An Investigation into the Efficacy of Students-as-partner Pedagogy in a Singapore University Education Learning Context

Hui Min Chon, Nanyang Technological University, Singapore
Chun Chau Sze, Nanyang Technological University, Singapore
Wilson Wen Bin Goh, Nanyang Technological University, Singapore

The Asian Conference on Education 2019
Official Conference Proceedings

Abstract
In Students-as-partner (SAP), students work in partnership with staff members in higher learning institutions to facilitate deeper learning in students by promoting student engagement. While SAP's impact on student consultants and staff members directly involved in partnership is generally well and widely researched, relatively little is reported about its application on student learning in an Asian university context. This case study reports the efficacy of SAP-produced teaching resource on a cohort of biology major students in Nanyang Technological University, Singapore (NTU). Participants were either exposed to the student-centric (collaborative work between faculty and student) or tutor-centric (produced by the faculty) teaching resource and their test performance and perception of the respective materials compared. Our data shows that students generally prefer tutor-centric material with consistent and higher improvement in test scores when they knew material as tutor-centric compared to the learner-centric alternative. Although SAP is a high-impact pedagogical practice, this study suggests cultural context can confound outcome, and that at least in NTU's predominantly conservative Asian setting, and where collaboration in content creation is concerned between faculty and student, stronger buy-in and gradual introduction is necessary.

Keywords: Outcome-based teaching and learning (OBTL), Students-as-partner (SAP) Technology enhanced learning (TEL)

Introduction

Motivation for study

In Students-as-partner (SAP), students adopt an active producer role alongside teaching faculty in the process of teaching and learning as opposed to the consumerist role typically expected of them (Bovill et al., 2015). SAP involves engaging students as collaborators, specifically acting as co-inquirers or co-producers, which has the effect of levelling the power imbalance between faculty and students such that students have more of a say in what they are learning (Cook-Sather, 2013).

There had been much research on the benefits and efficacy of SAP in promoting deeper learning in students (Healey et al., 2016). However, these mainly revolved around qualitative aspects of SAP intervention with heavy focus on the perception and reflections of student partners themselves and the staff involved in the SAP projects.
This study aimed to fill the gap in knowledge regarding the efficacy of SAP to enhance learning in higher education students in an Asian University setting by evaluating the impact of SAP-generated teaching material in a cohort of biology students through the use of mixed method analysis.

Literature review

In this section, we first discuss the concept of student-as-partner and summarize the benefits of SAP in practice. We then discuss one relevant case study done on the effectiveness of resources created by students as partners as perceived by the general student population, followed by an overview of the differences between learners from Western and Asian societies.

Student-as-partner conceptual models and benefits

While partnership falls under the umbrella of student engagement, it is important to note that there are qualitatively different kinds of student engagement and not all involve partnership (Healey et al., 2016). For example passive collection of end-term student feedback in not in itself a true form of partnership.

The four-stage model of student engagement from the NUS/HEA student engagement toolkit (Healey et al., 2014) is useful for describing the different forms of engagement and highlighting the differences among them - mainly consultation, involvement, participation and partnership (Figure 1). Partnership is defined as a “reciprocal process through which all participants have the opportunity to contribute equally, although not necessarily in the same ways, to curricular or pedagogical conceptualization, decision-making, implementation, investigation, or analysis” (Cook-Sather et al., 2014, p. 6-7). Benefits of student-faculty partnerships in teaching and learning have been well documented mainly in three clusters: engagement, awareness and enhancement (Lubricz-Nawrocka, 2018). While students are not disciplinary or pedagogical experts, they do come with the knowledge and experience of being a student, a position faculty members have not occupied for a long time, and possibly disconnected from (Healey et al., 2016). Consequently, collaborative intervention may bridge the gap and disconnect between faculty and students, and precipitate in a better and more meaningful learning experience for everyone involved

(Healey et al, 2016). Furthermore, partnership can transform the student-faculty relationship into one where both parties become colleagues, challenging the constraints of traditionally hierarchical student-teacher relationships.

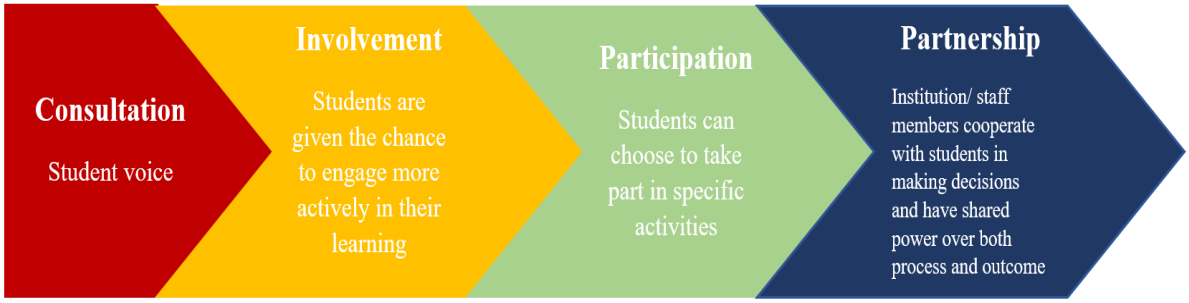


Figure 1. Ladder model showing the four stages of student engagement (Inspired from Healey et al., 2014).

Case study on the effectiveness of student-produced resources perceived by the general student population

Dunn et al., (2018) investigated the effectiveness of student-produced resources as perceived by the general student population in an Australian university. Student responses were split equally into 'yes' or 'no' when asked if it affected them to know that student consultants were involved in producing their material. Of respondents who replied "no", most expressed a preference for more authority in the material produced. Moreover, a small percentage indicated 'no' even when they acknowledged the benefits of student-produced material such as relevance to other students. Interestingly however, the majority of respondents liked the idea of students and faculty co-producing material the most, over other provided options of 'students only' and 'teachers only', an indication that students were open to the concept of collaboration between staff and students in research. Additionally, quantitative analysis of the results after the intervention of the student-produced informative animated videos showed that they were useful in increasing feelings of relevance towards mathematics in first-year engineering students, which was the intended goal of the project.

Differences between the Western and Asian learning environment

The ingrained cultures of the East and West do determine behaviour and thought-process amongst individuals raised in these respective societies. As such, we anticipate culture-related differences in educational values that should be considered while implementing educational reform (Ee and Tan, 2008).

