Available online at www.sciencedirect.com

ScienceDirect

Biomedical Journal

journal homepage: www.elsevier.com/locate/bj


Original Article

# Academic outcome and moderator of flipped classroom learning program "*Teaching on the Run*"


*Kuo-Su Chen* [a,f,*], *Ming-Ju Hsieh* [b,f], *Min-Ping Huang* [c,f], *Chih-Ken Chen* [d,f,*], *Ming-Jui Hung* [e,f]

[a] *Division of Nephrology, Chang Gung Memorial Hospital at Keelung, Keelung, Taiwan*
[b] *Division of Cardiovascular Surgery, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan*
[c] *Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan*
[d] *Division of Psychiatry, Chang Gung Memorial Hospital at Keelung, Keelung, Taiwan*
[e] *Division of Cardiology, Chang Gung Memorial Hospital at Keelung, Keelung, Taiwan*
[f] *College of Medicine, Chang Gung University, Taoyuan, Taiwan*

ARTICLE INFO

*Article history:*
Received 30 July 2019
Accepted 4 May 2020
Available online 1 June 2020

*Keywords:*
Flipped classroom
*Teaching on the Run* (TOTR)
Outcome
Moderator

ABSTRACT

*Background:* Flipped classroom (FC) style Australian faculty development program *Teaching on the Run* (TOTR) was introduced into Chang Gung Memorial Hospital since 2014. However, its effectiveness in Taiwan has not been formally assessed. This work intended to examine the learning gain of TOTR and identify the moderators of FC outcome by using TOTR as a representative model of FC.

*Methods:* A non-controlled before-after study was undertaken by retrospective analysis of learning data collected during TOTR workshop. Multiple choice questions were tested at baseline (pre-test), after pre-class learning (mid-test) and after classroom activity (post-test) to assess the learning gain. All available demographic and learning variables were included in the moderator analysis.

*Results:* Stepwise and significant improvement in exam scores was noted from pre-test to mid-test and post-test ($p < 0.001$ for both). Univariate analysis showed pre-test scores, mid-test scores, class participation and session of TOTR were significantly associated with post-test scores. However, multivariate analysis by general linear model showed only mid-test scores and session of TOTR were significant predictor of post-test score. Generalized estimating equations analysis showed that class participation is a significant moderator that influence the scores change from mid-test to post-test.

*Conclusion:* TOTR is effective in improving knowledge of teaching skills for clinical teachers in Taiwan. Achievement in pre-class learning, class participation and learner factor are potential moderators of the FC outcome. Thus, facilitators should try their best to promote a good achievement in pre-class learning and engagement in classroom activity in FC style learning.

* *Corresponding author*. Department of Nephrology and Psychiatry, Chang Gung Memorial Hospital at Keelung, 222, Mai Chin Rd., Keelung 204, Taiwan.
E-mail addresses: cksdavid@adm.cgmh.org.tw (K.-S. Chen), kenchen@cgmh.org.tw (C.-K. Chen).
Peer review under responsibility of Chang Gung University.
https://doi.org/10.1016/j.bj.2020.05.009

## At a glance commentary

### *Scientific background on the subject*

*Teaching on the Run* (TOTR) is a flipped style Australian faculty development program. It was introduced to Chang Gung Medical Foundation in 2014. This study examined the effectiveness of TOTR program in Taiwan and explored the moderators of flipped classroom learning outcome by analyzing the learning data of TOTR program.

### *What this study adds to the field*

TOTR program is effective in improving the knowledge of teaching skills for clinical teachers in Taiwan. Achievement in pre-class learning, level of class participation and learner factors are potential moderators that might influence the academic outcome of flipped style learning program TOTR.

Flipped classroom (FC) [1,2] is a hybrid pedagogical model that combines online e-learning and face to face classroom activities. It is characterized by two stages of learning process, including a pre-class self-learning outside the classroom and followed by active learning process in the classroom. In FC, lecture is delivered at the pre-class period, and classroom is reserved for students to apply the lecture material [3–5]. With the knowledge acquisition before the classroom and problem solving activity at classroom, a better learning gain in higher cognitive level is expected. Following the wide application of FC across multiple disciplines all over the world, research publications on FC were growing rapidly [6,7]. Most research studies of FC focused on its effectiveness toward satisfaction and learning gain [8–13]. Studies that addressed the moderators of FC outcomes were relatively few [7,14–16]. A recent meta-analysis [17] demonstrated that FC had a better academic outcome than the traditional lecture. However, great heterogeneity was found in both the direction and the effect sizes of outcomes estimate. Such inconsistency between studies might be caused by many factors. One possible cause is the big methodological diversities between studies. This diversity is caused by the fact that there is still no consensus regarding the standard format of FC [18]. Another potential cause is that most published studies were observational design and confounding factors were not well controlled between the study and the control groups [17,19]. However, information was scanty regarding the moderators of FC outcome.

