THE ASIA-PACIFIC SCHOLAR
MEDICAL AND HEALTH PROFESSIONS EDUCATION

ORIGINAL ARTICLE

Check for updates

Insights for medical education: via a mathematical modelling of gamification

De Zhang Lee[1], Jia Yi Choo[1], Li Shia Ng[2], Chandrika Muthukrishnan[1] & Eng Tat Ang[1]

[1]Department of Anatomy, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [2]Department of Otolaryngology, National University Hospital, Singapore

Abstract

Introduction: Gamification has been shown to improve academic gains, but the mechanism remains elusive. We aim to understand how psychological constructs interact, and influence medical education using mathematical modelling.

Methods: Studying a group of medical students (n=100; average age: 20) over a period of 4 years with the Personal Responsibility Orientation to Self-Direction in Learning Scale (PRO-SDLS) survey. Statistical tests (Paired t-test) and models (logistic regression) were used to decipher the changes within these psychometric constructs (Motivation, Control, Self-efficacy & Initiative), with gamification as a tool. Students were encouraged to partake in a maze (10 stations) that challenged them to answer anatomical questions using potted human specimens.

Results: We found that the combinatorial effects of the maze and Script Concordance Test (SCT) resulted in a significant improvement for "Self-Efficacy" and "Initiative" (p<0.05). However, the "Motivation" construct was not improved significantly with the maze alone (p<0.05). Interestingly, the "Control" construct was eroded in students not exposed to gamification (p<0.05). All these findings were supported by key qualitative comments such as "helpful", "fun" and "knowledge gap" by the participants (self-awareness of their thought processes). Students found gamification reinvigorating and useful in their learning of clinical anatomy.

Conclusion: Gamification could influence some psychometric constructs for medical education, and by extension, the metacognition of the students. This was supported by the improvements shown in the SCT results. It is therefore proposed that gamification be further promoted in medical education. In fact, its usage should be more universal in education.

Keywords: *Psychometric Constructs, Medical Education, Motivation, Initiative, Self-efficacy*

Practice Highlights
- Student's enjoyment (interest) of the curriculum will determine the eventual academic outcome.
- Metacognition (defined as the "learning of learning", "knowing of knowing" and/ or the awareness of one's thought processes) was improved with SCT and gamification.
- Gamification is useful as a form of augmentation for didactic teaching but should never replace it.
- Different type of psychometric scale (e.g. LASSI versus PRO-SDLS) used in your research will produce varying results.
- Gamification is resource intensive and needs extra time to prepare compared to didactic approaches.

I. INTRODUCTION

Psychology is integral to healthcare and education but has often been overshadowed, compared to the other basic disciplines (Choudhry et al., 2019; Pickren, 2007). This is ironical because human psyche needs to be properly understood in order to manage them effectively (Wisniewski & Tishelman, 2019). Presently, the study of psychology does not feature prominently in the medical curriculum (Gallagher et al., 2015) with the exception of psychiatry (Douw et al., 2019). This gap needs to be addressed (Paros & Tilburt, 2018). In this research, we seek to understand the constructs for good medical

learning via gamification which has wide ranging effects (Mullikin et al., 2019). The psychometric constructs to be analysed were as follows: 1) "Motivation"; define as the desire to learn out of interest or enjoyment (Yue et al., 2019). 2) "Initiative"; refers to how proactive a student is to learning (Boyatzis et al., 2000). 3) "Control"; is how much influence one has over the circumstances (Sheikhnezhad Fard & Trappenberg, 2019). 4) "Self-Efficacy"; relates to how confident one is, to do what needs to be done (Michael et al., 2019). We believe that these constructs contribute to the student's awareness of their own thought processes (metacognition) towards their medical education.

Gamification" is defined as a process of adding game-like elements to something so as to encourage more participation (Rutledge et al., 2018; Van Nuland et al., 2015). The idea of using games to "lighten up" medical education in the clinical setting was first proposed in 2002 (Howarth-Hockey & Stride, 2002). The authors observed increased engagement and participation during lunchtime medical quizzes in the hospital. They therefore concluded that medical education could be fun, and since then, gamification has been taken seriously by the community (Evans et al., 2015; Nevin et al., 2014). In essence, gamification could be something as simple as having board games (Ang et al., 2018) but importantly, its impact on students' learning must be evaluated and validated. Most studies in the literature did not fulfil this requirement (Graafland et al., 2012). The impact of games on the behavioral and/or psychological outcomes should be studied (Graafland et al., 2017; Graafland et al., 2014).

A PubMed search would reveal that there are numerous self-reporting tools such as LASSI (Learning and Strategies Study Inventory (Muis et al., 2007), MSLQ (Motivated Strategies for Learning Questionnaire (Villavicencio & Bernardo, 2013), and the SRLPS (Self-regulated Learning Perception Scale) (Turan et al., 2009) etc. Given the choices, how does one decide which one to adopt for their studies? In our research, we chose to use the PRO-SDLS survey questions with some modifications. The choice was both serendipitous and practical, as we have previously validated it via the Cronbach alpha (>0.7). In our earlier work, feedback scores and results yielded inconclusive evidence to support enhanced motivation among our students. Furthermore, was this due to gamification? With the current endeavour, we aim to prove via mathematical modelling that there are indeed alterations to the psychometric constructs. Hence, we re-analyse the old data set together with additional new information, using

statistical analysis tools such as the logistic regression model, Wilcoxon tests, and the Paired t-test.

Medical teaching and learning is a complex endeavour based on an apprenticeship model (Cortez et al., 2019), which may or may not be an ideal arrangement (Sheehan et al., 2010). Furthermore, the decision making is often delegated to the seniors (Chessare, 1998). Conversely, gamification could empower the students to take charge of one's learning, including decision making (Shah et al., 2013). Furthermore, one needs to understand what works from what is empirical (Cote et al., 2017). While our initial research addressed the impact of the games on academic performance, we now sought to further understand its effects on the psychometric dimensions. This will help to understand the psychology of self-directed (or regulated) learning. We hypothesize that the amount of gamification will impact these constructs. In summary, we hope to achieve the following:
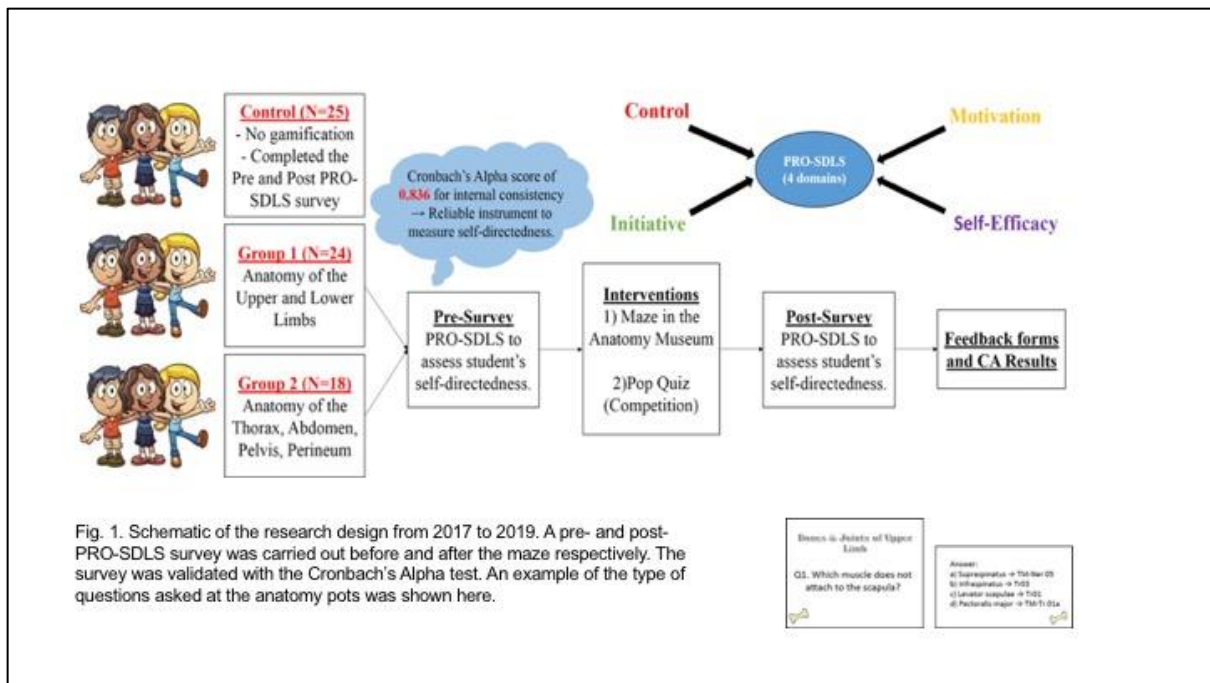
Aims:
• Understanding the role of psychometric constructs and gamification in medical education via suitable mathematical modelling.
• To decipher the interaction of different psychometric constructs (Motivation, Self-efficacy, Control and Initiatives) in producing desired learners' behaviours (metacognition) via the anatomy maze.