Asian societies such as that of Singapore remain tightly organised by strict social rules and acceptable behaviours that constitute norms. Status, age and even gender are typical factors which influence social interaction between two individuals (Fang and Gopinathan, 2009). Additionally, a high level of importance is placed on hierarchy between superior and subordinate. For example, local teachers assume authoritarian roles while students are relegated to the role of the submissive/passive learner in the classroom. The Confucian heritage that permeates across the Asian-Chinese culture

also places a strong emphasis on learning, instruction and the status of teachers (Ee and Tan, 2008).
In contrast, social interaction within the Western society is relatively less structured, with less importance attributed to rank and social status (Fang and Gopinathan, 2009). Westerners are generally more individualistic and less concerned with winning approval from their in-groups compared to their Asian counterparts (Ee and Tan, 2008) (Figure 2). We expect that differences in social norms to also act as a strong confounder when it comes to the application of high-impact practices such as SAP, which relies on a certain degree of extraversion.

VARIABLES	EASTERN SOCIETY	WESTERN SOCIETY
CULTURE	• Tightly organised • Hierarchical • Greater emphasis on social order and harmony • More concerned with "saving face" or reputation	• Loosely organised • Egalitarian • Individualistic • Greater emphasis in open and democratic exchange of ideas • More concerned with realising one's creative potential in life
SCHOOL	• Classes highly structured • More teacher directed	• Classes less structured • Less teacher directed
FAMILY	• Family pressure serves as a catalyst • Socialise child to be dependent on ingroup	• Expects individual to be a catalyst • Socialise child to be independent

Figure 2. Influence of culture on creative styles and motivational orientation (Modified from Ee and Tan, 2008).

Introduction to the study

We were interested to find out how SAP may impact our local student population and whether the learning benefits extended beyond the student partners directly involved in the partnership. Student consultants were first recruited and tasked to produce their version of a teaching resource on a topic irrelevant to their major of study (data science) after consultation and mentoring with a faculty member. Both learner-centric (developed by both faculty and students) and tutor-centric (developed by the mentoring faculty only) resources were then fielded to discrete cohorts of student participants (who were all from the same biology major as the student consultants themselves). Quantitative and qualitative results from the trials were then analysed.

Materials and Methods

Recruitment and Selection of Student Consultants

A recruitment email was sent to the student body in the School of Biological Sciences (SBS), NTU. Interview and trial sessions were then conducted for respondents. Of the 8 respondents, 5 were selected based on 2 criteria- availability for the duration of the research and their performance on the trials. 3 were disqualified based on their unsatisfactory performance on the trials which did not meet the standards expected of a student consultant. Of the 5 whom offers were made to, only 3 accepted while the other 2 declined, citing the job scope as being a poor fit with their personal interests as the main reason.

Interview and Trial Process

The trial session stimulated what was expected of a practicing student consultant. Interviewees worked in small groups (of 2s and 3s) to design a set of teaching material in the form of PowerPoint slides based on a bio-related topic (Concept of emulsion) with the added twist that it must be suitable for educating a non-bioscience audience. Their work was evaluated for its creativity, originality and palatability as teaching material for a non-bioscience audience. Additionally, several strategic questions were posed to them to assess their suitability and motivation for joining the project (Table 1).

1. How did you approach the problem?
2. How did you distribute the work?
3. What are your intentions and thought processes while doing the task?
4. Did you feel comfortable doing this task?
5. Was it difficult to put yourself into the roles of both learner and teacher?
6. If you had more time, what else would you have tried to incorporate?

Table 1. List of interview questions asked to applicants for the Student Consultant position

Lastly, the basic rules-of-engagement and expectations were established during the interview. In this study, student consultants were paid for their work due to the high time commitment required and high expectations. To protect the rights of both student consultants and the faculty, a clear list of payable and non-payable items was established (Table 2). Such strict delineation on payable items reduced the likelihood of recruiting someone mainly motivated by monetary incentives. The student consultants were paid per hour, with a pre-agreed cap on the time required to complete each task, established and discussed with the faculty prior to execution.

Payable items	Non-payable items
1. Appropriate number of hours on a payable task to be agreed upon before starting it 2. Extensions given where justified 3. Tasks included - Preparation of teaching material - Team discussions (minutes must be recorded and submitted) - Field work or conducting simulation - Writing paper/poster/report (cap of 15 payable hours each) - Reading and presentation of relevant research material (hours agreed upon by PI and student consultants) - Carrying out experiments - Creating software/hardware	1. Training sessions and generic meetings with the Principal Investigator (PI) 2. Training sessions for e-learning software 3. Transport reimbursement 4. Time spent self-studying or researching generic materials not previously approved by the PI

Table 2. List of payable and non-payable items as agreed upon by the student consultants and faculty.

Training of Student Consultants and Production of Student-centric Teaching Material

Student consultants had to create their own set of Data Science-related teaching material based on the topic "Machine Learning". They were given a comprehensive lecture on the topic itself by the faculty first before crafting their own as they had little to no prior exposure to 'Data Science' (the students have no precognition bias, and their knowledge of the field comes due to interact with the faculty involved in this project). Additionally, they were expected to keep journal logs of their thought processes and ideas during the process, which were submitted to the PI for review and to keep the PI updated on their progress.

The student-centric teaching material was then presented as a team to the faculty by the student consultants and any changes discussed were implemented before the final product was produced. Throughout the entire process, the faculty adopted a collaborative mentoring role as opposed to an authoritative one to ensure minimal personal influence on the consultants' work while also making sure no technical mistakes were made in the learner-centric material.

Recruitment of Participants for Paid Trial Sessions and Exclusion Criteria

24 students were recruited from the School of Biological Sciences, NTU across different years of study from year 2 onward. This study deliberately excluded freshmen and seniors who have completed the module BS211S: Equations of Life. This is because freshmen would not have completed the core module BS1008: Biostatistics, which was required of participants as necessary basic statistical knowledge for the trial and for those seniors who have taken BS211S, they would already be more aware and likely more knowledgeable in the content covered by the trial teaching material. Since we were looking for learning gains, it was important that test participants had limited knowledge of the chosen test topic such that any

improvement made on the trial post-test would have been attributed to the intervention of the test material, and not the participant's prior knowledge on the subject matter.

A recruitment notice was disseminated informally through WhatsApp and word-of-mouth and participants signed up by indicating their availability and their particulars on GoogleForms. The sign-up closed after the quota for each session had been reached.

Trial Procedure

To ensure controllability of the trial, participants were split into groups of 4 according to their availability as indicated during sign-up. Each group was enrolled into a trial session covering a unique scenario. Altogether six trials were conducted (Table 3).

During the study, participants were first briefed on what to expect. They were shown instructions on-screen reminding them not to use their phones to minimize distraction and encouraged to do their best on the pre- and post-tests. The pre-test was then fielded, followed by a 15-minute recorded lecture covering the learning material, and finally, administration of the post-test. Additionally, a Likert-scale type survey was included at the end of the post-test to evaluate students' reception towards the lecture material. The participants were then remunerated for their time. The design principles and reasoning for the instruments (pre/post-tests and learning materials) are described in Table 4 (The actual pre- and post-test questions are listed in Appendix A and B respectively). The questions for the Likert-scale palatability evaluation are listed in Table 5.