Several potential confounders of FC outcome have been proposed [20–24], including format and design of FC, learner factors and instructor factors. Jesurasa et al. [14] in a qualitative interview identify learning style, teaching style, course design and expectation as barrier or facilitator to engagement in FC learning. Ihm et al. [16] reported that learning readiness significantly influenced learning outcomes in flipped-learning. DeLozier et al. [5] in a review examined the value of out-of-class activities (e.g., video lectures) and in-class activities (e.g., quizzes, student discussions). They pointed out that future work should examine the influence of these individual activities on learning outcome, particularly when objective measures of learning, such as quizzes and exams, are held constant. Liu et al. [7] identify seven factors that might influence the outcome FC, including overall design, design of information, design of technology use, active learning, motivation, special guidance and self-regulated learning from literature review. They used logistic regression to establish model that could predict the FC outcome by incorporating these variables and found overall design, design of information, design of technology use, active learning and motivation were statistical significant. They advocated more explorations to identify factors that may influence the success of flipped learning experiences. Thus research evidence regarding the moderators of FC is still insufficient. More research evidence regarding the confounders of FC is needed.

*Teaching on the Run* (TOTR) is a national faculty development program developed by staff of the Education Centre, Faculty of Medicine and Dentistry at the University of Western Australia [25]. The purpose of TOTR program is to improve the quality of clinical teaching. It was introduced into Chang Gung Medical Foundation in Taiwan since 2014 with an aim to promote the teaching skills of clinical teachers. The program in Taiwan contains five main topics, including learning plan, clinical teaching, assessment, supporting the learner and effective group teaching. These topics cover the major fields of clinical teaching skills. TOTR was implemented by a flipped style design, starting with a 3-wks period of pre-class online learning and followed by a one-day classroom time for active learning. To date, literature evidence regarding the effectiveness of TOTR program is very rare. Its effectiveness in Taiwan context has also not been examined yet. Since TOTR program is executed by using a typical FC format, it is a suitable model to explore the outcome and moderators of FC.

In this work, we retrospectively analyzed the learning data obtained from TOTR workshop to determine the effectiveness of TOTR intervention and identify moderators of FC outcome. The research question of this work is (1) What's the learning gain of TOTR intervention in Taiwan? (2) What are the predictors of final learning gain of FC style TOTR program ? (3) What are moderators that influence the longitudinal change of knowledge gain in each stage of FC learning?

## Methods

### *Design*

This work is a non-controlled before-after study that retrospectively analyzed the learning data of a FC style training program TOTR. All the participants received only TOTR intervention. Learning outcome was assessed by multiple choice questions (MCQ) at baseline (pre-test, T1), after pre-class learning (mid-test, T2) and after classroom activity (post-test, T3). Effect of TOTR intervention was evaluated by before-after comparison of exam scores. Factors that influence the final outcome (post-test scores) was examined by GLM (general linear model). Alternatively, factors that

influence the longitudinal change of exam scores in each stage of learning were examined by the interaction effect in GEE (generalized estimating equations) analysis. Given that this work is an explanatory analysis, all the possible moderators with available information were included in the analysis. Independent variables used in moderator analysis included demographic parameters, hours spent on pre-class learning, completion rate of pre-class homework, and class participation. This study was performed with the approval of the ethical committee of the Chang Gung Memorial Hospital.

### *Participants*

Potential candidates for the TOTR program are clinical teachers of any discipline who worked in Chang Gung Memorial Hospital, Keelung branch. All the clinical teachers who signed up and completed TOTR program between July 2015 and July 2017 were included in the data analysis.

### *TOTR program*

TOTR workshop was held regularly in Chang Gung Memorial Hospital at Keelung. The frequency was four times a year. When a clinical teacher signed up for TOTR course, a baseline MCQ test was performed to evaluate his background knowledge in teaching skills. Then, the participants were guided to undergo online e-learning at their free time within a 3-week period. Facilitators would monitor the process of online learning and give assistance whenever necessary. Three topics of pre-class homework, including "learning plan", "mini clinical teaching" and "Powerpoint presentation", were assigned to complete before class. After pre-class learning, a second MCQ test was undertaken to evaluate the achievement in pre-class learning. A full day face-to-face classroom activity was held after pre-class learning. During the classroom activity, most time was reserved for the participants to undergo active learning, such as discussion, teaching, presentation, exercise and sharing. Active learning skills, including brain storming, think-pair-share, buzz group discussion and role play, were practiced during the classroom time. Additionally, a mini-lecture that briefly reviewed the key concepts of pre-class learning material was given in the classroom time. The purpose of incorporating mini-lecture is to enforce the knowledge retention, and to make up for the knowledge deficit for learners who did not have well pre-class prepare.

