

Pretest and posttest reliability to measure the impact of online team-based learning workshops during the Covid-19 pandemic as a continuing faculty development activity among health professional educators

Shahid Hassan

Department of Surgery School of Medicine/IMU Centre for Education, International Medical University, Malaysia.

Author; Email: ShahidHassan@imu.edu.my.

Accepted 25th November, 2022.

Abstract. Medical education has been facing unprecedented challenges due to the COVID-19 pandemic leaving aftermath, which needs continuing reparative work. Imperative to exploring solutions faced by Covid pandemic challenges, various immediate steps were taken by medical schools globally including the faculty development activities transformed into online strategies. Workshops are the most common methods of the FDA to enhance knowledge and skills in specific areas of teaching and learning and assessment. A retrospective analysis of three of the TBL online FDA workshops is presented to determine its immediate impact on teaching and learning as well as an assessment model using pretest and posttest designs of a 3-point Likert scale of participants attending a half-day online workshop during the Covid19 pandemic for faculty development in a teaching hospital in Malaysia. Data was collected and analysed using online links created in Microsoft Google form at the beginning and towards the end of each workshop on a 3-point Likert scale of agreed, disagreed and not sure. A 15 items questionnaire was administered with additional information of participants' trends in studying the reading materials sent out a week prior to the actual workshop. An interventional strategy between the two tests were four plenaries and a break-out session with hands-on group work followed by presentations. A collated quantitative data was gathered from 82 participants attending the online TBL workshops. Data were analysed using the statistical test of paired t-test, one-way ANOVA and ANCOVA with effect size in SPSS version 24. The percentage correct pretest and posttest score difference was significant at $p = .013$. The mean difference between pretest and posttest scores was significant at $t = 31.345$, $p < 0.01$. A significant difference in mean scores between groups was found, $F(2,79) = 4.923$, $P = .010$ using one-way ANOVA. Post-hoc analysis revealed significantly high posttest mean scores compared to partially read and not read groups at $p = .033$ and $.014$, respectively. However, the difference between partially read and not read was insignificant. A difference in mean posttest scores among the three groups using one-way ANCOVA on removing the effect of pretest was found significant at $F(2, 80) = 6.211$, $p = 0.003$. The effect size using Cohen's d and Partial Eta Squared have a large practical effect of 2.072 and a moderate effect of 11.1% and 13.70% shown in paired t-test, ANOVA and ANCOVA respectively. A significant difference in the mean pretest and posttest scores was found between the groups. A post-test score, controlling on pretest score, was also significant, with a large Cohen's d and moderate practical effect size suggesting an effectively delivered TBL online workshop.

Keywords: Faculty development, online workshops, pretest, posttest design, quantitative evaluation, team-based learning.

INTRODUCTION

It is a daunting task to measure the impact, be it an immediate or delayed impact of any faculty development activity or workshop. FDP plays a significant role in undertaking on-the-job learning, beneficial to the career of the faculty (Jehanzeb and Bashir, 2012). FDPs are also important for faculty development in health professions education as well as in the implementation of medical curricula efficiently (Steinert et al., 2016; World Health Organization, 2013). Besides teaching and learning, the role of the faculty is equally important in assessment for the successful implementation of the curriculum in medical education. Assessment developed as learning helps in many ways, from monitoring the students' progress in training to providing feedback to students and insight to teachers to plan and guide instructional strategies (Shahid, 2020). Faculty development activities (FDA) have been a regular feature in most of the institutions involved in health professions education. However, what has been ignored is the evaluation of those training and workshops for their impact to determine their effectiveness (Lancaster et al., 2014). Gillespie and colleagues recommend 10 steps to be considered when building a Faculty Development. Higher education recommends strategies for improving faculty knowledge and skills across all levels of teaching faculty, especially in teaching and learning inclusive of assessment (Cook and Kaplan, 2011).

Assessment as a learning may provide an ongoing diagnostic approach to learning, especially in a crisis like the Covid-19 pandemic (Shahid, 2020). Faculty should consider and practice teaching and learning methods that incorporate assessment and Team-Based Learning (TBL) provides one such option. TBL is an active learning approach designed for small group instructional sessions, which can be applied to a larger class. TBL is the instructional strategy that provides students with opportunities to apply factual and conceptual knowledge to solve problems in sequentially designed activities that includes pre-class individual work assessed through readiness assurance test both at individual and group level, in-class application of knowledge and feedback (Shahid et al., 2018). Motivating students to work as a team this method of teaching aims to explore students' critical thinking by using carefully designed in-class activities and assignments as collaborative teamwork, which is a characteristic feature of TBL. TBL provides frequent and immediate feedback which is vital for learning and content retention. This can take place at several stages such as GRAT or TRAT (Group/Team Readiness Assurance Test) (Shahid et al., 2018), peer evaluation and in-classroom application of assignment from the theme of clinical problems identified by each discipline in basic surgical posting. The success of teaching through this method like other small group discussions such as PBL and CBL depends on how the faculty perceive this method of teaching and assessment and how they implement TBL.