Both pre- and post-tests were conducted using TurningPoint self-paced polling, a Technology-enhanced learning (TEL) device. Participants were reminded (and encouraged) to choose the "I do not know this concept" option in both tests should they not know the answer. The test material fielded was either designed by the faculty (tutor-centric), or in collaboration with the student consultants (learner-centric).

Session	Scenario
1.	Tutor-centric teaching material fielded with participants unaware about it being tutor-centric
2.	Student-centric teaching material fielded with participants unaware about it being student-centric
3.	Tutor-centric teaching material fielded with it deliberately made known to participants that the teaching material was designed by a professor
4.	Student-centric teaching material fielded with it deliberately made known to participants that the teaching material was designed by student t consultants who underwent training before producing the teaching slides
5.	Tutor-centric teaching material fielded but participants deliberately misinformed that it was produced by trained student consultants
6.	Student-centric teaching material fielded but participants deliberately misinformed that it was produced by a professor

Table 3. Description of scenarios for the six separate trials conducted.

Materials list and rationale behind instrument design

Instrument	Details of instrument design	Rationale behind instrument design
Pre-test questionnaire (Appendix A)	1. Number of questions: 6 2. Test type: Single answer MCQ with 'I do not know this concept' answer option available for all the questions 3. Nature of questions: To assess conceptual knowledge	1. Each question was designed to test participant's knowledge of the subject matter covered by each learning objective of the learning material. 2. MCQ format was most familiar to students and reduced performance anxiety that may hinder their performance. 'I do not know this concept' option eliminated the possibility of participants' guessing the answer when they truly did not know. 3. Students' prior knowledge of the concepts covered in the learning material was evaluated and compared against their ability to apply their knowledge (application skills) in the post-test after the learning material was fielded to them.
Post-test questionnaire (Appendix B)	1. Number of questions: 8 2. Test type: Same as for pre-test 3. Nature of questions: To assess participants' ability to ap ply their knowledge of the subject matter after watching the learning material in the form of video lecture	1. All learning objectives from the learning material were covered 2. Same as for pre-test 3. The ability of participants to apply what they have learned is indicative of effective learning. Furthermore, even if their answers were wrong, analysis of their pre and post-tests could reveal the gap in knowledge that prevented them from getting it right.
Evaluation for palatability – Likert-scale type questionnaire survey	1. Scale of 1 to 5- Strongly agree, Agree, Neutral, Disagree, Strongly disagree. Each option is given a number of points: from 5 for Strongly Agree to 1 for Strongly disagree	1. Scale of 5 was sufficient to evaluate the depth of feeling participants had towards the entire study while a scale of 10 would be excessive and redundant.

	2. Number of questions: 4 3. Nature of questions: To gauge participants' reception of the learning material in regard to its content, organisation, pace of delivery and usefulness in teaching the learning objectives	
Teaching material- "Introduction to Machine Learning"	1. 2 versions: Tutor-centric and Student-centric; both were PowerPoint presentations 2. Lecture content adhered strictly to the learning objectives which were also included in the first slide of the presentation for both versions 3. Mode of delivery: Video of lecture material with voiceover done by the same person for both versions 4. The narrator of the lecture videos was also responsible for writing the transcripts for both of them	2. Done in accordance with NTU's pedagogy of Objective-based Teaching and Learning (OBTL) 3. Presenting it as a voiceover instead of a conventional recorded lecture video with the presenter seen on-screen minimizes distraction to the test participants that may affect their concentration on the lecture material. 4. Transcripts were written strictly based on what was presented in the PowerPoint slides and no additional information was included. The narrator was a third party- neither the PI nor a Student Consultant. Any bias associated with difference in tone of voice or additional information delivered orally was eliminated.

Table 4. List of instruments used in the trial sessions and the rationale behind the instrument design.

On a scale from 1-5 from “Strongly agree” to “Strongly disagree”, rate the following statements that agree best with your sentiments. 1. The concepts explained in the video lecture were clear and easy to follow 2. The pace of the presentation was just right. 3. The presentation slides were well organized and comprehensive 4. If the lecture was presented live and no recordings were made available, the presentation slides contain sufficient material to allow me to understand the concepts based on the learning objectives

Table 5. Questions for the evaluation of palatability on the Likert Survey. The order of questions corresponded to that used in the trials.

Results

Analysis of general performace across trials on the pre- and post-tests

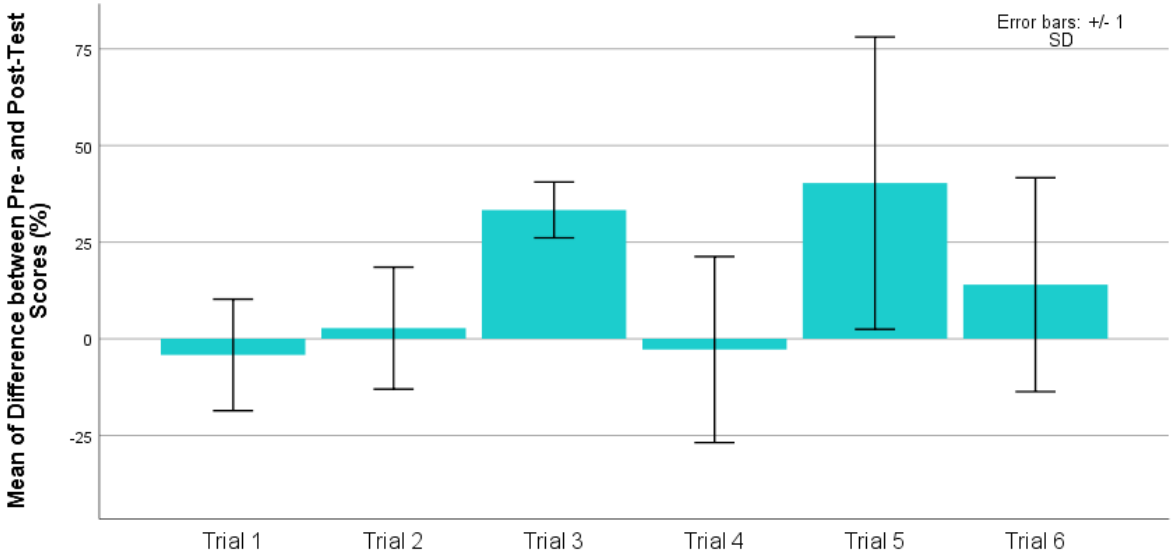


Figure 3. Comparison of the mean of (Post-test – pre-test) scores across all trials with error bars for 1 standard deviation included. Trial 1- ***Normal, tutor-centric****, Trial 2-* ***Normal, learner-centric****, Trial 3-* ***Told tutor-centric****, Trial 4-* ***Told learner-centric****, Trial 5-* ***Told learner-centric, but material is tutor-centric****, Trial 6-* ***Told tutor-centric, but material is learner-centric.***

The mean of the difference between post- and pre-test scores were calculated for all members in each trial, with the exception of trial 1 where one test participant scored full marks in the pre-test (Figure 3). In this situation, there is no possible learning gain. While only trials 1 and 4 showed mean negative difference scores, the mean difference for trial 1 is more negative than trial 4 (negative learning gain). It is possible to do worse in the post-test as the pre-test questions test only knowledge on theory, while the post-test questions evaluate the ability to apply the knowledge (if it has been successfully learnt). This ensured that learners do not pass by simply regurgitating facts from memory (which is already a skill finely honed in the Asian academic setting).