The format of FC, pre-class learning material, learning time, in-class learning activities and learning environment was kept same for all sessions of TOTR. The tutor group was comprised of 12 experienced clinical teachers who are familiar with FC learning technique. There were eight tutors/facilitators in each session of classroom activity. Given the small group and active learning nature in classroom activity, each session of TOTR could accommodate at most 20 learners.

### *Outcome measurement*

The outcome of interest was knowledge gain in TOTR training course. The exam scores of post-test were used as the final outcome of TOTR intervention. Instrument used to assess knowledge gain was 14 MCQ developed by the tutor group of TOTR program. Content validity analysis of the MCQ showed that 13 items had I-CVI (item-level content validity index) values greater than 0.8, and 1 had a I-CVI value of 0.75. The overall S-CVI (scale-level content validity index) for the 14 test questions was 0.949. Reliability analysis showed Cronbach's Alpha of the 14 test questions was 0.616.

### *Potential moderators analyzed*

Demographic data of participants were obtained at the time of enrollment. Time spent on pre-class learning was reported by the learners themselves in the feedback questionnaire. Completion rate of pre-class homework and level of participation in classroom activity were assessed by facilitator using a regular assessment form developed by our educational department. The assessment form included 12 items, 3 were evaluating the extent of pre-class homework completion and the other 9 were assessing the level of participation in the classroom activities, including class attendance, concentration in class activity, and extent of engagement in discussion, sharing, role playing, presentation, and feedback. The completion rate and class participation indicators were mainly used to evaluate the learner's engagement.

### *Statistical analysis*

Continuous variables were recorded as mean ± SD. To account for within-individual correlation, GEE model were used to assess the trends of scores changes from pre-test to mid-test and post-test.

Simple linear correlation, independent t test and ANOVA were used in univariate analysis to identify variables significantly associated with the final outcome (post-test scores). Then, multi-variate analyses by GLM was conducted, using the post-test scores as the dependent variable and all significant variables at the univariate analysis as factor or covariates, to determine the significant predictors of final outcome.

GEE model was also used in moderator analysis by examining the interaction of time by various potential confounders. The variables used in moderator analysis included age, gender, academic position, administrative position, hours spent on pre-class learning, completion rate of pre-class material, and class participation.

SPSS software version 24.0 (IBM, Armonk, New York) was used for data analysis. A $p$ value of less than 0.05 is considered statistically significant.

## Results

### *Demographic characteristic of the participants*

From 2015 to 2017, eight sessions of TOTR training workshop were held (July 2015, Sep 2015, April 2016, July 2016, Oct 2016, Nov 2016, April 2017, July 2017). Totally, 164 clinical teachers signed up and 155 completed the training course. However, only 115 took all three MCQ tests (pre-test, mid-test and post-

test). Quantitative data analysis in this study was performed by using the data of these 115 participants. Of these 115 participants, 42 were males. Their occupational categories included 45 physicians, 12 pharmacists, 32 nurses, 10 radiological technicians and 16 other disciplines. The academic level of these participants includes 1 professor, 3 associate professors, 11 assistant professors, 11 lecturers, and 89 non-academic teachers [Table 1].

### *Effect of TOTR intervention on knowledge gain*

Exam scores showed a stepwise increase from pre-test to mid-test (64.84 ± 15.43 to 72.95 ± 13.55, $p = 0.001$) and post-test [80.75 ± 13.0, $p < 0.001$ when compared with both pre-test and mid-test, Fig. 1]. GEE analysis confirmed a stepwise improvement in exam score after each stage of learning activity [Table 2]. The results of GEE analysis for the main effect of time showed that regression coefficient beta was 6.94 when mid-test (T2) score was compared with pre-test (T1) score and beta was 4.32 when post-test (T3) score was compared with mid-test score.