II. METHODS

First-year medical students (M1) took part in this retrospective analytical research. Two randomised groups of medical students (n=75, median age: 20 years) consented to the study (Group 1 & 2). A randomised group of students (n=25) exposed to no gamifications (Group 0) served as the control. Every student was required to complete a pre- and post- PRO-SDLS for the research. There were no penalties for withdrawing from the IRB-approved project (See IRB: B-16-205).

Gamification was carried out according to the scheme in Figure 1. Each group was divided into 3 to 4 subgroups that would enter the maze with a clue card (see example in Figure 1) linked to a specific pot specimen. They were required to explore the museum for the next clue and had to answer the hidden questions (see examples in Appendix) which would provide further directions. At the conclusion, students were given a competitive pop quiz that had no impact on their summative academic grades.

Fig. 1. Schematic of the research design from 2017 to 2019. A pre- and post-PRO-SDLS survey was carried out before and after the maze respectively. The survey was validated with the Cronbach's Alpha test. An example of the type of questions asked at the anatomy pots was shown here.

The main purpose was to assess formative knowledge acquisition. The validated PRO-SDLS include the following psychometric constructs: "Motivation" (7 questions), "Initiative" (6 questions), "Control" (6 questions) and "Self-Efficacy" (6 questions). (See Sup. Materials). The responses are then collapsed into an average accordingly. A higher score indicated more agreeability towards that construct for self-directed learning (Ang et al., 2017). The survey was designed with backward scoring to ensure accuracy. For quantification purposes, we subtracted the pre-feedback from the post-feedback scores for each question. An increased score for a particular construct suggests improvement (Cazan & Schiopca, 2014). Furthermore, students in Group 2 were given Script Concordance Test (SCT) quizzes (See Sup. Materials) as part of gamification (Lubarsky et al., 2013; Lubarsky et al., 2018; Wan et al., 2018). SCT were meant to enhance clinical reasoning. All data were analysed from two perspectives:

(a) The magnitude of score increase (or decrease) of the post- PRO-SDLS survey responses, with respect to the pre- responses.

(b) The odds of a student reporting an increased score in the post- PRO-SDLS survey responses.

In (a), the paired differences for each student's response were studied using a parametric approach (paired t-test). In (b), we studied the odds of increased score for each construct, and investigate if grouping affected these odds. More formally, for each construct k (where k is one of the four constructs), we define variable p_{kj} as the probability of a student from group j showing an increase in score for construct k (and hence, $1 - p_{kj}$ as the probability that the student's score decreased or remained unchanged). The value of p_{kj} can be estimated by dividing the number of students from group j with an increased score for construct k by the total number of students from group j. If the interventions are unsuccessful, we would expect p_{kj} to be around 0.5 since a student's score would likely either increase or decrease at random, with an equal probability. This can be tested using the t-test.

An alternative approach would be to study the odds of success, which can be written as $\frac{p_{kj}}{1 - p_{kj}}$. A common mathematical model used to study these odds is the logistic regression model. For each construct, the logistic regression model studies the odds of a student from a given group showing an increased or decreased score. The overall significance of the model can be tested using the p-value obtained from the likelihood ratio test, while the significance of the individual odds can be tested using the t-test. For more details on the logistic regression model, we refer the reader to (Agresti, 2003).

We utilised the open source software R (Team, 2019) to perform our statistical analysis.

III. RESULTS

Participation rate in the gamification endeavour was consistently 90±5%, and there was zero withdrawal from it, accompanied by reported favourable qualitative comments.

A. Studying the Absolute Scores

The average change in scores across all the groups for each construct is given in Table 1. From these scores, we believe that our gamification exercises may have had a positive impact on "Self-Efficacy" and "Initiative". To visualize the spread of responses, we have prepared box plots of the post – pre scores (available in Supplementary Materials).

Groups	Constructs			
	Self-Efficacy	Initiative	Motivation	Control
0	0.07	-0.05	0.11	0.03
1	0.13	0.26	0.05	0.01
2	0.13	0.20	0.12	0.07

Table 1: Average post-pre scores

To determine if the construct scores pre and post intervention were different, we used the paired t-test, under the null hypothesis that there is no change. The p-values obtained are summarized in Table 2.

Groups	Constructs			
	Self-Efficacy	Initiative	Motivation	Control
0	0.46	0.57	0.43	0.79
1	0.07	0.00	0.54	0.86
2	0.01	0.00	0.09	0.14

Table 2: p-values of t-test (to 2 decimal places)

We observe that the null hypothesis of no difference between pre and post intervention levels for all constructs are not rejected (under p=0.05) for the control group. Both tests also failed to show any significant change for the "Control" construct.

There is strong evidence that the classroom interventions employed by Groups 1 and 2 had an impact on "Initiative" levels of students, reflected by the small p-values obtained using both tests. The average increase in "Initiative" scores for students in Groups 1 and 2 are 0.71 and 0.67, respectively, which are similar. Recall that the students in Group 2 participated in the SCT, in addition to the maze which is common across both groups. This suggests that the SCT has a negligible impact on "Initiative".

There is also strong evidence (p=0.05) to show that the games enhanced the "Self-Efficacy" levels among the students. The t-test also gives strong evidence (p=0.05) that there is a significant change in Group 2, and milder evidence for Group 1 (p=0.10). The average increase in "Self-Efficacy" levels for Groups 1 and 2 are 0.63 and 0.56, respectively. Again, the differences are negligible, and this suggests that the SCT has a negligible impact. Finally, there is mild evidence (p=0.10) of a significant change in "Motivation" for Group 2, but no such evidence for Group 1. The average increase in "Motivation" score for Group 2 is 0.55. This time, the SCT might have helped to improve students' motivation.

B. Studying the Odds of Score Improvement

We will now turn our attention to modelling the odds of a student reporting an increase in construct score. Earlier, we defined p_{kj} as the probability of a student from group j showing an increase in score for construct k, and explained why we would expect p_{kj} to be around 0.5 if the games have no impact on the odds of "success". The t-test was used to test this, under the null hypothesis that $p_{kj} = 0.5$ for all groups and constructs. The p-values obtained are summarised in Table 3.