Trials 2, 3, 5 and 6 showed positive mean differences (positive learning gain), with the highest positive mean difference observed in trial 5 at approximately 37.5%, followed closely by trial 3 at approximately 30%. Trial 2 experienced the lowest positive mean difference close to 0.

In this study, we looked at learning gains solely as the distribution of absolute difference between post-test and pre-test (delta) amongst individuals. No normalisation of the results was done since most participants performed similarly on the pre-test anyway and there was little variance in basal capability among individuals within and across trials.

Analysis of performance on one specific learning objective across trials

The net improvement score for each trial as a whole and the performance for individual participants on the pre- and post-tests matched to learning objective 5: 'Interpreting the Receiver Operator Curve (ROC) curve' are presented in Figure 4. Performance on this learning objective was chosen for analysis as it had the highest percentage of people choosing 'I do not know this concept' out of all the pre-test questions at 50% of the entire cohort across all trials.

Tutor-centric: Not told (trial 1) and Perceived as (Trials 3 and 5)

Trial 1 (not told, tutor-centric)				Net improvement: +/- 0%
Student	A	B	C	D
Pre-test (qn 6)	X	✓	N/A	X
Post-test (qn 8)	X	✓	X	X

Trial 3 (told tutor-centric)				Net improvement: +25%
Student	A	B	C	D
Pre-test (qn 6)	N/A	N/A	N/A	✓
Post-test (qn 8)	X	X	✓	✓

Trial 5 (student-centric, but told tutor-centric)				Net improvement: -25%
Student	A	B	C	D
Pre-test (qn 6)	N/A	✓	N/A	✓
Post-test (qn 8)	X	X	✓	X

Student-centric: Not told (trial 2) and Perceived as (Trials 4 and 6)

Trial 2 (not told, student-centric)				Net improvement: +50%
Student	A	B	C	D
Pre-test (qn 6)	N/A	N/A	N/A	N/A
Post-test (qn 8)	X	X	✓	✓

Trial 4 (told student-centric)				Net improvement: +/- 0%
Student	A	B	C	D
Pre-test (qn 6)	✓	X	N/A	✓
Post-test (qn 8)	X	X	✓	✓

Trial 6 (tutor-centric, but told student-centric)				Net improvement: -25%
Student	A	B	C	D
Pre-test (qn 6)	N/A	✓	✓	✓
Post-test (qn 8)	X	X	✓	✓

Figure 4. Comparison of individual test participants' pre- and post- test results matched ***to learning objective 5: Purpose of ROC curve*** *across trials. Results are separated into two columns:* ***tutor-centric (perceived and not told)- left, student/learner-centric (perceived and not told)-right****. The grey boxes represent wrong answers that are marked with 'X' or when students chose 'I do not know this concept', marked as 'N/A' and the green boxes represent correct answers, marked with '✓'.*

The net improvement score was derived by summing up the positive improvement and negative improvement scores of all members in each group. Trial 2's net improvement score of 50% was the highest while trials 5 and 6 had the lowest score at -25%. The learner/student-centric category had a higher total net improvement score of 25% combined compared to tutor-centric at 0%. Additionally, the percentage of people who improved from an 'N/A' score (which represented 'I do not know this concept') for trials 2 and 4 combined was higher at 60% compared to the combined percentage of 25% for trials 1 and 2.

However, there was an increase in net improvement from 0% when test participants were unaware material was tutor-centric (trial 1) to when they were aware (trial 3) at 25% with no negative improvement observed in both trials. In contrast, net improvement decreased from 50% to 0% between when they were unaware material was learner-centric (trial 2) and when they were told (trial 4). Although there was a positive improvement of 25% by student C of trial 4, it was cancelled out by student A's negative improvement in the calculation of net improvement.

Qualitative comparison of palatability towards both teaching materials

Figures 5 and 6 showed the proportion of test participants per group who chose from the 5 options available on the palatability questionnaire (Table 5) for each question. All participants' responses were included in this qualitative analysis. None of them chose 'Strongly disagree' for any of the questions. The total score on the Likert scale for each group was calculated by summing up the scores of the group members corresponding to the option they chose.

In situations where participants were unaware of who produced the teaching material, the responses were slightly more positive for the tutor-centric group (trial 1) than the learner-centric group (trial 2) across all questions. For questions 1 to 3, trial 1 scored