### *Outcome predictor analysis*

Among the multiple variables analyzed, only pre-test scores, mid-test scores, class participation, and session of TOTR program were significantly associated with the post-test scores in univariate analysis (Table 3). Multivariate analysis by GLM with post-test scores as the independent variable and significant factors in univariate analysis as covariates showed

**Table 1 Demographic characteristic of study participants.**

| | Total Participants (n = 115) |
|---|---|
| Age (yrs) | 38.9 ± 7.56 |
| Male gender | 42 (36.5%) |
| Length of clinical teaching (yrs) | 7.83 ± 6.4 |
| Administrative position | |
| Yes | 12 (10.4%) |
| No | 103 (89.6%) |
| Professional category | |
| Physician | 45 (39.1%) |
| Pharmacist | 12 (10.4%) |
| Nurse | 32 (27.8%) |
| Radiation technician | 10 (8.7%) |
| Other category | 16 (13.9%) |
| Academic position | |
| Professor | 1 (0.9%) |
| Associate Professor | 3 (2.6%) |
| Assistant professor | 11 (9.6%) |
| Lecturer | 11 (9.6%) |
| None | 89 (77.4%) |
| Workshop Session | |
| Session 1 | 14 (12.2%) |
| Session 2 | 13 (11.3%) |
| Session 3 | 17 (14.8%) |
| Session 4 | 13 (11.3%) |
| Session 5 | 18 (15.7%) |
| Session 6 | 13 (11.3%) |
| Session 7 | 16 (13.9%) |
| Session 8 | 11 (9.6%) |

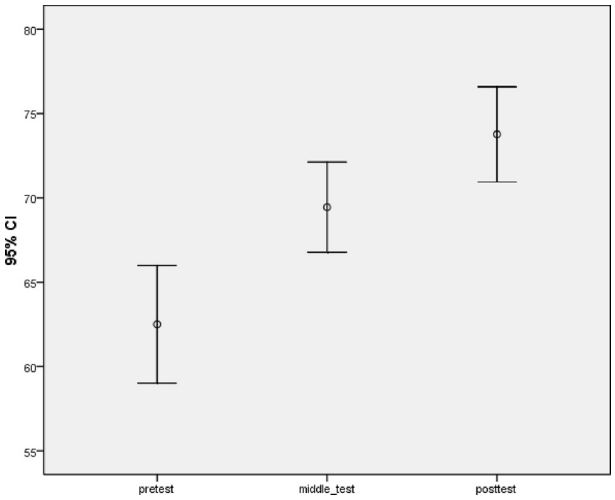


Fig. 1 Examination scores increased significantly from pre-test to mid-test ($p < 0.001$), and post-test ($p < 0.001$ when compared with both pre-test and mid-test).

that only mid-test scores and session of TOTR were significant predictors of final outcome [Table 3].

### *Moderator analysis*

GEE model showed that age, gender, academic position, administrative position, hours spent on pre-class learning had no significant interaction with time [Table 2]. Only class participation by time interaction was statistically significant when post-test score was compared with mid-test score. This indicated that class participation will influence the knowledge gain in classroom learning.

## Discussion

This work for the first time demonstrated the effectiveness of TOTR intervention in enhancing the knowledge on teaching skills for clinical teaching in Taiwan. In this work, some potential confounders of final FC outcome were also identified, including achievement in pre-class learning (mid-test scores), level of class participation, and sessions of TOTR.

The findings that TOTR intervention causes a significant and stepwise improvement in knowledge gain after both pre-class and class learning indicates that both learning stages of FC contributed significantly to the knowledge gain. In the original concept of FC, knowledge transfer is expected to occur mainly in the first stage pre-class online learning. Class time is reserved for active learning with an aim of higher cognitive learning rather than low level knowledge acquisition. The significant improvement in knowledge gain after classroom activity in our data (from mid-test to post-test) might be explained by the inclusion of mini-lecture in the classroom time. In our FC format, a mini-lecture was included in class time to give a brief review of background knowledge. In this way, we can enhance the knowledge retention of well-prepared learners and make up the knowledge gap for un-prepared students. Our results were consistent with the

**Table 2 Potential moderators that might influence the longitudinal change of exam scores by generalized estimating equations analysis.**

