for engineering education, 2010. Engineers Australia. Available at [https://www.researchgate.net/profile/Iouri\\_Belski/publication/228460756\\_The\\_Impact\\_of\\_Self-Assessment\\_and\\_Reflection\\_on\\_Student\\_Learning\\_Outcomes/links/02e7e5338b26ab3089000000.pdf](https://www.researchgate.net/profile/Iouri_Belski/publication/228460756_The_Impact_of_Self-Assessment_and_Reflection_on_Student_Learning_Outcomes/links/02e7e5338b26ab3089000000.pdf).
  25. Eva KW. Assessing Tutorial-Based Assessment. *Adv Health Sci Edu.* 2001;6(3): 243-57.
  26. Langendyk V. Not knowing that they do not know: self-assessment accuracy of third-year medical students. *Med Edu.* 2006;40(2):173-9.
  27. Whitfield CF, Xie SX. Correlation of problem-based learning facilitators' scores with student performance on written exams. *Adv Health Sci Edu.* 2002;7(1):41-51.
  28. Shaban S, Aburawi EH, Elzubeir K, Elango S, El-Zubeir M. Factors Influencing Medical Students' Self-assessment of Examination Performance Accuracy: A United Arab Emirates Study. *Edu Health.* 2016;29(2):75-81.
  29. Guide to the completion of a personal development plan. Available at <https://www.managers.org.uk/~media/ET/checklist/Personal-Development-Plan-Example-Guide.pdf>.
  30. Personal development planning: guidance for institutional policy and practice in higher education. Available at <http://www.recordingachievement.ac.uk/images/pdfs/pdpguide2009.pdf>.
  31. Bollela VR, Gabarra MH, da Costa C, Lima RC. Students and tutors' social representations of assessment in problem-based learning tutorials supporting change. *BMC Med Educ.* 2009;7:9:30.
  32. Jaikaprash H, Min AK, Ghosh S. Increased correlation coefficient between the written test score and tutors' performance test scores after training of tutors for assessment of medical students during problem-based learning course in Malaysia. *Korean J Med Educ.* 2016;28(1):123-5.
  33. Fitzgerald JT, White CB, Gruppen LD. A longitudinal study of self-assessment accuracy. *Med Edu.* 2003;37:645-9.
  34. Watts WE, Rush K, Wright M. Evaluating first-year nursing students' ability to self-assess psychomotor skills using videotape. *Nursing Edu Perspectives.* 2009;30(4):214-9.
  35. Mehrdad NN, Bigdeli S, Ebrahimi H. A comparative study on self, peer and teacher evaluation to evaluate clinical skills of nursing students. *Procedia-Social Behavioural Sci.* 2012;47: 1847-52.
  36. Ingchatcharoena S, Tangdhanakanonda K, Pasiphola S. Testing measurement invariance of quality rating causal models in tutorial-based assessment. *Procedia - Social and Behavioural Sciences.* 2016;217:867-77.
  37. Hood S, Chapman E. A. Tutorial Programme to Enhance Psychiatry Learning Processes within a PBL-Based Course. *Edu Res Perspectives.* 2011;38(2):69-86.

**Cite this article as:** Vuma S, Bidyadhar Sa. Self-assessment: how do third year medical students rate their performance during problem-based learning? *Int J Res Med Sci* 2017;5:3044-52.