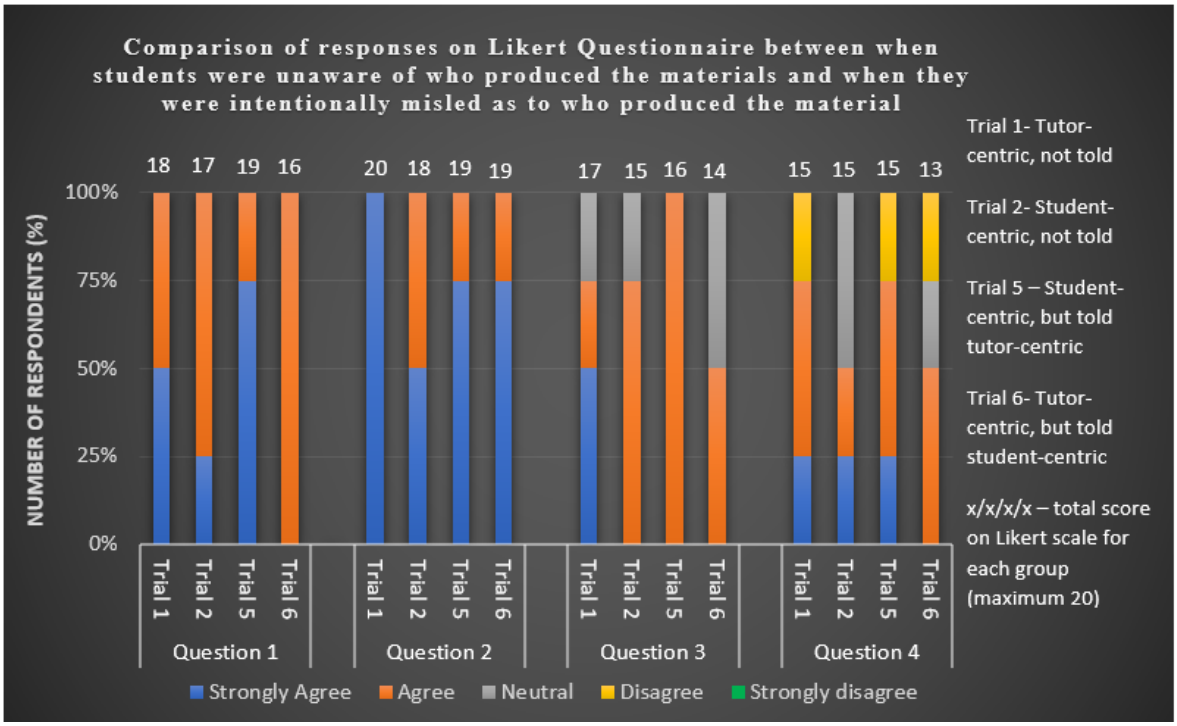
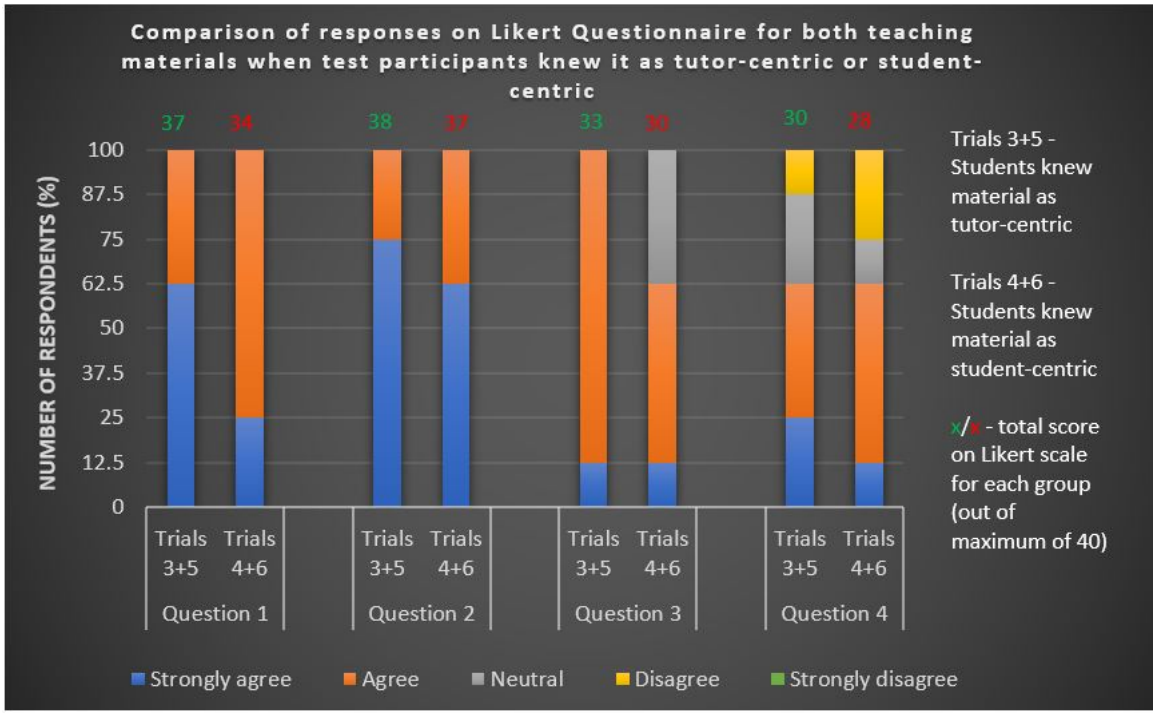


Figure 5. Palatability scores across trials in which students were unaware of who produced the material and when they were intentionally misled. Tutor-centric material-perceived or otherwise- scored higher than learner-centric material in all questions except question 4, in which trials 1,2 and 5 got the same score.

higher on the Likert scale than trial 2. Additionally, a higher proportion of respondents in trial 1 chose 'Strongly agree' than 'Agree' whereas in trial 2, 'Agree' made up the bulk of their responses. In question 3 where both trials had the same proportion of people choosing 'Neutral' at 25%, trial 1 was still comparatively more positive as 50% of the participants chose 'Strongly agree' while none chose that option in trial 2. For question 4 however, although the score on the Likert scale was the same for both groups at 15, trial 1 still showed a more positive response with a median score of '4' as opposed to '3.5' in trial 2.

The comparison between when participants were not told who produced the teaching material (Trials 1 and 2) and when they were intentionally misled as to who produced the teaching material (Trials 5 and 6) shows 2 obvious results. Firstly, responses were more positive when test participants were told the teaching material was tutor-centric even when it was really learner-centric (trial 5) as compared to when the opposite is true (trial 6) for all questions except question 2.

Secondly, barring question 2, trial 5 was more similar to trial 1 than trial 2 and trial 6 was more similar to trial 2 than trial 1 when comparing across the two categories. For question 4, trial 5's median score of 4 was identical to trial 1 even though trials 1, 2 and 5 had the same score of 15. In contrast, trial 6 scored lower at 13 yet its median score was the same as trial 2's at 3.5.



38 37 33 30 30 28
34

Figure 6. Palatability scores between combined trials of when students knew material as tutor-centric or learner/student-centric, irrespective of whether that was true. Perceived tutor-centric scored higher than learner-centric in all questions.

Similarly, responses were also more positive across all questions when test participants were told teaching material was tutor-centric as opposed to when they were told it was learner-centric, regardless of whether they were truly tutor-centric or learner-centric respectively. For questions 1 and 2 where participants only chose either 'Strongly agree' or 'Agree' for both clusters, more respondents chose "Strongly agree" in the cluster which knew material as tutor-centric (trials 3 and 5) than those who knew material as learner-centric (trials 4 and 6).

Discussion

Students showed a much higher improvement on post-test when told teaching material was tutor-centric as opposed to when told learner-centric

In the first two trials conducted, little difference in improvement was observed when test participants were unaware of who produced what between when tutor-centric material was fielded (trial 1) and when learner-centric material was fielded (trial 2) (Figure 6).

However, students benefitted from knowing that their teaching material was tutor-centric, with an approximate five-fold increase in improvement when they knew material as tutor-centric, irrespective of whether that was indeed and truly the case (trials 3 and 5), compared to when they were not told.