Groups	Constructs			
	Self-Efficacy	Initiative	Motivation	Control
0	0.56	0.56	0.85	0.07
1	0.08	0.00	0.78	0.25
2	0.57	0.08	0.57	0.85

Table 3: p-values of t-test (to 2 decimal places)

We first notice that the p-values reported using both tests are almost identical. Interestingly, there is mild evidence (p=0.10) that the p_{kj} value of "Control" construct for Group 0 deviates significantly from 0.5, and it is estimated to be 0.32. This means that the students in the control group reported a drop in "Control" levels.

There is also mild evidence (p=0.10) that the probability of a student reporting an increase in "Self-Efficacy" for Group 1 deviates significantly from 0.5. This probability is estimated to be 0.63, which indicates that the odds of a student from Group 1 reporting an increase in "Self-Efficacy" levels is higher compared to the others.

Finally, there is evidence that the probability of reporting an increase in "Initiative" levels for students from Groups 1 and 2 deviates significantly from 0.5. The probabilities for Group 1 and Group 2 are 0.71 and 0.67, respectively.

Next, we will model our data using the logistic regression model. We will fit four models, one for each construct. For each model, we calculate the odds of a student from a given group showing an increased or decreased score.

An odds of greater than 1 means that the student is more likely to show an increased score, while an odds of less than 1 means the opposite. An odds of exactly 1 means that the student is neither more nor less likely to show a changed score. The statistical significance of the individual odds and the overall model fit for each construct was computed using the t-test, and likelihood ratio test, respectively. The results are summarised in Table 4, with the statistically significant (p=0.10) odds highlighted in blue, together with their respective p-values.

| Odds | Constructs | | | |
	Self-Efficacy	Initiative	Motivation	Control
Group 0	0.92	0.79	1.27	0.47 (0.08)
Group 1	1.08	2.41 (0.02)	1.67	0.72
Group 2	1.25	1.99 (0.10)	1.26	0.93

| | Constructs | | | |
	Self-Efficacy	Initiative	Motivation	Control
p-value	0.93	0.01	0.22	0.29

Table 4: (top) Coefficients for each construct (significant odds in blue, p-values in brackets) (bottom) p-values to assess logistic regression model fit using likelihood ratio test

Under the logistic regression model, not rejecting the null hypothesis for a given odds means that we assume it takes on the value 1. It should be noted that the individual coefficients should be examined when the model is determined to be significant under the likelihood ratio test, as the coefficients obtained under a poor model fit may not be meaningful.

We notice that the significant terms flagged out by the t-test (Table 3) largely agree with the significant terms of the logistic regression model, except for the "Self-Efficacy" odds for Group 1. However, the "Self-Efficacy" model was not determined to be a good fit using the likelihood ratio test.

The only model which was deemed to be a good fit was the one for the "Initiative" construct. The odds for Group 0 is deemed to be insignificant (and hence assumed to be 1), while the odds for Groups 1 and 2 are statistically significant. We can interpret this model as follows,

a. Since the odds for Group 0 is statistically insignificant under the t-test, we assume the odds to be 1. In other words, it is equally likely for a student from the control group to show an increase or decrease in score.

b. The odds for both Group 1 and 2 are statistically significant. The odds of success of Group 1 is 2.41, which can be translated to a roughly 7 in 10 chance (probability of 0.71) for a student in this group showing an improved score. A similar interpretation can be made for Group 2, which showed an odds of 1.99. This translates to a slightly lower probability of 0.67 for a student from Group 2 displaying an improved score.

With this, we have presented a logistic regression approach of mathematically modelling these odds. A search on Google Scholar and PubMed yielded no previous work which made us of this mathematical modelling approach on the PRO-SDLS survey data. With the derived odds, we can compare the degree of success of the various classroom interventions. The logistic regression modelling approach is therefore, proposed as a complement of the t-test approach, which is restricted to detecting the presence of statistically significant differences.

C. Qualitative Comments (underlined words underpinning for metacognition)

1) Positive feedbacks:

• *"The maze games were the most helpful as they helped me to consolidate my learning, and also enables me to ask the tutor any questions that I have from class. They allowed me to learn anatomy in a fun, enjoyable and memorable way"*

• "It allowed me to visualize the things that I was learning and helped with clarifying doubts"

• "The extra question posted at each station was helpful"

• "Wanting to be able to identify things in the museum makes me more motivated to prepare beforehand"

• "Allows me to identify the knowledge gap so that I can work on it"

• "I like the quiz as it motivates me to study beforehand and shows me the gaps in my knowledge"

• "The clinically relevant questions made me think a *lot*"

2) Negative feedbacks:

• "I prefer didactic teaching"

• "We did not interact much with the exhibits"

• "The maze was more of a mini quiz or test to check if we remember anything"

• "Perhaps we could go into more complex concepts"

• "More challenging questions"

• "Students just follow each other around the anatomy museum and it defeats the purpose of the maze"

• "The maze could have a competitive element to make it more exciting. Maybe more MCQ questions per model so we can make use of it more"

IV. DISCUSSION

We undertook this research to decipher how gamification as a concept helps medical students learn a basic subject like human anatomy. We also want to understand how psychometric constructs interact to produce behavioural changes towards self-directed learning. This was done by analysing the data from the PRO-SDLS via statistical tests. Put simply, one needs to understand that medical education is a very complex process that demands balance between apprenticeship (fellowship) (Sheehan et al., 2010), and a dose of self-directed learning (van Houten-Schat et al., 2018). With our initial research into gamification of anatomy education (Ang et al., 2018), there were other studies suggesting similar benefits (Felszeghy et al., 2019; Nicola et al., 2017; Van Nuland et al., 2015). We are therefore convinced that gamification could help to engage students and improve academic gains. However, the notion of gaming can be very broad (Virtual Reality, board games, digital apps etc.), so there is a need to understand the underlying psychology. With that in mind, we re-analyse our previous data with the existing, using proven statistical tools to decipher the learning psychology of these medical students, and their awareness of their own thought processes (metacognition).

We earlier hypothesised that gamification would influence these dependent constructs differently and indeed this was the outcome. In our analysis, we found that the combinatorial effects of the maze and SCT resulted in a significant improvement for "Self-Efficacy" and "Initiative". While the maze alone did not significantly improve "Motivation", we saw mild evidence of an improvement in terms of psychometric scores, when the SCT and maze were used in combination. In lay terms, the maze encouraged these students to learn on their own. By extension, one could also argue that gamification will help the students in making decisions since "Motivation" and "Initiatives" are key attributes (Vohs et al., 2008). The ability to make a simple clinical judgement, and the courage to act on them, are the virtues that we should be imbuing in the medical students, and some junior doctors. Interestingly, there is mild evidence that the "Control" construct was undergoing erosion in the students not exposed to gamification, as the course progresses. This adverse result is not seen in both groups exposed to the games. Perhaps the more relaxed classroom setting with gamification helped students to feel more in control of their learning process. Logically, this made a lot of sense across the education landscape.

A follow up question would be, does the feedback confirm the results given in our qualitative analysis? Recall that in our logistic regression model, students from both non-control groups displayed a statistically significant improvement in "Initiative" levels. This is supported by some of the positive feedback received for our endeavours, such as being "motivated to prepare beforehand", "identify the knowledge gap" and work on them, as well as helping them to "think a lot" about the course content. Furthermore, some of the negative feedback, such as requests for more challenging questions, or more questions in general, suggests that the students are taking the initiative to learn more. This certainly adds credence to the findings of our proposed logistic regression model, as well as highlighting the importance of studying both qualitative and quantitative feedback.