Therefore, mandatory faculty training in TBL is as important as the reliability of training in terms of its immediate impact. For immediate impact, a pretest/posttest questionnaire using a 3-5 point Likert scale provides a good tool to collect the data. However, the authentication of such an assessment requires reliable evidence. There is a lack of evidence supporting the authentication of such instruments in faculty development workshops (Arunava et al., 2019).

A robust set of statistical methods used to analyse the data collected on a pre and posttest questionnaire-based instrument used for online TBL workshop provide reliable evidence. This procedure provides feedback to facilitators by measuring the knowledge and skills that participants have prior to attending the training programme and what they gained after attending the workshop (Barge, 2007). This will ensure continuing faculty improvement in health profession education. The learning outcome of the workshop can be measured using pre/posttest questionnaires to collect and analyse the data using appropriate statistical tools depending on the hypothesis set for the training programme. Participants may perform poorly in the initial pre-test, and some even hesitate to respond. The subsequent posttest captures the participants' improvement as an immediate impact. Facilitators must reassure participants that pre/posttest questionnaire aims to determine the workshop impact and not the test of their knowledge for any other purpose and that this information will be used to monitor the objective and the intended learning outcome of the workshop (Jayachandran and Balaji, 2016).

METHODOLOGY

TBL, an ideal teaching and assessment learning method was designed to develop faculty with enhanced knowledge and skills using online workshops for its impact. A retrospective quantitative pretest and posttest study designed to determine educators' acquisition of required knowledge and skills has been held during the earlier days of the Covid-19 pandemic from 2019 to 2022. The data was collected and analysed using Microsoft Form, provided by our institution. Using an institution account, any respondent could be tracked and identified (via their ID) by default. Pre/posttest form was administered at the beginning and towards the end of each workshop, on a 3-point Likert scale of agreed, disagreed and not sure. A 15 items questionnaire was developed and administered that covered various aspects of TBL philosophy, structure, implementation strategies, pre-class and in-class application of knowledge and assessment based on the individual readiness assurance test (IRAT) and group readiness assurance test (GRAT) as well as peer evaluation as the special features of TBL. Additional

Table 1. The % correct pretest and posttest score results are significant at $p \leq 0.05$.

Pre/posttest items	Pretest result	Posttest result
	% Correct score	% Correct score
Question 1	13.462	26.923
Question 2	48.462	96.154
Question 3	61.154	94.231
Question 4	32.692	36.538
Question 5	28.846	28.846
Question 6	46.538	86.154
Question 7	69.231	88.462
Question 8	65.385	84.615
Question 9	80.769	88.462
Question 10	25.000	86.538
Question 11	50.000	67.308
Question 12	67.308	88.462
Question 13	57.692	88.462
Question 14	75.000	88.462
Question 15	42.154	48.077
Average	50.913	72.795

Table 2. Paired t-Statistics to determine the significant difference of means between pretest and posttest knowledge of participants.

Pair pretest and posttest score	Pretest mean (SD)	Posttest mean (SD)	Mean (SD) score difference (95% CI)	t-statistics (df)	p-value
N = 82	4.27 (.930)	11.44 (2.103)	7.171 (2.072) (6.716, 7.626)	31.345 (81)	<0.001

information on participants' trends in studying the reading materials sent out a week prior to the actual workshop was also collected.

An interventional strategy between the two tests was delivered through four specified plenaries and a break-out session with hands-on group work followed by presentations. Task accomplished in group work was presented by each group in the main room for peer and facilitator feedback and question-answer session to clear the ambiguities associated with this format of teaching and learning and assessment model. A collated quantitative data was gathered from 82 participants attending the online TBL workshops delivered every year as faculty development activity by the IMU Centre for Education. Two sets of data sheets were developed from the Excel file of Microsoft Form after decoding as the numeric raw data and % correct answer. The % correct data and the raw data collected were analysed using independent t-test, paired t-test, one-way ANOVA and ANCOVA respectively. ANCOVA was used to analyse the effect of groups on posttest removing the impact of pretest. The effect size associated with each of these statistical methods was also determined using Cohen's *d* in Paired t-test and Partial Eta Squared in ANOVA and ANCOVA to estimate the practical effect and its impact on participants learning through online workshops using Microsoft SPSS (Version 28,

SPSS Inc., Chicago, IL, USA).

