Medical Teacher

ISSN: 0142-159X (Print) 1466-187X (Online) Journal homepage: informahealthcare.com/journals/imte20

PBL curriculum improves medical students' participation in small-group tutorials

Y. T. Wun, Eileen Y. Y. Tse, T. P. Lam & Cindy L. K. Lam

To cite this article: Y. T. Wun, Eileen Y. Y. Tse, T. P. Lam & Cindy L. K. Lam (2007) PBL curriculum improves medical students' participation in small-group tutorials, Medical Teacher, 29:6, e198-e203, DOI: 10.1080/01421590701294349

To link to this article: https://doi.org/10.1080/01421590701294349

2007; 29: e198–e203 MEDICAL TEACHER

WEB PAPER

PBL curriculum improves medical students' participation in small-group tutorials

Y. T. WUN, EILEEN Y. Y. TSE, T. P. LAM & CINDY L. K. LAM
Family Medicine Unit, The University of Hong Kong, Hong Kong

Abstract

Background: Group learning is the core of problem-based learning (PBL) but has not been extensively studied, especially in Asian students.

Methods: This study compared students of PBL and non-PBL curricula in students' talking time and participation in small-group tutorials in a medical school in Asia. The proportions of student talking of 46 tutorials in three teaching rotations of the PBL curriculum and those of 43 corresponding tutorials in the non-PBL curriculum were counted. Twelve videotapes of tutorials (six from each curriculum), stratified for tutor, case scenario and students' learning stage, were randomly selected and transcribed. They were rated with the group-interaction (5 items) and active-participation (four items) tutorial assessment scales developed by Valle et al. These outcomes were compared between the students of PBL and non-PBL curricula.

Results: Students from the PBL curriculum talked significantly more. In only two (4.7%) of 43 tutorials in the non-PBL curriculum did the students talk more than the tutors; but students talked more than the tutors in 17 (37.0%) of 46 tutorials in the PBL curriculum. PBL students scored significantly higher than non-PBL students in all items except one item (respect to peers) of the tutorial assessment scales, and in the mean scores of both the group interaction scale (items 1–5) and the active participation scale (items 6–9).

Conclusions: The results suggested that PBL starting from the early years of a medical curriculum was associated with more active student participation, interaction and collaboration in small-group tutorials.

Introduction

Problem-based learning (PBL) enables students to learn group-work skills and attitudes, and improves their communication skills (Wood 2003). These skills and attitudes include teamwork, cooperation, respect for colleagues' views, chairing a group, and interaction with group members (Wood 2003). Hence PBL is thought to be good for group learning (Dolmans et al. 2005). PBL medical graduates are more likely than non-PBL graduates to indicate that they have learned communications skills and teamwork (Prince et al. 2005). However, there is not much research evidence on the effect of PBL on group learning, especially among students in Asia. In fact, one study observed that undergraduates in psychiatry had no difference in their learning style whether they were in a PBL or non-PBL curriculum (McParland et al. 2004).

It is often difficult to get students to work in a group or as a team (Distlehorst et al. 2005; Walton et al. 1997). There is also a general impression that group learning is even more difficult in Asian students who tend to be passive and unwilling to challenge others. Our previous study on consecutive videotaped tutorials of medical students in Hong Kong in 1995 (Dixon et al. 1997) seemed to support this hypothesis. We found that students were not much engaged in discussion, which was connected mainly through the tutor. In one third of the tutorials, the students spoke less than 30% of the time, in another one third 31–40%, and the remaining one third more than 40%. In only two out of 58 tutorials did the students spoke more than the tutors. This pattern was irrespective of whether the tutorials were conducted in the students' mother tongue (Cantonese) or in English. On the other hand, Kember & Gow (1996) argued that the apparent stereotype of passive, superficial and rote learning behaviour of Asian students might be explained by the teaching method and environment rather than inherent characteristics.

Since September 1997, our medical school adopted PBL starting from year one of the curriculum. This provided an opportunity to compare students of the new PBL-curriculum with those of the traditional non-PBL curriculum in their performance in small-group tutorials. The aim of the study was

Practice points

- Asian students from a non-PBL medical curriculum participated little in discussion in small-group tutorials.
- PBL introduced in the early years of a medical curriculum could enable Asian students to participate more actively in small-group tutorials.
- Objective criterion-based assessment of student performance in small-group tutorials is feasible.
- Further research on the effect of a PBL curriculum on the quality of student discussion and learning is required.

Correspondence: Cindy L. K. Lam, Family Medicine Unit, the University of Hong Kong, 3rd Floor, Ap Lei Chau Clinic, 161 Main Street, Ap Lei Chau, Hong Kong. Tel: 852-25185653; fax: 852-28147475; email: clklam@hku.hk

ISSN 0142–159X print/ISSN 1466–187X online/07/060198–6 © 2007 Informa UK Ltd.
DOI: 10.1080/01421590701294349

to evaluate the effects of PBL on student participation and interaction in small-group tutorials.