There are caveats that one should be aware when implementing gamification. The formative part of the endeavour could be variable, and dependent on numerous factors such as the tutor involved, and the type of games, interventions, and reporting scales used. In the feedback, 76% of the participants felt that the maze should continue as an adjunct but not to totally replace didactic tutorials. In other words, introducing gaming elements into the curriculum should be done judiciously. With reverse scoring, it was shown that "Self-Efficacy" fell as the level of gamification is increased. In lay terms, students might be feeling that the maze trivialize the learning of the subject. As a counter measure, and to maintain quality assurance, we could introduce video lectures from previous years to allay these fears. In summary, we now confirmed that gamification works, and it influences learning outcomes as demonstrated by others (Burgess et al., 2018; Goyal et al., 2017; Kollei et al., 2017; Kouwenhoven-Pasmooij et al., 2017; Kurtzman et al., 2018; O'Connor et al., 2018; Patel et al., 2017; Savulich et al., 2017). Separately, there were

criticisms as to why SCT was introduced into the research. We believed that such augmentation will add "fun" for the pre-clinical students to tackle the various clinical scenarios and clinical anatomy.

V. LIMITATIONS OF THE STUDY

Our research necessitated that the students take part in the maze and the SCT. Although it was not compulsory, no students opted out of it. Some critics would misconstrue this to be a form of forced play. According to Jane McGonigal, gamification should ideally not be mandated (Roepke et al., 2015).

VI. CONCLUSION

Through statistical modelling, we have shown how the "Initiative", "Motivation", and "Self-Efficacy" constructs could potentially benefit from gamification. The before-after experimental set up allowed for powerful comparisons to be made. Studying the odds of construct score improvement, alongside the raw scores, allowed us to study the data from different perspectives. Though this approach, we discovered how the potential benefits of our gamification exercises outweigh the potential adverse effects. Gamification had resulted in improved "Initiative" in these medical students. We believe that their decision-making skills will also be boosted if existing culture allows for more self-discovery (to improve "Initiative", "Control" and "Self-efficacy") and autonomy. If these recommendations are duly considered and implemented thoughtfully, there is little doubt that our future doctors will be better equipped to serve humanity. This may also help to avoid possible burnout in residents (Hale et al., 2019).

Stronger conclusion and potential for applications are as follows: In a continuum, we started gamifying anatomy education and proven that academic grades could be improved by the process (Ang et al., 2018). We then asked a fundamental question in how exactly it happened. This was done by carrying out a psychometric analysis on the participants. We discovered that psychometric constructs were important, and this was proven in this manuscript. The impact of gamification is now elevated given the COVID-19 pandemic that necessitated more online teaching. Moving forward, we believe that gamification should move towards creating an electronic application that the students may access 24/7. This will ensure that medical teaching will be fortified and be somewhat protected from further disruptions.

Declaration of Interest

All authors have no conflict of interest to declare.

References

Agresti, A. (2003). *Categorical data analysis*. John Wiley & Sons.

Ang, E. T., Abu Talib, S. N., Samarasekera, D., Thong, M., & Charn, T. C. (2017). Using video in medical education: What it takes to succeed. *The Asia Pacific Schola*r. *2*(3), 15-21.

Ang, E. T., Chan, J. M., Gopal, V., & Li Shia, N. (2018). Gamifying anatomy education. *Clinical Anatomy*, 31(7), 997-1005. https://doi.org/10.1002/ca.23249

Boyatzis, R. E., Murphy, A. J., & Wheeler, J. V. (2000). Philosophy as a missing link between values and behaviour. *Psychological Reports*, *86*(1), 47-64. https://doi.org/10.2466/pr0.2000.86.1.47

Burgess, J., Watt, K., Kimble, R. M., & Cameron, C. M. (2018). Combining Technology and Research to Prevent Scald Injuries (the Cool Runnings Intervention): Randomized Controlled Trial. *Journal of Medical Internet Research, 20*(10), e10361. http://doi.org/10.2196/10361

Cazan, A. M., & Schiopca, B. A. (2014). Self-directed learning, personality traits and academic achievement. *Procedia-Social and Behavioral Sciences* (127), 640-644.

Chessare, J. B. (1998). Teaching clinical decision-making to pediatric residents in an era of managed care. *Pediatrics*, *101*(4 Pt 2), 762-766; discussion 766-767. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9544180

Choudhry, F. R., Ming, L. C., Munawar, K., Zaidi, S. T. R., Patel, R. P., Khan, T. M., & Elmer, S. (2019). Health literacy studies conducted in australia: A scoping review. *International Journal of Environmental Research and Public Health*, *16*(7). https://doi.org/10.3390/ijerph16071112

Cortez, A. R., Winer, L. K., Kassam, A. F., Hanseman, D. J., Kuethe, J. W., Quillin, R. C., 3rd, & Potts, J. R., 3rd. (2019). See none, do some, teach none: An analysis of the contemporary operative experience as nonprimary surgeon. *Journal of Surgical Education*, 76(6), e92-e101. https://doi.org/10.1016/j.jsurg.2019.05.007

Cote, L., Rocque, R., & Audetat, M. C. (2017). Content and conceptual frameworks of psychology and social work preceptor feedback related to the educational requests of family medicine residents. *Patient Education and Counseling*, *100*(6), 1194-1202. https://doi.org/10.1016/j.pec.2017.01.012

Douw, L., van Dellen, E., Gouw, A. A., Griffa, A., de Haan, W., van den Heuvel, M., Hillebrand, A., Van Mieghem, P., Nissen, I. A., Otte, W. M., & Reijmer, Y. D. (2019). The road ahead in clinical network neuroscience. *Network Neuroscience*, *3*(4), 969-993. https://doi.org/10.1162/netn_a_00103

Evans, K. H., Daines, W., Tsui, J., Strehlow, M., Maggio, P., & Shieh, L. (2015). Septris: a novel, mobile, online, simulation game that improves sepsis recognition and management. *Academic Medicine*, *90*(2), 180-184. https://doi.org/10.1097/ACM.0000000000000611

Felszeghy, S., Pasonen-Seppänen, S., Koskela, A., Nieminen, P., Härkönen, K., Paldanius, K. M., Gabbouj, S., Ketola, K., Hiltunen, M., Lundin, M., & Haapaniemi, T. (2019). Using online game-based platforms to improve student performance and engagement in histology teaching. *BMC Medical Education*, *19*(1), 273. https://doi.org/10.1186/s12909-019-1701-0

Gallagher, S., Wallace, S., Nathan, Y., & McGrath, D. (2015). 'Soft and fluffy': medical students' attitudes towards psychology in medical education. *Journal of Health Psychology*, *20*(1), 91-101. https://doi.org/10.1177/1359105313499780

Goyal, S., Nunn, C. A., Rotondi, M., Couperthwaite, A. B., Reiser, S., Simone, A., Katzman, D. K., Cafazzo, J. A., & Palmert, M. R. (2017). A mobile app for the self-management of Type 1 Diabetes among adolescents: A randomized controlled trial. *Journal of Medical Internet Research mHealth and uHealth*, *5*(6), e82. https://doi.org/10.2196/mhealth.7336

Graafland, M., Bemelman, W. A., & Schijven, M. P. (2017). Game-based training improves the surgeon's situational awareness in the operation room: a randomized controlled trial. *Surgical Endoscopy*, *31*(10), 4093-4101. https://doi.org/10.1007/s00464-017-5456-6

Graafland, M., Schraagen, J. M., & Schijven, M. P. (2012). Systematic review of serious games for medical education and surgical skills training. *British Journal of Surgery*, *99*(10), 1322-1330. https://doi.org/10.1002/bjs.8819

Graafland, M., Vollebergh, M. F., Lagarde, S. M., van Haperen, M., Bemelman, W. A., & Schijven, M. P. (2014). A serious game can be a valid method to train clinical decision-making in surgery. *World Journal of Surgery*, *38*(12), 3056-3062. https://doi.org/10.1007/s00268-014-2743-4