A better performance on their tests after knowing the teaching material they were going to be exposed to was tutor-centric could be attributed to the high regard for teachers that most students have in a predominantly Chinese Confucian-based society such as Singapore (Ee and Tan, 2008). Making participants conscious about the fact that the material was tutor-centric right before the start of the trial could have emphasized the credibility of said teaching material to them or made them aware that the material had the possibility of not being tutor-centric, thus making them more appreciative that it was actually tutor-centric.

In contrast, no observable difference in improvement was made between when students were told teaching material was student-centric (trial 4) and when they were not told (trial 2). Although a larger improvement was seen in trial 6 when students were fielded tutor-centric material but told it was student-centric compared to trials 2 and 4, it was still notably lower than that in either trials when students were told material was tutor-centric. The results of trial 4 were surprising given that SAP was reportedly successful in not just increasing deeper learning in student consultants involved in the co-creation process (Carey, 2013), but some cases also demonstrated how student contribution to work fielded to fellow students was well-received (Dunn et al., 2018). We expected a deeper engagement with the lecture material in the test participants after knowing it was student-centric that may be reflected by a higher improvement in scores compared to when they were not told, which was not the case in this study. One reason could be that because the Asian culture emphasises hierarchy in relationships and respect towards authority, people who are seen as equal in standing may not be afforded the same level of regard compared to those in higher

positions. Consequently and unfortunately, they might not be receptive towards learning from a material prepared by "SBS students like them but trained by a Computer Science professor".

However, as trial 6's results conflicted with trial 4, it cannot be inferred that telling students material was student-centric was ineffective. At the same time, because two confounding variables were possibly introduced in trial 6, the increase in improvement as compared to trial 2 cannot be related clearly to either being told learner-centric or if it was due to the material being tutor-centric.

All trials, with the exception of trials 1 and 4, followed the expected notion that any intervention in learning which serves to educate (be it in the form of a lecture, video presentation or reading for example) will lead to positive learning that may be different for each individual (Conn, 2017). In both groups, most of them actually did worse in the post-test after exposure to the teaching material as compared to the pre-test.

Key reasons for such a disparity in both trials could be due to disengagement from the lecture material or a larger practice-theory gap in the students of these trials compared to the other trials.

Students found tutor-centric material to be more palatable than student-centric, be it perceived or real

Palatability refers to the general perception of agreeability the test participants have towards the teaching materials that reveals the magnitude of their feelings regarding a certain aspect of the material.

When students were unaware of who produced what, they rated the tutor-centric material slightly more positively than the learner-centric one in various ways (Figure 5). However, this does not mean that the tutor-centric material was truly superior to the learner-centric version in all these categories. The palatability survey only revealed the test participants' subjective perception of the teaching material, which did not account for the effectiveness of the material. Furthermore, the test participants rated the tutor-centric material less favorably when told it was student-centric (trial 6), suggesting that general notions of palatability are easily manipulated.

Similarly, students found teaching material to be more palatable when told it was tutor-centric than learner-centric, regardless of whether that was truly the case (Figure 6). This indicated that students generally preferred tutor-centric teaching material over learner-centric and this bias may be one reason accounting for their better improvement on post-test scores when told tutor-centric compared to when told student-centric.

Learner-centric material was more effective than tutor-centric material in educating students on learning objective 5, more so when students were unaware of it being student-centric

Question 6 on the pre-test relating to the concept of ROC curve was the most unfamiliar to test participants based on results analysis and feedback to the facilitator about the questions after the trials (Figure 4).

Interestingly, more people improved on the corresponding post-test question from an N/A score in the pre-test in the student-centric trials combined (2 and 4) compared to those in the tutor-centric trials (1 and 3) at a ratio of 2.4:1. However, while telling students the material was tutor-centric increased the net improvement rate by 25% compared to when not told with only positive improvement made, telling students material was learner-centric actually decreased the net improvement with a negative improvement accounting for the decrease.

Although learner-centric material was more effective than tutor-centric material in bridging the gap in knowledge regarding ROC curve in test participants, the effectiveness was diminished by telling them it was learner-centric. On the other hand, students did better when told tutor-centric compared to not told, which agreed with the findings on the overall improvement.

While the content of both teaching materials regarding the ROC curve was similar,there were some differences in the presentation of information. For example, the learner-centric material had a larger picture of the curve, and equations were separate from the curve itself. Such presentation could have appealed more to the participants and facilitated learning of the concept better. This thus demonstrated the usefulness of using student input in co-designing teaching material on the basis that students understand better how other students would like to learn on account of them being in the same position themselves compared to professors (Healey et al., 2016).

Limitations and recommendations for future study

The test design did not account for the theory-practice gap which varied uncontrollably amongst individuals (i.e., the post-test does not reward superficial learners). Thus, people with a wider theory-practice gap would naturally be at a disadvantage. However, weak application skills do not necessarily translate to a failure to learn as successful learning may also manifest in other forms (aside from test performance). Therefore, the post-test results could be limited in the interpretation of deeper learning since they only captured one form of successful learning. Future studies may include surveys to determine participants' learning styles and tailoring the evaluation of successful learning accordingly.

Additionally, while the study revealed students perceived tutor-centric material more favourably than learner-centric, that their perception of a material may change when told who produced it and that their performance can be affected by whether they knew the teaching material as tutor-centric or learner-centric, it was limited in showing the extent of influence of perception over the individual's performance or how it affects it. Future studies may be done to investigate the extent of correlation between both factors by including a post-hoc analysis of the students' feelings and though processes

during the trial in future studies (a list of potential questions are shown in Appendix C).

Conclusion

Students generally found tutor-centric material more appealing than learner-centric and telling them it was tutor-centric clearly improves their test performance. However, learner-centric material did have merit and was even more useful than tutor-centric material in teaching test participants something which they previously knew nothing about (e.g. the section on ROC curves), provided they were not told it was in fact, learner-centric.

The placebo effect of telling students a teaching resource was tutor-centric was consistent and much stronger than that of telling them it was learner-centric. While students found learner-centric material generally palatable, there was an inherent disconnect in them when faced with material perceived to be learner-centric which may be responsible for the poorer test performance observed.

While it seems intuitive to suggest incorporating SAP in terms of co-creation between student consultants and professors, but not to tell the general student population that students contributed to the production of material to get the best of both worlds, doing so violates the ethos of SAP itself. A partnership refers to equal contribution from both parties (though not necessarily in the same ways) and equal amount of respect (and recognition) should be accorded to both of them. Additionally, it is unethical to deceive the student population and it may become counter-productive when the student consultants feel resentful and discouraged.

Though SAP is gaining traction as a viable pedagogical method to enhance learning among higher education students in Western societies, there is limited information available about it applied in the Asian context. The differences between learners from the Western and Asian societies should be considered during the implementation of SAP and customised for these respective contexts. It may be prudent to introduce SAP slowly and in ways which students can visualise how it is done. We believe that SAP can work, but it needs to be introduced gradually and purposefully.