Methods

Study setting and subjects

Before 1997, our medical school had a traditional curriculum that was lecture and discipline based. The curriculum was revised with PBL playing a significant part since the first year in 1997 but the student selection process had not changed. Since 1994, our final-year medical students during their family medicine clerkship had weekly small-group (seven to eight students) tutorials based on scenarios of common problems in family practice. These problem-based tutorials differed from the usual PBL tutorials in that the problems were mostly dealt with in one session (for clinical analysis and management) with the main aim for the students to learn problem-solving skills and a secondary aim of generating new learning objectives. The family medicine clerkship continued the problem-based tutorials with the same problem scenarios until 2002 when the first batch of students of the new PBL curriculum had the family medicine clerkship.

Data collection

Forty-eight tutorials over three clerkship rotations (with four student groups in each rotation and four tutorials per group) in 2002 (PBL group) were videotaped. Two trained research assistants, after standardization on three tapes, viewed separately these tapes and recorded whether it was the student or tutor talking at each one-minute interval. The one-minute interval was used after a pilot test showing that counting at half-minute intervals was technically difficult and gave similar results to those obtained at one-minute intervals. The tutorials lasted for different lengths of time and there were silence at some one-minute stops. The student and tutor talking proportions of each tutorial were expressed as the percentages of the total talking incidents. Silent incidents were not counted. The method of counting the proportion of student or tutor talking was the same as that used in our previous study in 1995 (Dixon et al. 1997). Similar data on the proportions of student and tutor talking in the tutorials of three matched clerkship rotations in the academic year of 1995 (non-PBL group) were retrieved from the data collected in our previous study (Dixon et al. 1997).

Students' group interaction and participation were assessed qualitatively in six tutorials randomly selected from each of 1995 and 2002 batches stratified for the tutor, case scenario, and student stage of training (i.e. the students had undergone similar specialty clerkships before the tutorials took place, to minimize the effects of knowledge on students' performance in group tutorials). The videotapes of these tutorials were transcribed in verbatim and each student's performance was assessed with a rating form adapted from Valle's tutorial assessment questionnaire developed and validated by a team of physicians trained in PBL teaching methods as part of a project to evaluate PBL outcomes at the National Autonomous University of Mexico (Valle et al. 1999). The original Valle's assessment form consists of 24 items in four scales: independent study, group interaction, reasoning skills, and active participation. The four scales had demonstrated internal construct validity with scale inter-item correlations ranging from 0.59–0.88, and adequate internal reliability Cronbach alphas of greater than 0.8. (Valle et al. 1999) Since we aimed to assess only students' participation and collaboration, we used only the rating scales on group interaction and active participation, which consist of nine items as shown in Appendix A. Items 1 to 5 formed the group interaction scale (ability to communicate and fit into the group) while items 6 to 9 formed the active participation scale. Each item was rated on a 6-point Likert scale from '1' (never) to '6' (always). A pilot study was done to confirm the applicability of Valle's items and to standardize the assessment criteria on the transcripts and videotapes of two tutorials that were not included in the main study were each assessed by the rater of the main study and a teacher who was experienced in PBL teaching. Valle's items were found to be applicable and there was good inter-rater correlation in the scores.

The 12 tutorials included in the main study were all rated by a trainer in Family Medicine who had extensive experience in undergraduate teaching but had not taught in any of the curricula under study. He first rated the performance of the students on all the items except item 4 by reading through the transcripts in which the identity of the year (1995 or 2002) was concealed. Since item 4 (listens attentively to other members of the group) could not be rated from the transcripts, the rater reviewed the videotape to score the students on this item after he had scored all the other items. The scores of the other items were not to be altered.

Data analysis

Students' participation in the tutorials was measured by the students' talking proportions. We expected that students' talking proportion would be less than 40% in two-thirds of the non-PBL groups (Dixon et al. 1997). Hence, we compared the proportion of tutorials between the two study years by student talking proportions of <20%, 21–30%, 31–40%, 41–50% and above 50% to see if the students from the PBL curriculum made any improvement. We used the Fisher's Exact test to analyse the statistical difference between the two groups.

For Valle's assessment scales, we analysed the scores of each item and scale as ordinal data and used the Mann–Whitney U test to detect any statistical difference between the non-PBL and PBL students. We also compared the mean scores by tutorials, in which we took the mean of the scores of all students in each tutorial as the tutorial score. Mann–Whitney U test was also used to analyse the difference between tutorials from the two years. Two-tailed p-values <0.05 were considered statistically significant. We used the Cronbach alpha to evaluate the internal consistency of the items in each scale. All data analyses were done by the SPSS for Windows 11.0 programme (SPSS Inc. Chicago).