Hale, A. J., Ricotta, D. N., Freed, J., Smith, C. C., & Huang, G. C. (2019). Adapting Maslow's Hierarchy of Needs as a Framework for Resident Wellness. *Teaching and Learning in Medicine*, *31*(1), 109-118. https://doi.org/10.1080/10401334.2018.1456928

Howarth-Hockey, G., & Stride, P. (2002). Can medical education be fun as well as educational? *British Medical Journal*, *325*(7378), 1453-1454. https://doi.org/10.1136/bmj.325.7378.1453

Kollei, I., Lukas, C. A., Loeber, S., & Berking, M. (2017). An app-based blended intervention to reduce body dissatisfaction: A randomized controlled pilot study. *Journal of Consulting and Clinical Psychology*, *85*(11), 1104-1108. https://doi.org/10.1037/ccp0000246

Kouwenhoven-Pasmooij, T. A., Robroek, S. J., Ling, S. W., van Rosmalen, J., van Rossum, E. F., Burdorf, A., & Hunink, M. G. (2017). A blended web-based gaming intervention on changes in physical activity for overweight and obese employees: Influence and usage in an experimental pilot study. *Journal of Medical Internet Research Serious Games*, *5*(2), e6. https://doi.org/10.2196/games.6421

Kurtzman, G. W., Day, S. C., Small, D. S., Lynch, M., Zhu, J., Wang, W., Rareshide, C. A., & Patel, M. S. (2018). Social incentives and gamification to promote weight loss: The lose it randomized, controlled trial. *Journal of General Internal Medicine*, *33*(10), 1669-1675. https://doi.org/10.1007/s11606-018-4552-1

Lubarsky, S., Dory, V., Duggan, P., Gagnon, R., & Charlin, B. (2013). Script concordance testing: from theory to practice: AMEE guide no. 75. *Medical Teacher*, *35*(3), 184-193. https://doi.org/10.3109/0142159X.2013.760036

Lubarsky, S., Dory, V., Meterissian, S., Lambert, C., & Gagnon, R. (2018). Examining the effects of gaming and guessing on script concordance test scores. *Perspectives on Medical Education*, *7*(3), 174-181. https://doi.org/10.1007/s40037-018-0435-8

Michael, K., Dror, M. G., & Karnieli-Miller, O. (2019). Students' patient-centered-care attitudes: The contribution of self-efficacy, communication, and empathy. *Patient Education and Counseling*. https://doi.org/10.1016/j.pec.2019.06.004

Muis, K. R., Winne, P. H., & Jamieson-Noel, D. (2007). Using a multitrait-multimethod analysis to examine conceptual similarities of three self-regulated learning inventories. *British Journal of Educational Psychology*, *77*(Pt 1), 177-195. https://doi.org/10.1348/000709905X90876

Mullikin, T. C., Shahi, V., Grbic, D., Pawlina, W., & Hafferty, F. W. (2019). First year medical student peer nominations of professionalism: A methodological detective story about making sense of non-sense. *Anatomical Sciences Education*, *12*(1), 20-31. https://doi.org/10.1002/ase.1782

Nevin, C. R., Westfall, A. O., Rodriguez, J. M., Dempsey, D. M., Cherrington, A., Roy, B., Patel, M., & Willig, J. H. (2014). Gamification as a tool for enhancing graduate medical education. *Postgraduate Medical Journal*, *90*(1070), 685-693. https://doi.org/10.1136/postgradmedj-2013-132486

Nicola, S., Virag, I., & Stoicu-Tivadar, L. (2017). vr medical gamification for training and education. *Studies in Health Technology and Informatics*, 236, 97-103. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28508784

O'Connor, D., Brennan, L., & Caulfield, B. (2018). The use of neuromuscular electrical stimulation (NMES) for managing the complications of ageing related to reduced exercise participation. *Maturitas*, 113, 13-20. https://doi.org/10.1016/j.maturitas.2018.04.009

Paros, S., & Tilburt, J. (2018). Navigating conflict and difference in medical education: insights from moral psychology. *BMC Medical Education*, *18*(1), 273. https://doi.org/10.1186/s12909-018-1383-z

Patel, M. S., Benjamin, E. J., Volpp, K. G., Fox, C. S., Small, D. S., Massaro, J. M., Lee, J. J., Hilbert, V., Valentino, M., Taylor, D. H., & Manders, E. S. (2017). effect of a game-based intervention designed to enhance social incentives to increase physical activity among families: The BE FIT randomized clinical trial. *Journal of the American Medical Association Internal Medicine*, *177*(11), 1586-1593. https://doi.org/10.1001/jamainternmed.2017.3458

Pickren, W. (2007). Psychology and medical education: A historical perspective from the United States. *Indian Journal of Psychiatry*, *49*(3), 179-181. https://doi.org/10.4103/0019-5545.37318

Roepke, A. M., Jaffee, S. R., Riffle, O. M., McGonigal, J., Broome, R., & Maxwell, B. (2015). Randomized controlled trial of superbetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. *Games for Health Journal*, *4*(3), 235-246. https://doi.org/10.1089/g4h.2014.0046

Rutledge, C., Walsh, C. M., Swinger, N., Auerbach, M., Castro, D., Dewan, M., Khattab, M., Rake, A., Harwayne-Gidansky, I., Raymond, T. T., & Maa, T. (2018). Gamification in action: Theoretical and practical considerations for medical educators. *Academic Medicine*, *93*(7), 1014-1020. https://doi.org/10.1097/ACM.0000000000002183

Savulich, G., Piercy, T., Fox, C., Suckling, J., Rowe, J. B., O'Brien, J. T., & Sahakian, B. J. (2017). Cognitive training using a novel memory game on an ipad in patients with amnestic mild cognitive impairment (aMCI). *International Journal of Neuropsychopharmacology*, *20*(8), 624-633. https://doi.org/10.1093/ijnp/pyx040

Shah, A., Carter, T., Kuwani, T., & Sharpe, R. (2013). Simulation to develop tomorrow's medical registrar. *The Clinical Teacher*, *10*(1), 42-46. https://doi.org/10.1111/j.1743-498X.2012.00598.x

Sheehan, D., Bagg, W., de Beer, W., Child, S., Hazell, W., Rudland, J., & Wilkinson, T. J. (2010). The good apprentice in medical education. *New Zealand Medical Journal*, *123*(1308), 89-96. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20201158

Sheikhnezhad Fard, F., & Trappenberg, T. P. (2019). A novel model for arbitration between planning and habitual control systems. *Frontiers in Neurorobotics*, *13*, 52. https://doi.org/10.3389/fnbot.2019.00052

Team, R. C. (2019). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing.