| | Predictor variables included in the model | Effect estimate reported | Comparison | β(95%CI) | *p* value |
|---|---|---|---|---|---|
| Model 1 | Time | Main effect of time (without interaction) | T3 vs. T1 | 11.26 (7.83,14.69) | <0.001 |
| | | | T2 vs. T1 | 6.94 (3.58, 10.29) | <0.001 |
| | | | T3 vs. T2 | 4.32 (6.10, 2.55) | <0.001 |
| Model 2 | Time, Gender | Interaction of time by gender | female * T3 vs. male * T1 | −2.74 (−9.47, 3.99) | 0.425 |
| | | | female * T2 vs. male * T1 | −2.72 (−9.43, 3.99) | 0.427 |
| | | | female * T3 vs. male * T2 | −0.02 (−3.54, 3.5) | 0.992 |
| Model 3 | Time, Academic position | Interaction of time by academic position | Academic (+) *T3 vs. Academic (−) *T1 | 0.36 (−6.78, 7.49) | 0.922 |
| | | | Academic (+)*T2 vs. Academic (−) *T1 | −0.21 (-6.85, 6.44) | 0.951 |
| | | | Academic (+)*T3 vs. Academic (−) *T2 | 0.56 (−3.81, 4.94) | 0.801 |
| Model 4 | Time, Administrative position | Interaction of time by administrative position | Administrative (−) *T3 vs. Administrative (+) *T1 | 4.11 (−3.62, 11.83) | 0.297 |
| | | | Administrative (−) *T2 vs. Administrative (+) *T1 | 6.45 (−0.41, 13.30) | 0.065 |
| | | | Administrative (−) *T3 vs. Administrative (+) *T2 | −2.34 (−8.12, 3.45) | 0.428 |
| Model 5 | Time, Age | Interaction of time by age | Age * T3 vs. Age * T1 | −0.02 (-0.47, 0.43) | 0.921 |
| | | | Age * T2 vs. Age * T1 | −0.06 (−0.54, 0.41) | 0.789 |
| | | | Age * T3 vs. Age * T2 | 0.04 (−0.19, 0.27) | 0.724 |
| Model 6 | Time, Hours spent on pre-class learning | Interaction of time by hours spent on pre-class learning | Hours spent on pre-class *T3 vs. Hours spent on pre-class *T1 | 0.08 (−0.50, 0.66) | 0.794 |
| | | | Hours spent on pre-class *T2 vs. Hours spent on pre-class *T1 | 0.26 (−0.37, 0.89) | 0.422 |
| | | | Hours spent on pre-class *T3 vs. Hours spent on pre-class *T2 | −0.18 (0.05, −0.04) | 0.117 |
| Model 7 | Time, Completion rate of pre-class homework | Interaction of time by completion rate of pre-class homework | Completion rate * T3 vs. Completion rate * T1 | −2.20 (−25.31, 20.92) | 0.852 |
| | | | Completion rate * T2 vs. Completion rate * T1 | 2.77 (−19.76, 25.30) | 0.810 |
| | | | Completion rate * T3 vs. Completion rate * T2 | −4.97 (−16.15, 6.22) | 0.384 |
| Model 8 | Time, Class participation | Interaction of time by class participation | Class participation* T3 vs. class participation* T1 | 3.87 (−1.89, 9.62) | 0.188 |
| | | | Class participation* T2 vs. class participation* T1 | −0.58 (−6.64, 5.48) | 0.851 |
| | | | Class participation* T3 vs. class participation* T2 | 4.451 (1.15, 7.75) | 0.008 |

Dependent variable was exam score and within subject effect was time (pre-test, mid-test and post-test) for all eight models of GEE analysis.
T3 vs. T1 indicates a comparison in the score change from pre-test to post-test, T3 vs. T2 indicates the score change from mid-test to post-test.
Abbreviations: T1: pre-test; T2: mid-test; T3: post-test; academic (+): participants with academic position, including lecture, assistant professor, associate professor and professor; academic (−): participants without academic position; administrative (+): participants with administrative position; administrative (−): participants without administrative position.

**Table 3 Univariate and multivariate (by General linear model) analysis to identify variables that significantly predict the post-test scores (T3).**