Acknowledgements

WWBG and CCS gratefully acknowledge the EdeX and ACE Grants from Nanyang Technological University, Singapore.

References

Bovill, C., Cook-Sather, A., Felten, P., Millard, L., & Moore-Cherry, N. (2015). Addressing potential challenges in co-creating learning and teaching: Overcoming resistance, navigating institutional norms and ensuring inclusivity in student–staff partnerships. *Higher Education,71*(2), 195-208. doi:10.1007/s10734-015-9896-4

Carey, P. (2013). Student as co-producer in a marketised higher education system: A case study of students' experience of participation in curriculum design. *Innovations in Education and Teaching International,50*(3), 250-260. doi:10.1080/14703297.2013.796714

Conn, K. M. (2017). Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-Analysis of Impact Evaluations. *Review of Educational Research,87*(5), 863-898. doi:10.3102/0034654317712025

Cook-Sather, A. (2013). Student-faculty partnership in explorations of pedagogical practice: A threshold concept in academic development. *International Journal for Academic Development,19*(3), 186-198. doi:10.1080/1360144x.2013.805694
Cook-Sather, A., Bovill, C., & Felten, P. (2014). *Engaging students as partners in learning and teaching a guide for faculty*. San Francisco: Jossey-Bass.

Dunn, M., Loch, B. & Scott, W. (2018) The effectiveness of resources created by students as partners in explaining the relevance of mathematics in engineering education, *International Journal of Mathematical Education in Science and Technology*, 49:1, 31-45, DOI: 10.1080/0020739X.2017.1338771

Ee, J., & Tan, O. (2008). Cultural influences of the East & West: Implications on the educational reforms in the Singapore context. *KJEP,5*(1), 49-62.

Fang, Y., & Gopinathan, S. (2009). Teachers and Teaching in Eastern and Western Schools: A Critical Review of Cross-Cultural Comparative Studies. *International Handbook of Research on Teachers and Teaching,*557-572. doi:10.1007/978-0-387-73317-3_36

Healey, M., Flint, A., & Harrington, K. (2014). Engagement through partnership: Students as partners in learning and teaching in higher education. *York: Higher Education Academy*.

Healey, M., Flint, A., & Harrington, K. (2016). Students as Partners: Reflections on a Conceptual Model. *Teaching & Learning Inquiry: The ISSOTL Journal,4*(2). doi:10.20343/10.20343/teachlearninqu.4.2.3

Lubicz-Nawrocka, T. M. (2018). Students as partners in learning and teaching: The benefits of co-creation of the curriculum. *International Journal for Students as Partners,2*(1), 47-63. doi:10.15173/ijsap.v2i1.3207

Shor, I. (2014). *Empowering education: Critical teaching for social change.* Chicago: The University of Chicago Press.

Contact email: wilsongoh@ntu.edu.sg

Appendix A

List of questions in Pre-test for the topic 'Machine Learning'

Instructions to test participants

Thank you for taking part in our study, your responses are important to us. There are 5 single-answer MCQ questions in this pre-test. Please answer to your best abilities and where a concept is unfamiliar to you, please fill in "I do not know this concept".

1. In statistical testing and machine learning, what does a "true negative" mean to you?

    a. Happens with a positive prediction, when in fact it is negative
    b. Happens with a negative prediction, when in fact it is positive
    c. Happens with a positive prediction and it is in fact, positive
    d. Happens with a negative prediction, and it is in fact, negative
    e. I do not know this concept.

2. In statistical testing and machine learning, what does a "False negative" mean to you?

    a. Happens with a positive prediction, when in fact it is negative
    b. Happens with a negative prediction, when in fact it is positive
    c. Happens with a positive prediction and it is in fact, positive
    d. Happens with a negative prediction, and it is in fact, negative
    e. I do not know this concept.

3. Knowledge Discovery is the process of extracting useful information from data. Using your own intuition, list the order of the steps which you find to be most reasonable.

    a. Data gathering> Variable selection> Variable generation> Variable integration
    b. Data gathering> Variable integration> Variable selection> Variable generation
    c. Variable selection> Variable generation> Data gathering> Variable integration
    d. Data gathering> Variable generation> Variable selection> Variable integration
    e. I do not know this concept.

    Glossary*: **Variable integration** is to use selected information to make a conclusion about something, **variable selection** is to determine the correct variable, **variable generation** is to classify information into specific categories, **data gathering** is to obtain all available information
    *arranged in a random order

4. Which of the following best describes the relationship among variables, attributes and traits?

    a. An attribute is a variable

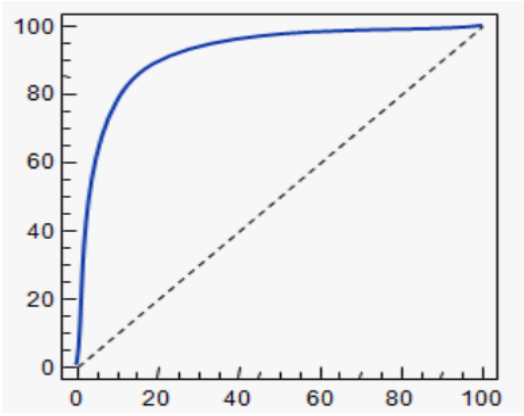
b. A trait is a variable that distinguishes one thing from another
c. An example of a variable is “Shape” while an example of an attribute is “circle”
d. B and C only
e. I do not know this concept

5. Statistical prediction is a form of statistical inference. Which of the following are conditions of a good prediction?

I. High sensitivity
II. High specificity
III. High false positive rate
IV. Low true positive rate

a. I and II only
b. I, II, III only
c. III and IV only
d. II only
e. I do not know this concept

6. The Receiver Operator Curve (ROC) is shown as below with its X and Y axes removed.



ROC curve present?

What information does the

a. Degree of specificity of the test
b. Degree of sensitivity of the test
c. The relationship between sensitivity and specificity of the test for every possible cut-off
d. None of the above
e. I do not know this concept

Appendix B

List of questions in Post-test for the topic 'Machine Learning'

Read the scenarios given carefully and tick the box that you think correctly describes the situation. You may tick 'X' if you do not know the answer. **Each box may be ticked once, more than once, or not at all.**

Scenario	Outcome				
	TP	TN	FP	FN	X
1. A pregnancy test is positive when you are not pregnant					
2. Detective Jake Peralta arrested a criminal suspect who was indeed the ringleader in a diamond heist					
3. The surveillance system of the jewellery store where the diamond heist took place did not detect malicious activities during the heist					
4. The saleswoman told the rude customer that there were no more new pieces of the blouse displayed on the mannequin when there were actually a lot of it stored in the storeroom.					