RESULTS

The pretest and posttest data were divided into three groups of well-read, partially read or not read participants of reading materials sent out to each participant a week prior to the workshop. The % correct pretest and posttest data opened in Excel in Microsoft Form was downloaded and analysed using an independent t-test and found significant at $p = .013$ (Table 1). The result of pair t-statistics indicates a mean difference of 7.171 between the pretest score = 4.27 (.930) and post-test score = 11.44 (2.103) respectively significant at, $t = 31.345$, $p < 0.001$ (Table 2) associated with a significant correlation of 0.255.

The result of ANOVA (Tables 3 and 4) showed Levene's test of homogeneity of variance significant at .406 with a significant difference between groups; $F(2,79) = 4.923$, $P = .010$. Bonferroni post-hoc analysis (Table 4) revealed, well-read ($n = 27$, $M = 8.148$, $SD = 1.955$) with significantly high posttest score on average than both partially read ($n = 29$, $M = 6.724$, $SD = 2.169$) and not read ($n = 26$, $M = 6.724$, $SD = 2.169$) at $p = .033$ and .014 respectively. Partially read and those not read groups are not significantly different from each other. A difference in mean

Table 3. ANOVA F statistics to determine the significant difference among the three independent variable groups.

Group-based on pre-session reading	Number	Mean Score (SD)	F-statistics (df)	p-value
Well-read	27	8.148 (1.955)	4.923 (2,79)	.010
Partially read	29	6.724 (2.169)		
Not read	26	6.653 (1.765)		

Table 4. Post-hoc test to determine the group's differences in mean scores among the three groups of well-read, partially read and not read through the reading materials.

(I) Group	(J) Group	(I-J) Mean Difference	p-value	95% CI	
				Lower bond	Upper bond
Well-read	Partially read	1.424	.026	.130	2.717
	Not read	1.494	.022	.164	2.823
Partially read	Well-read	1.424	.026	-2.717	-.130
	Not read	.070	1.000	-1.236	1.377
Not read	Well-read	1.494	.022	-2.823	-1.649
	Partially read	.070	1.000	1.377	1.236

Table 5. ANCOVA between subject effects after removing the effect of covariance of pretest.

Group	Number	Mean (SD)	F-statistics (df)	Significance	Partial Eta Squared
Well-read	27	12.59 (1.824)	6.211 (2,79)	.003	.137
Partially read	29	11.03 (2.195)			
Not read	26	10.69 (1.860)			

Table 6. Effect Size as Cohen's d and Partial Eta Squared determined to know the practical effects.

Paired t-Test (Cohen's d)	One Way ANOVA (Partial Eta Squared)	ANCOVA (Partial Eta Squared)
Low = .2, Medium = .5, High = .8		
2.072	.111 (11.10%)	.137 (13.70%)

posttest scores among the three groups using one-way ANCOVA (Table 5) on removing the effect of covariant of pretest was found significant, $F(2,80) = 6.211$, $P = .003$. The effect size associated with paired t-test as Cohen's d, one-way ANOVA without controlling for pretest and ANCOVA controlling for pretest using Partial Eta Squared was also estimated and a substantial effect size suggests its practical impact (Table 6).

DISCUSSION

In IMU, FDAs are run throughout the year to enhance faculty knowledge and skills in teaching and assessment. Most of these FDAs as workshops were switched over to online sessions with an additional step of introducing the pretest and posttest assessment of participating faculty. This helped to measure the workshop's immediate impact

both statistically and practically as the effect size. TBL was the first of those online workshops developed and delivered by the author with pretest and posttest measures. A retrospective analysis of three of the TBL online faculty development workshops (2019-2022) is presented to determine its immediate impact on teaching and learning as well as the assessment incorporated in this method. The current study is the second of a research project approved by the IMU Joint Committee (JC) on Research and Ethics entitled, "Evaluation of the immediate impact of faculty development programme using pretest-posttest study design in a facilitator guide format". The first study undertook the analysis of data collected from the standard-setting workshop and has been published (Shahid *et al.*, 2021).

The pretest and post-test scores of 82 participants based on the inclusion/exclusion criteria were analysed. The independent variable of pre-workshop preparation by

the participants was categorized into three groups and analyzed for any significant effect on the dependent variable of posttest scores. For differences within the group, the current study ran a paired t-test, and for the difference between the group, 15 items questionnaire across pretest and posttest, independent t-test among the three groups, a one-way ANOVA was run. The result of ANOVA (Tables 3 and 4) showed a significant difference between participants prepared for the workshop (well-read, partially read, and not read) and its impact on their posttest scores. The null hypothesis of no difference in posttest mean scores between the attainment of reading materials as well-read, partially read or not read is rejected to accept the alternative hypothesis.

The current study like the earlier published study (Shahid *et al.*, 2021) investigated the pretest/posttest format utilised to determine the immediate impact of the workshop. Evaluation reveals that there is a difference in the mean scores between the two with posttest predominantly high than the pretest score in all the three statistical methods used (Tables 1, 2 and 3). The study also investigated the impact of participants categorised into three groups of well-read, partially read and not read in all those reading materials sent out a week prior to the workshop, on pretest and posttest scores.