| | Univariate | | Multivariate | | |
|---|---|---|---|---|---|
| | Statistics | *p* value | β | 95% CI | *p* value |
| Age (yrs) | 0.095[a] | 0.320 | | | |
| Teaching duration (yrs) | 0.138[a] | 0.166 | | | |
| Pre-test (T1) | 0.408[a] | <0.001 | 0.035 | (-0.059, 0.129) | 0.463 |
| Mid-test (T2) | 0.787[a] | <0.001 | 0.761 | (0.631, 0.890) | <0.001 |
| Hours spent on pre-class learning | 0.160[a] | 0.091 | −0.137 | (-0.399, 0.125) | 0.307 |
| Completion rate of pre-class homework | 0.054[a] | 0.583 | | | |
| Class participation | 0.198[a] | 0.034 | 0.306 | (-4.99, 5.60) | 0.910 |
| Gender | | 0.146 | | | |
| Male (n = 42) | 76.5 ± 14.2 | | | | |
| Female (n = 73) | 72.2 ± 15.8 | | | | |
| Administrative position | | 0.332 | | | |
| Yes (n = 12) | 77.8 ± 13.6 | | | | |
| No (n = 103) | 73.3 ± 15.5 | | | | |
| Professional category | | 0.471 | | | |
| Physician (n = 45) | 75.3 ± 14.6 | | | | |
| Pharmacist (n = 12) | 75.5 ± 15.1 | | | | |
| Nurse (n = 32) | 70.5 ± 16.8 | | | | |
| Radiation technician (n = 10) | 79.0 ± 13.6 | | | | |
| Other category (n = 16) | 71.6 ± 15.2 | | | | |
| Academic position | | 0.171 | | | |
| Lecturer or higher (n = 26) | 77.4 ± 13.5 | | | | |
| None (n = 89) | 72.7 ± 15.7 | | | | |
| Session of TOTR workshop | | <0.001 | | | 0.046 |
| July, 2015 (n = 14) | 87.00 ± 8.17 | | Reference category | | |
| Sep, 2015 (n = 13) | 76.31 ± 14.16 | | −1.749 | (-8.193, 4.696) | 0.595 |
| April, 2016 (n = 17) | 79.00 ± 14/00 | | −1.885 | (-9.795, 6.024) | 0.640 |
| July, 2016 (n = 13) | 72.00 ± 12.78[b] | | −7.176 | (-15.860, 1.508) | 0.105 |
| Oct, 2016 (n = 18) | 61.50 ± 11.33[b,c] | | −10.974 | (-19.842, −2.106) | 0.015 |
| Nov, 2016 (n = 13) | 80.08 ± 13.40 | | −7.755 | (-16.555, 1.045) | 0.084 |
| April, 2017 (n = 16) | 72.00 ± 9.90 | | −5.406 | (-14.051 3.238) | 0.220 |
| July, 2017 (n = 11) | 63.09 ± 21.71[b] | | −10.105 | (-18.772, −1.438) | 0.022 |

Univariate analysis was performed by using correlation, t-test and ANOVA analysis. Multivariate analysis was undertaken by general linear model.

[a] Correlatio coefficient.

[b] $p < 0.05$ when compared with July, 2015.

[c] $p < 0.05$ when compared with April, 2016.

findings of Taylor et al. [26] who reported that pre-class exercise in combination with in-class reinforcement improved the student performance. Thus, the addition of mini-lecture in class time is highly warranted when using FC style learning.

Our moderator analysis by GLM revealed that mid-test score is the most significant predictor of final knowledge gain. The results of the mid-test scores might reflect the overall achievement in pre-class learning. Therefore, the significance of mid-test score in predicting final knowledge gain points out the importance of achievement in pre-class learning. Flipped courses require students to obtain course content well before class so that class time can be used for active learning exercises. It is believed that poor pre-class learning will hamper the final effect of FC. To obtain the optimal outcome in FC learning model, effort should be made to maximize the achievement of pre-class learning. Berg et al. [27] claimed that a prerequisite for a successful FC curriculum is that students come to class well-prepared. Coulter et al. [28] demonstrated that mandatory pre-class readings and pre-class quizzes were beneficial to students' examination and quiz grade in a therapeutics course. Lemoncello et al. [29] found a significant correlation between time spent on online pre-class learning tutorials and the final grades. Gross et al. [30] also reported that increased pre-class preparation underlies student outcome improvement in FC. Ihm et al. [16] reported that learning readiness had a significant influence on learning outcomes in multivariate regression analysis. These finding were consistent with our results. The FC model requires completion of pre-reading assignments in preparation for in-class activities. However, students often come to class without preparation [31]. Multiple factors might affect the willing of learner to actively undergo pre-class learning, such as attitude, expectation, motivation and engagement of the learners, easiness and quality of pre-class material, format of pre-class learning, familiarity with FC format and etc. Lieu R et al. [32] created pre-class reading guides that provided students with a way to actively engage with the required reading for each day of class. The results showed that optional pre-class reading guides may help students stay on track to acquire course content in introductory biology and thus result in

improved exam performance. Thus, facilitators should be aware of these confounders and encourage the learner well-prepared in pre-class learning.

The session of TOTR is another significant predictor of final outcome in our GLM analysis. In this study, all the elements of FC, including learning materials and environments, were kept the same in each session of TOTR, except the learners and tutors were different in each session. However, the eight tutors in each TOTR session came from the same TOTR tutor group comprised of 12 experienced teachers. Thus, tutor factor seems to play little role on the different outcome between different sessions. We postulate that the difference in learning outcome between sessions mainly result from the learner factors.