TP- true positive

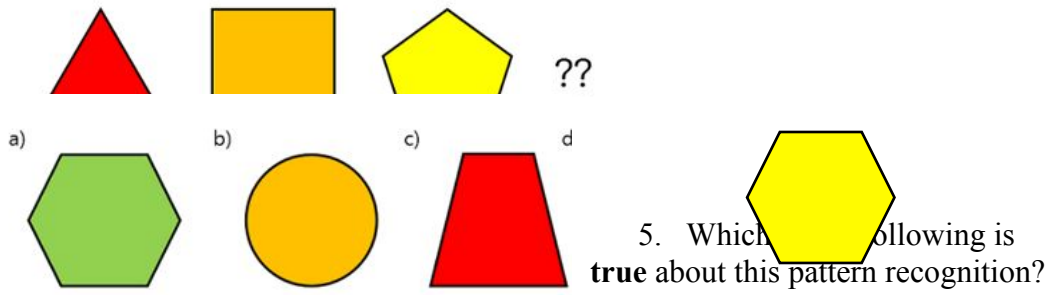
TN- true negative

FP- false positive

FN- false negative

X- I do not know the concept

Question 5 is based on the diagram shown below.



5. Which of the following is **true** about this pattern recognition?

a. 2 traits are used to determine which comes next in the pattern
b. 'Shape' is an example of an attribute while 'Hexagon' is an example of a variable
c. Both a and b
d. None of the above
e. I do not know this concept

6. You were an eye-witness in the diamond heist case and was called in to identify the culprit from a line-up as shown below:

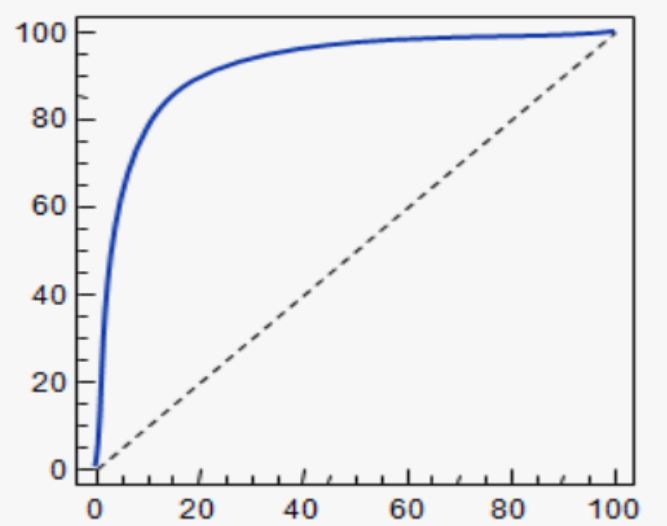
Using what you have learnt about the process of knowledge discovery, identify the appropriate heading for each classification in the table below.

I	II	III	IV
Suspect 3 seems to fit the profile of the culprit you saw that day	From memory, the culprit was a very tall, white male, bald, had a nose ring and a rose tattoo on his left forearm	You looked specifically at the men' height (about 185cm tall), hairstyle (bald), race (Caucasian),	Height, race, facial accessories, body modifications

a. I-Variable integration, II- Variable selection, III- Data gathering, IV-Variable generation
b. I- Variable integration, II-Variable selection, III-Variable generation, IV-Data gathering
c. I- Variable integration, II-Data gathering, III-Variable selection, IV-Variable generation
d. I- Variable selection, II-Data gathering, III-Variable integration, IV-Variable generation
e. I do not know this concept

7. Which of the following is true about 'Sensitivity' in statistics?
   a. Sensitivity is otherwise known as the accuracy of the test
   b. It measures how well the prediction captures relevant results
   c. 'False negatives' are not considered in the calculation of Sensitivity
   d. I do not know this concept

8. The picture below shows a ROC curve with its X and Y axes removed.



Which of the following statements is/are **false** about the ROC curve?

I. The Y axis is labelled 'Sensitivity' and the X axis is labelled 'Specificity'
II. The top left corner of the ROC curve is the point where the results are most sensitive and specific
III. The diagonal line represents results that were obtained by chance
IV. Sensitivity and specificity have a directly proportional relationship i.e. as sensitivity increases, so does specificity

a. I only
b. I, III and IV
c. I and IV
d. I, II and III
e. I do not know this concept

Appendix C

Potential List of questions for palatability survey

On a scale from 1 to 5 (1 being "Strongly Disagree" to 5 being "Strongly Agree"), please choose the option that corresponds most closely to how you feel about the following statements:

1i) The questions on the pre-test were challenging to answer
ii) Describe one or more particularly memorable question to you from the pre-test and elaborate more on why you say so.
iii) I feel confident that I did well on the post-test compared to the pre-test

2i) I was fully engaged during the entirety of the lecture
ii) Think back on the lecture. Which part of it did you think was confusing for you or failed to engage you enough? Please elaborate and give as many details as possible.
iii) The lecture prepared me well enough for the questions on the post-test

3i) Rank in order of importance to you the factors **(1 being least important, to 5 being most important)** that optimize your learning in SBS.
Interest in the subject, having ample time to learn the content, quality of the teaching material, having trained teaching assistants available for assistance, technology-enabled learning (use of clickers, smartboard etc)
3ii) Describe the ideal teaching material you would appreciate to be made available to you.

4) If NTU professors were to collaborate with students that have already taken the module in designing learning material, which type of material would you want to learn from the most?
a) Student consultants produced material only
b) Student and professors produced material
c) Professor produced material only
d) Please explain your choice in terms of why your chosen option appealed to you most and not the other two.

Question 5 and its subparts are tailored to the different trials (T1, T2, T3 etc) according to the unique scenario involved in each one.

(T1,2)

5a) How surprised are you to know that the lecture material was produced by a professor/ student consultants?

- **Extremely surprised, quite surprised, moderately surprised, slightly surprised, not at all surprised**

5b) Please elaborate on your answer in terms of what you experienced vs what you expected from a professor-produced/student-produced material (content, usefulness of analogies etc).

(T3,4)

5a) Knowing my material was student-produced/professor-produced made me look forward to it.

5b) Please elaborate on your answer, giving as many details about how you perceived the material as you can.

5c) How much more excited will you be if the material presented to you was student-produced/professor-produced instead?

- **Very excited, quite excited, moderately excited, a little excited, not at all excited**

(T5,6)

5a) How surprised are you to know that the lecture material was produced by a professor/ student consultants instead of by a professor/ student consultant as mentioned to you during the trial itself?

- **Extremely surprised, quite surprised, moderately surprised, slightly surprised, not at all surprised**

5b) What are your expectations of a professor-produced material (for those fielded faculty but told student). Did the material live up to your expectations? Please elaborate in terms of content, relatability, analogies used for etc.