Within-group difference has been found between pretest and posttest mean scores significant at $p \leq .001$. Looking at the data downloaded from the Microsoft Form Excel sheet, the % correct answer through all 15 questions also indicates a significant ($p = .013$) difference between pretest and posttest scores (Table 1). These differences in the mean scores and % correct score have been shown to have a positive impact on intervention observed in participants' gain in knowledge and skills through plenaries and hands-on workshops. Pretest and posttest questionnaires have been useful instruments in rejecting the null hypothesis to accept the research hypothesis that there is a significant difference in knowledge and skills acquired in these workshops. The current study is comparable with a couple of other studies with similar outcomes (Arunava *et al.*, 2019; Hartley, 1973).

A significant difference has been observed between participants who prepared for the workshop (completely read, partially read, and not read) with its impact on their posttest scores (Tables 3 and 4). Post-hoc analysis (Table 4) revealed that well-read has significantly high posttest scores on average than both partially read and not read respectively. Partially read and those not read groups are not significantly different from each other. The null hypothesis that there is no difference in posttest mean scores between the attainment of reading materials as well-read, partially read or not read would be rejected.

Overall pretest and posttest questionnaires have been found to be effective instruments to analyze the data collected in two different measures prior to intervention and after the intervention. Positive impact indicates that intervention between the two measures has been effective

for participants' immediate retention of knowledge and skills acquired through the workshops.

CONCLUSION

A significant difference between the mean pretest and posttest scores was established within and between the groups. The application of a pretest and posttest questionnaire was found to be a feasible and effective tool in estimating the participants' posttest scores after the interventional strategy adopted in a workshop in achieving the workshop learning outcome. A posttest score, controlling on pretest score, was also significant, with a large Cohen's d and moderate practical effect size suggesting an effectively delivered TBL online workshop.

REFERENCES

- Arunava B, Sukanta S, Krishnangshu R (2019). Reliability assessment of pre – post test questionnaire on the impact of a daylong clinical pharmacology workshop among medical professionals. *Asian J. Med. Sci.* 10(6):93-97.
- Barge GL (2007). Pre- and Post-Testing with More Impact. *J. Ext.* 45(6):103.
- Cook CE, Kaplan M (2011). Advancing the culture of teaching on campus: How a teaching centre can make a difference. Stylus Pub.
- Hartley J (1973). The effect of pre-testing on post-test performance. *Instr. Sci.* 2(2):193.
- Jayachandran AL, Balaji J (2016). Introduction of pre and post lecture multiple choice question for second year undergraduate medical students in microbiology: a technique to assess knowledge acquired from the lecture. *Int. J. Res. Med. Sci.* 4:575-578.
- Jehanzeb K, Bashir NA (2012). Training and development program and its benefits to employees and organizations: A conceptual study. *Far East J. Psychol. Bus.* 9(2):58-71.
- Lancaster JW, Stein SM, Maclean LG, Amburgh JV, Persky AM (2014). Faculty Development Program models to advance teaching and learning within health science programs. *Am. J. Pharm. Educ.* 78(5):99. <https://doi.org/10.5688/ajpe78599>.
- Steinert Y, Mann K, Anderson B, Barnett BM, Centeno A, Naismith L, Prideaux D, Spencer J, Tullo E, Viggiano T, Ward H, Dolmans D (2016). A systematic review of faculty development initiatives designed to enhance teaching effectiveness: A 10-year update: BEME Guide No. 40. *Med. Teach.* 38(8):769-786. <https://doi.org/10.1080/0142159x.2016.1181851>.
- Shahid H (2020). Teaching and learning approaches in clinical education during Covid-19 pandemic: Time to take-on assessment as learning. *Int. J. Sci. Res. (IJSR).* 9 (11) November ISSN: 2319-706.
- Shahid H, Ibrahim MS, Nabiha GH (2018). The Structural Framework, Implementation Strategies and Students' Perception of Team-Based Learning in UGME of a Medical School in Malaysia. *Educ. Med. J.* 10(1):53-66.
- Shahid H, Imtiaz B, Nabiha GH (2018). Students' Perceptions about the Immediate Feedback Assessment Technique (IF-AT) in Team-Based Learning. *Educ. Med. J.* 10(1):23-32.
- Shahid H, Sunil PV, Vishna DN (2021). Evaluation of immediate impact of Faculty Development Programme using a pretest-post-test study design format. *Innovation and Education.* <https://doi.org/10.1186/s42862-021-00009-9>.
- World Health Organization. (2013). Transforming and scaling up health professional education and training: Policy brief on faculty development. World Health Organization.