We used hours spent on pre-class learning, completion rate of homework and class participation to evaluate the extent of commitment of the learner. In our univariate analysis, only class participation, but not hours spent on pre-class learning, was a significant predictor of the final knowledge gain. These results contradict to the finding of Lemoncello et al. [29] who showed that time spent on pre-class learning correlated significantly with the final grade. Our results failed to demonstrate the significant impact of time spent on pre-class learning might be due to information error. In this work, the information regarding the hours spent on pre-class learning was subjectively reported by the learner themselves. Such information might be over-estimated and biased. Another possible explanation is that the course design of TOTR does not require more than baseline participation to receive a learning benefit. In multivariate GLM model, all these three indicators were not significant predictor of final outcome. The failure of these indicators to predict the learning outcome might be again due to the reliability and validity of instruments used to measure these potential confounders. .

Despite class participation was not a significant predictor of final outcome, GEE model showed that class participation is a significant moderator influencing the score change from mid-test to post-test. This means that more engagement in classroom activity would cause a better final learning gain. Thus, engagement in classroom activity is an important moderator.

### *Limitation*

There are several weakness in this study. First of the all, this work is limited by the retrospective, non-control design. Without control group, the exact impact of various confounding variable cannot be examined without bias. Second, only a few potential confounding variables proposed by previous investigators were analyzed in this study. The cause for this limitation is due to the fact that this study is a retrospective data analysis. Third, some validated instruments that can effectively measure attitude, motivation, and engagement of learners, such as “student attitude to learning scale”, “motivation and engagement scale” or “motivation assessment scale”, were not used as the evaluation tools in TOTR program. Fourth, the use of scores of MCQ test as the outcome indicator to assess the effectiveness of TOTR might not reflect the real effect of higher cognitive learning outcome of FC. Fifth, the sample size is small in many subgroups. This may cause insufficient power to detect some significant moderators. Finally, the instrument used to measure class participation showed two dimensions in construct validity analysis. However, the second dimension contains only one item. This item was not removed from the measurement scale because it is theoretically important. The inclusion of this item in measuring class participation may pull apart from a single dimension and cause a small amount of measurement error.

### *Future studies*

The use of validated instruments to measure attitude, engagement, motivation, and higher cognitive level learning outcome in future studies might help us to confirm the advantage of FC in higher cognitive outcome and identify the confounders of FC outcome. Also, a non-FC control with the same pedagogy is necessary to examine whether it is FC that is the mechanism for learning or whether it is just the effect of exposure to the content.

## Conclusion

This work confirmed the effectiveness of FC style faculty development learning program TOTR on knowledge transfer of teaching skills for clinical teachers in Taiwan context. Also, we found that achievement in pre-class learning, class participation and learner factors are significant confounders of FC outcome. Based on the findings of this work, we recommend that tutor should be careful in pre-class material preparation and provide necessary assistance in pre-class learning to guarantee a good achievement in pre-class learning. At the same time, tutor should demonstrate good teaching skills to attract students' interest and engagement in classroom activities.

## Funding

This work is supported by grants from Chang Gung Memorial Hospital (grants No. CDRPG2E0011 and CDRPG2E0012).

## Ethical approval

The Institutional Review Board of the Chang Gung Memorial Hospital approved this study (IRB No: 104-3749B).

## Conflicts of Interest

The authors declare no conflicts of interest.

## REFERENCES

[1] Baker JW. The "classroom flip": using Web course management tools to become the guide on the side. *Selected Papers from the 11th International Conference on College Teaching and Learning*. Florida: Jacksonville; 2000. p. 9–17.