Turan, S., Demirel, O., & Sayek, I. (2009). Metacognitive awareness and self-regulated learning skills of medical students in different medical curricula. *Medical Teacher*, *31*(10), e477-483. https://doi.org/10.3109/01421590903193521

van Houten-Schat, M. A., Berkhout, J. J., van Dijk, N., Endedijk, M. D., Jaarsma, A. D. C., & Diemers, A. D. (2018). Self-regulated learning in the clinical context: A systematic review. *Medical Education*, *52*(10), 1008-1015. https://doi.org/10.1111/medu.13615

Van Nuland, S. E., Roach, V. A., Wilson, T. D., & Belliveau, D. J. (2015). Head to head: The role of academic competition in undergraduate anatomical education. *Anatomical Sciences Education*, *8*(5), 404-412. https://doi.org/10.1002/ase.1498

Villavicencio, F. T., & Bernardo, A. B. (2013). Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology*, *83*(Pt 2), 329-340. https://doi.org/10.1111/j.2044-8279.2012.02064.x

Vohs, K. D., Baumeister, R. F., Schmeichel, B. J., Twenge, J. M., Nelson, N. M., & Tice, D. M. (2008). Making choices impairs subsequent self-control: A limited-resource account of decision making, self-regulation, and active initiative. *Journal of Personality and Social Psychology*, *94*(5), 883-898. https://doi.org/10.1037/0022-3514.94.5.883

Wan, M. S., Tor, E., & Hudson, J. N. (2018). Improving the validity of script concordance testing by optimising and balancing items. *Medical Education*, *52*(3), 336-346. https://doi.org/10.1111/medu.13495

Wisniewski, A. B., & Tishelman, A. C. (2019). Psychological perspectives to early surgery in the management of disorders/differences of sex development. *Current Opinion in Pediatrics*, *31*(4), 570-574. https://doi.org/10.1097/MOP.0000000000000784

Yue, P., Zhu, Z., Wang, Y., Xu, Y., Li, J., Lamb, K. V., Xu, Y., & Wu, Y. (2019). Determining the motivations of family members to undertake cardiopulmonary resuscitation training through grounded theory. *Journal of Advanced Nursing*, *75*(4), 834-849. https://doi.org/10.1111/jan.13923
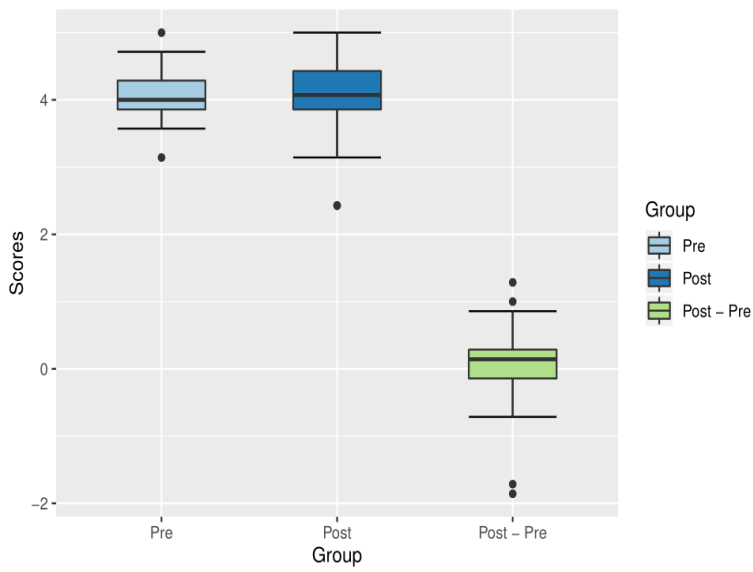
*Ang Eng-Tat
Department of Anatomy,
Yong Loo Lin School of Medicine,
MD10, National University of Singapore,
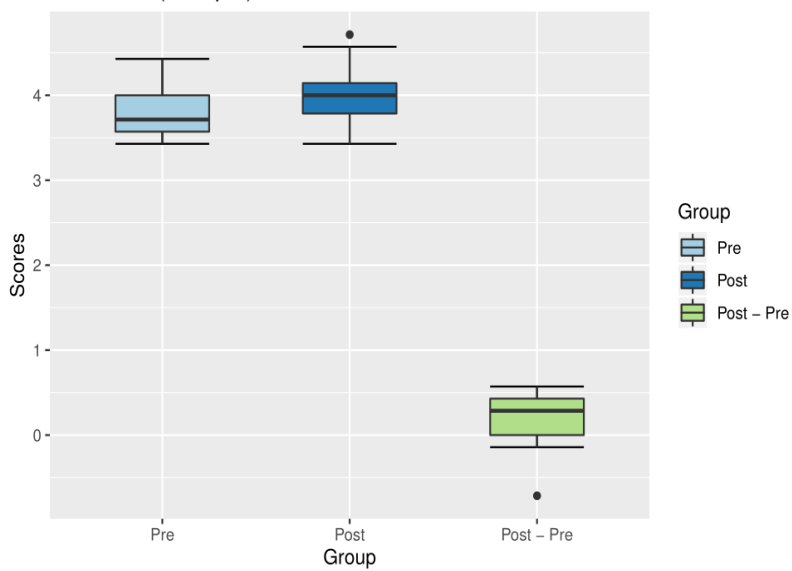Singapore 117599
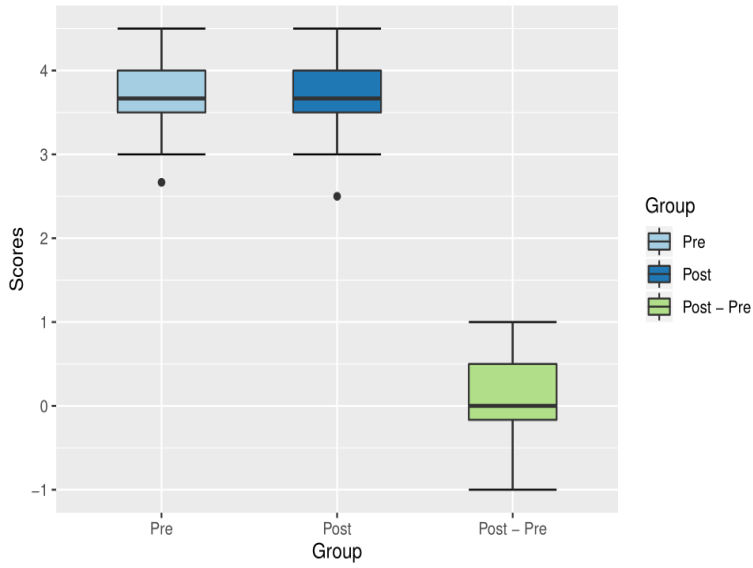Email: antaet@nus.edu.sg