[2] Lage M, Platt GJ, Treglia M. Inverting the classroom: a gateway to creating an inclusive learning environment. J Econ Educ 2000;31:30–43.
[3] Bergmann J, Sams A. Flip your classroom: Reach every student in every class every day. Eugene: International Society for Technology in Education; 2012.
[4] Pierce R, Fox J. Vodcasts and active-learning exercises in a "flipped classroom" model of a renal pharmacotherapy module. Am J Pharmaceut Educ 2012;76:196.
[5] DeLozier SJ, Rhodes MG. Flipped classrooms: a review of key ideas and recommendations for practice. Educ Psychol Rev 2017;29:141–51.
[6] Karabulut-Ilgu A, Cherrez NJ, Jahren CT. A systematic review of research on the flipped learning method in engineering education. Br J Educ Technol 2018;49:398–411.
[7] Liu L, Ripley D, Lee A. Flipped learning and influential factors: case analysis. J Educ Techno 2016;9:85–103.
[8] Bishop JL, Verleger MA. The flipped classroom: a survey of the research. Atlanta, United States: 120th ASEE Annual Conference & Exposition; 2013.
[9] Bonnes SL, Ratelle JT, Halvorsen AJ, Carter KJ, Hafdahl LT, Wang AT, et al. Flipping the quality improvement classroom in residency education. Acad Med 2017;92:101–7.
[10] Geist MJ, Larimore D, Rawiszer H, Al Sager AW. Flipped versus traditional instruction and achievement in a baccalaureate nursing pharmacology course. Nurs Educ Perspect 2015;36:114–5.
[11] Harrington SA, Vanden Bosch M, Schoofs N, Beel-Bates C, Anderson K. Quantitative outcomes for nursing students in a flipped classroom. Nurs Educ Perspect 2015;36:179–81.
[12] McLaughlin JE, Roth MT, Glatt DM, Gharkholonarehe N, Davidson CA, Griffin LM, et al. The flipped classroom: a course redesign to foster learning and engagement in a health professions school. Acad Med 2014;89:236–43.
[13] Tune JD, Sturek M, Basile DP. Flipped classroom model improves graduate student performance in cardiovascular, respiratory, and renal physiology. Adv Physiol Educ 2013;37:316–20.
[14] Jesurasa A, Mackenzie K, Jordan H, Goyder EC. What factors facilitate the engagement with flipped classrooms used in the preparation for postgraduate medical membership examinations? Adv Med Educ Pract 2017;8:419–26.
[15] Shinaberger L. Components of a flipped classroom influencing student success in an undergraduate business statistics course. J Stat Educ 2017;25:122–30.
[16] Ihm J, Choi H, Roh S. Flipped-learning course design and evaluation through student self-assessment in a predental science class. Korean J Med Educ 2017;29:93–100.
[17] Chen KS, Monrouxe L, Lu YH, Jenq CC, Chang YJ, Chang YC, et al. Academic outcomes of flipped classroom learning: a meta-analysis. Med Educ 2018;52:910–24.
[18] McLaughlin JE. Flipped classrooms, by design. Med Educ 2018;52:887–8.
[19] Hotle SL, Garrow LA. Effects of the traditional and flipped classrooms on undergraduate student opinions and success. J Prof Issues Eng Educ Pract 2016;142:05015005.
[20] Moranski K, Henery A. Helping learners to orient to the inverted or flipped language classroom: mediation via informational video. Foreign Lang Ann 2017;50:285–305.
[21] Day LJ. A gross anatomy flipped classroom effects performance, retention, and higher-level thinking in lower performing students. Anat Sci Educ 2018;11:565–74.
[22] Giuliano CA, Moser LR. Evaluation of a flipped drug literature evaluation course. Am J Pharmaceut Educ 2016;80:66.
[23] Persky A. Qualitative analysis of animation versus reading for pre-class preparation in a "flipped" classroom. J Excel Coll Teach 2015;26:5–28.
[24] Nawi N, Jawawi R, Matzin R, Jaidin JH, Shahrill M, Mundia L. To flip or not to flip: the challenges and benefits of using flipped classroom in geography lessons in Brunei Darussalam. Rev Eur Stud 2015;7:133–45.
[25] Brown T, Albert E, Catchpole M, Lake F. Teaching on the run-general practice training between consultations. Aust Fam Physician 2005;34:47–50.
[26] Taylor AT, Olofson EL, Novak WR. Enhancing student retention of prerequisite knowledge through pre-class activities and in-class reinforcement. Biochem Mol Biol Educ 2017;45:97–104.
[27] Berg IETvd. *Favoured pre-class preparation and associated learning strategies in a flipped classroom curriculum. AMEE.* Barcelona, Spain: AMEE 2016 Abstract Book; 2016. p. 27528.
[28] Coulter CJ, Smith S. The impact of preclass reading assignments on class performance. Curr Pharm Teach Learn 2012;4:109–12.
[29] Lemoncello R. Blended, active learning for anatomy & physiology: development & program evaluation. Perspectives on Issues in Higher Education 2015;18:62–75.
[30] Gross D, Pietri ES, Anderson G, Moyano-Camihort K, Graham MJ. Increased preclass preparation underlies student outcome improvement in the flipped classroom. CBE Life Sci Educ 2015;14:ar36.
[31] Alpaslan S, Baki C, Yunus EZ. Flipping a College calculus course: a case study. Edu Tech Soc 2015;18:142–52.
[32] Lieu R, Wong A, Asefirad A, Shaffer JF. Improving exam performance in introductory biology through the use of preclass reading guides. CBE Life Sci Educ 2017;16:ar46.