Appendix 1

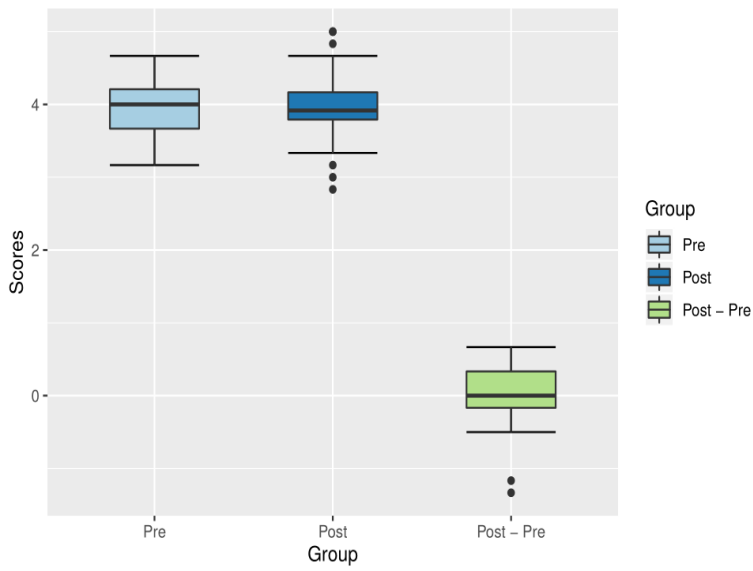## Motivation (Group 0)
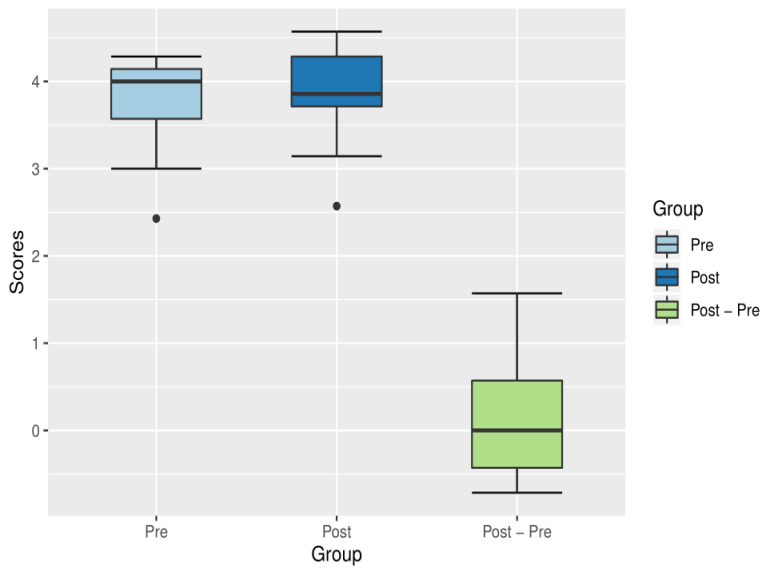
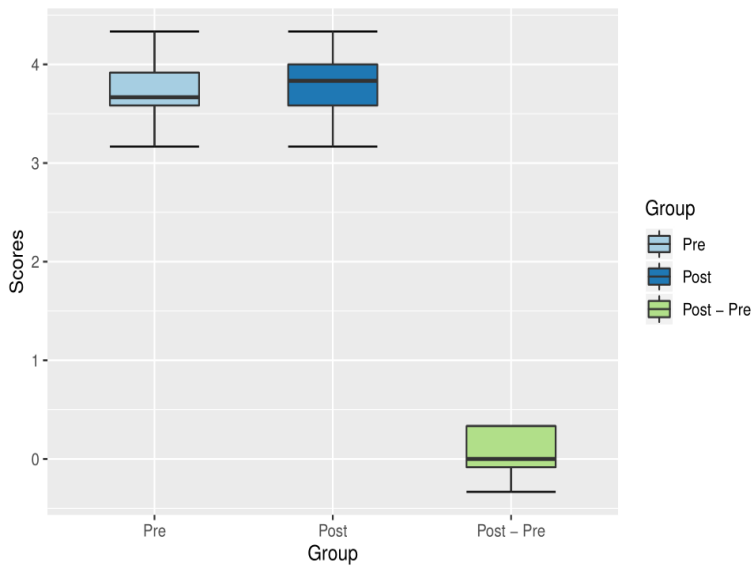## Motivation (Group 1)

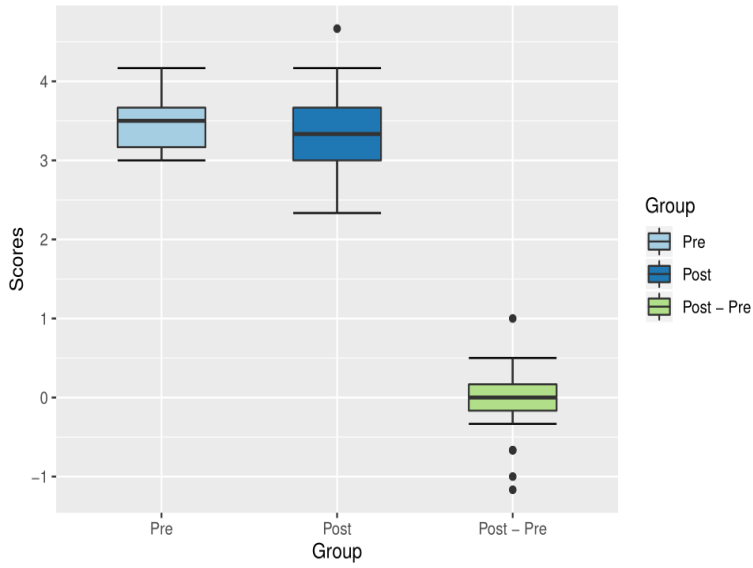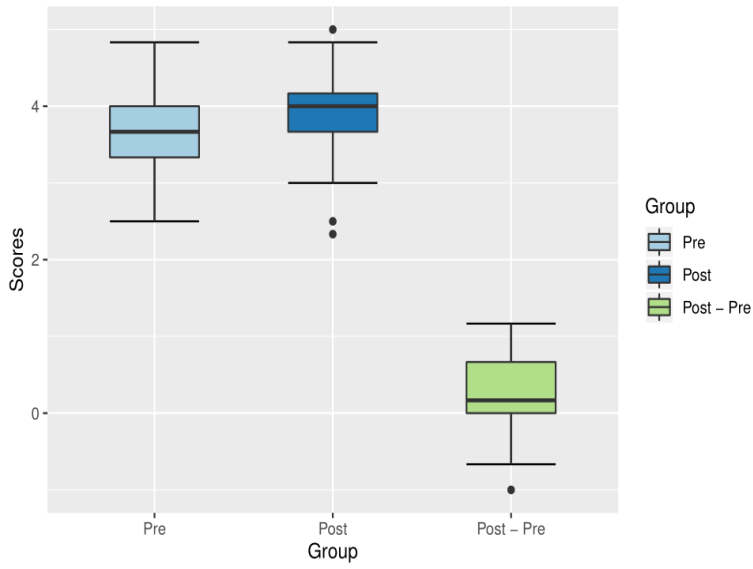## Motivation (Group 2)

Control (Group 0)
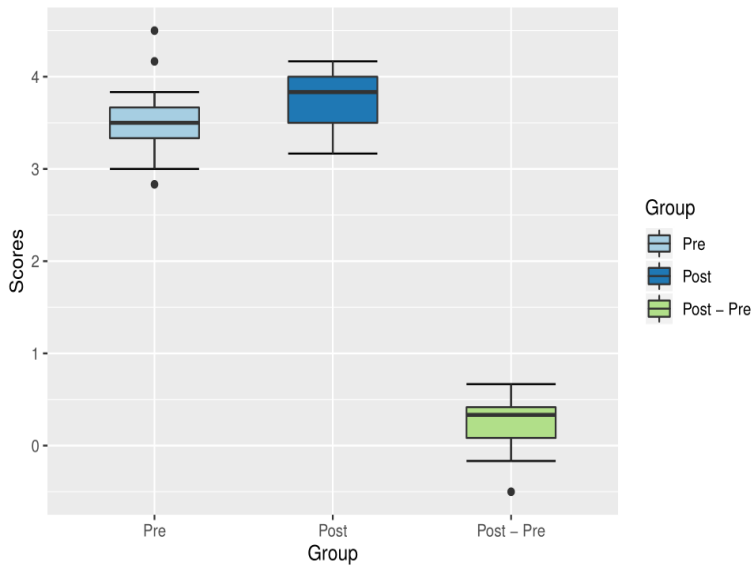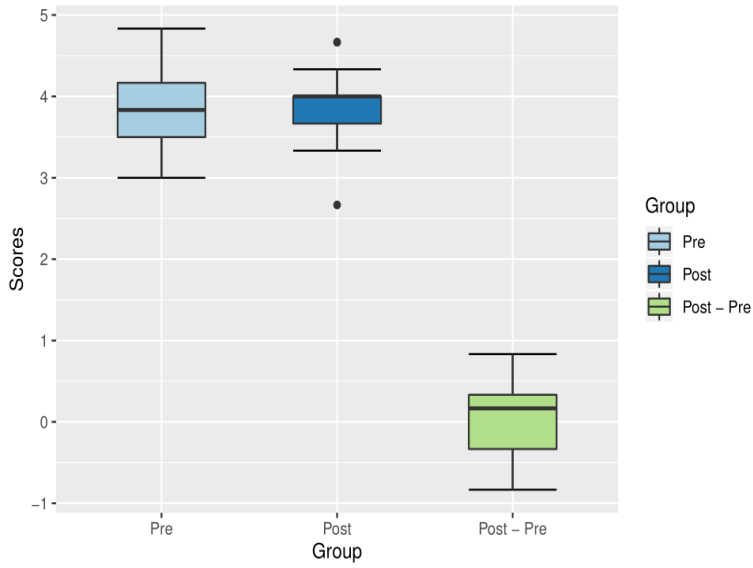

Control (Group 1)
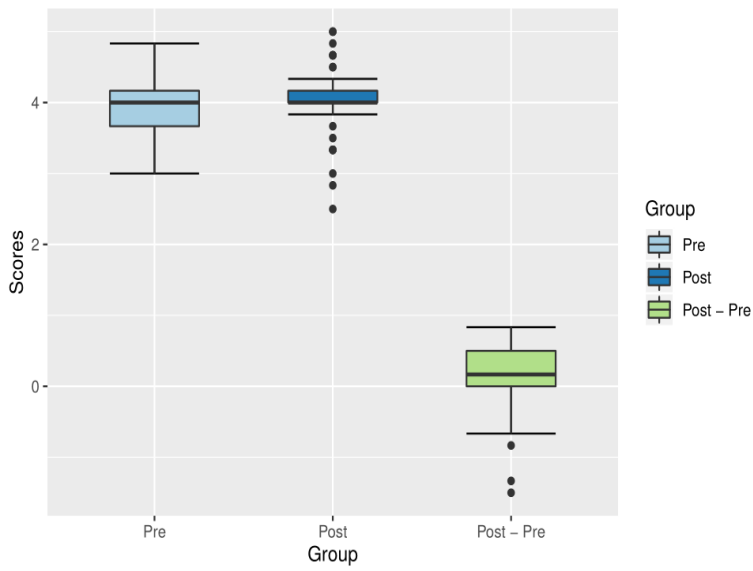

Control (Group 2)
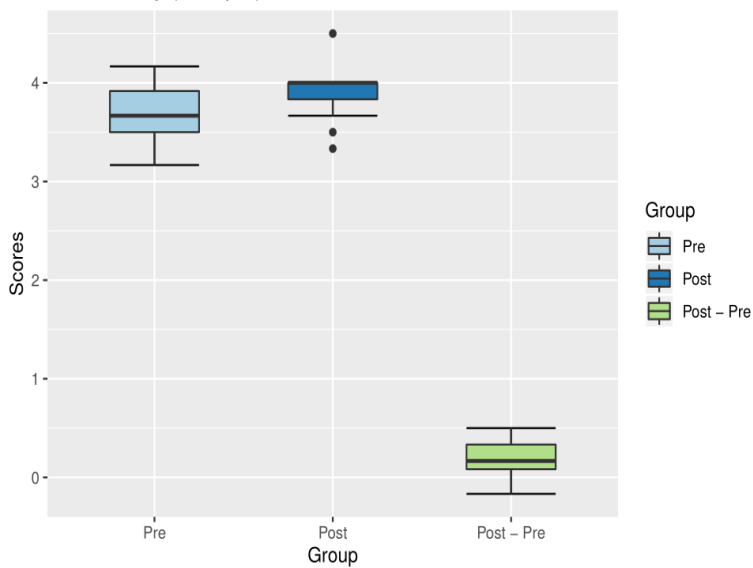
Initiative (Group 0)

Initiative (Group 1)

Initiative (Group 2)

Self Efficacy (Group 0)

Self Efficacy (Group 1)

Self Efficacy (Group 2)

Appendix 2

-2: much less likely
-1: less likely
0: neither more nor less likely
+1: more likely
+2: much more likely

Clinical scenario						
A: A man, with history of hypertension, ex-smoker, presents with a central neck lump.						
If you were thinking of…	**And then you find that…**	**This hypothesis becomes…**				
1. Thyroglossal cyst	The patient is 80 years old.	-2	-1	0	1	2
2. Thyroid nodule	The lump moves with swallowing.	-2	-1	0	1	2
3. Lymph node	The lump is fixed to skin.	-2	-1	0	1	2

Clinical scenario						
B: A 56 year old woman, with a past history of ear surgery, presents with ear discharge.						
If you were thinking of…	**And then you find that…**	**This hypothesis becomes…**				
1. Recurrent cholesteatoma	There are fungal spores in the external ear canal.	-2	-1	0	+1	+2
2. Infection of mastoid cavity	The external ear canal wall appears intact.	-2	-1	0	+1	+2
3. Chronic otitis media	The tympanic membrane is intact.	-2	-1	0	+1	+2

Clinical scenario						
C: A 40 year old woman presents with recurrent sore throat for past 4 months.						
If you were thinking of…	**And then you find that…**	**This hypothesis becomes…**				
1. Recurrent tonsillitis	Her tonsils appear normal.	-2	-1	0	+1	+2
2. Oropharyngeal cancer	There is no history of smoking.	-2	-1	0	+1	+2
3. Gastroesophageal reflux disease	She has no history of heartburn.	-2	-1	0	+1	+2

Clinical scenario		
D: A 46 year old man presents with recurrent bleeding nose for past 3 months.		
If you were thinking of…	**And then you find that…**	**This hypothesis becomes…**

1. Nasopharyngeal carcinoma	There is no palpable cervical lymph node.	-2	-1	0	+1	+2
2. Bleeding from the Little's area	Examination of the Little's area is normal.	-2	-1	0	+1	+2
3. Sinonasal tumor	He has a family business in wood industry.	-2	-1	0	+1	+2

Clinical scenario						
E: A 60-year-old man has sudden onset of left facial weakness.						
If you were thinking of…	**And then you find that…**	**This hypothesis becomes…**				
1. Stroke	The all the muscles in the entire half of the left face is weak.	-2	-1	0	+1	+2
2. Bell's palsy	Hearing and ear examination is normal.	-2	-1	0	+1	+2
3. Ramsay Hunt Syndrome	There are vesicles on the pinna.	-2	-1	0	+1	+2

Clinical scenario						
F: A 25-year-old woman presents with lateral neck lump.						
If you were thinking of…	**And then you find that…**	**This hypothesis becomes…**				
1. Papillary thyroid cancer	Fine needle aspiration cytology shows nuclear grooves.	-2	-1	0	+1	+2
2. Hashimoto's thyroiditis	Thyroid function test is normal	-2	-1	0	+1	+2
3. Follicular thyroid cancer	Ultrasound of the thyroid shows comet-tail artifact.	-2	-1	0	+1	+2

Legend
-2: much less appropriate
-1: less appropriate
0: neither more or less appropriate
+1: more appropriate
+2: much more appropriate

Clinical scenario						
G: A 45-year-old man presents with sudden onset of hearing loss.						
If you were thinking of…	**And then you find that…**	**Then your plan of action becomes…**				
1. Prescribing oral steroids	The hearing loss occurred 5 weeks ago.	-2	-1	0	+1	+2
2. Prescribing oral steroids	He has type 2 diabetes mellitus.	-2	-1	0	+1	+2
3. Prescribing oral steroids	He has a history of depression.	-2	-1	0	+1	+2

Clinical scenario						
H: A 78-year-old woman presents with ear discharge.						
If you were thinking of…	**And then you find that…**	**Then your plan of action becomes…**				
1. Prescribing antibiotics eardrops	There are spores seen in the external auditory canal.	-2	-1	0	+1	+2
2. Prescribing oral antibiotics	She has a facial nerve palsy.	-2	-1	0	+1	+2
3. Sending the ear discharge for culture and sensitivity	Her ear discharge has been persistent for 4 weeks.	-2	-1	0	+1	+2