Results

Out of 48 tutorials in Year 1995, the data on the proportion of student talking could be retrieved in only 43. Of the 48

videotaped tutorials of Year 2002, one tutorial was not completely recorded and the sound quality in another was too poor for assessment, leaving 46 tapes in the final analysis. Forty-four students in 1995 and 42 students in 2002 were assessed on 12 videotapes (six randomly selected from each year) with Valle's scales.

(a) Talking proportion in a tutorial

Table 1 shows the distribution of the proportions of student talking by years. In only two (4.7%) tutorials did the students of 1995 talked more than the tutors; but students of 2002 talked more than the tutors in 17 (37.0%) of the 46 tutorials. The difference was statistically significant ($P < 0.001$, Fisher's Exact Test). Students talked more than 40% in 16 (37.2%) out of the 43 tutorials in 1995, while students talked more than 40% in 29 (63.0%) out of 46 tutorials in 2002, the difference was statistically significant ($P = 0.01$, Fisher's Exact Test).

(b) Interaction and participation in tutorials

Table 2 shows the tutorial assessment scores of the tutorials by clerkship year. The tutorials of 2002 scored higher than those of 1995 in all items except item 5 (respect to peers). The differences were statistically significant (either two-tailed or one-tailed) for item 3 (adjusts to different group roles), item 6 (helps his/her peers to clarify ideas), and item 7 (participates in case discussions). The tutorials of 2002 had higher mean scores for both the group interaction scale (items 1–5) as well as the active participation scale (items 6–9) than the tutorials of 1995. The differences were statistically significant.

The mean scores of students of 2002 were higher those of students of 1995 in all items except item 5 (respect to peers) and both scales, when they were analysed as individuals. In contrast to the analysis by tutorials (Table 2), the differences between students of the two study years reached statistical significance (Table 3) in all items except item 5. The mean scores for item 5 were similar for students of both years.

The internal consistencies of the Valle's group interaction (items 1–5) scale, the active participation scale (items 6–9), and all 9 items as a whole were assessed. For all 86 students, the Cronbach alphas for group interaction, active participation, and all 9 items were 0.40, 0.70, and 0.72, respectively.

Discussions

This study evaluated the performance of students in small-group tutorials objectively and cross-validated the results with

Table 1. Distribution of talking proportions in tutorials.

	Number (%) of 1995 tutorials ($n = 43$)		Number (%) of 2002 tutorials ($n = 46$)	
Talking proportions	Teacher talking	Student talking	Teacher talking	Student talking
0–20%	0	4 (9.3%)	0	2 (4.4%)
21–30%	0	8 (18.6%)	1 (2.2%)	4 (8.7%)
31–40%	0	15 (34.9%)	4 (8.7%)	11 (23.9%)
41–50%	2 (4.7%)	14 (32.6%)	13 (28.3%)	12 (26.1%)
51–60%	16 (37.2%)	2 (4.7%)	14 (30.4%)	13 (28.3%))
61–70%	14 (32.6%)	0	11 (23.9%)	3 (6.5%)
71–80%	9 (20.9%)	0	1 (2.2%)	1 (2.2%)
81–90%	2 (4.7%)	0	2 (4.3%)	0
>90%	0	0	0	0

Table 2. Tutorial mean Valle's tutorial assessment item and scale scores.

Year		#1	#2	#3	#4	#5	#6	#7	#8	#9	Group interaction scale (#1–#5)	Active participation Scale (#6–#9)
1995	Mean	1.33	1.68	2.01	4.27	3.68	2.44	4.10	2.44	2.21	12.98	11.19
($n = 6$)	(SD)	(0.817)	(1.182)	(0.818)	(0.452)	(0.312)	(0.312)	(0.211)	(0.752)	(0.950)	(2.020)	(2.047)
	Range	1.0–3.0	1.0–4.0	1.0–3.0	3.6–4.8	3.3–4.0	1.3–3.4	3.9–4.4	1.6–3.4	1.5–3.9	11.0–16.3	8.75–13.86
2002	Mean	2.14	3.14	3.10	4.76	3.57	3.57	4.79	3.24	3.00	16.71	14.71
($n = 6$)	(SD)	(0.994)	(1.565)	(0.401)	(0.492)	(0.823)	(0.823)	(0.235)	(0.942)	(0.571)	(1.740)	(1.454)
	Range	1.0–3.9	1.0–5.0	2.4–3.6	4.4–5.7	2.1–4.4	2.9–4.3	4.4–5.1	1.4–4.1	2.3–4.0	14.43–19.71	12.71–17.00
MWU		7.5	8.0	4.0	8.0	15.5	3.0	0.5	8.0	7.5	2.0	2.0
P		0.073	0.102	0.024	0.107	0.686	0.016	0.005	0.107	0.092	0.010	0.010

SD = standard deviation; MWU = Mann–Whitney U statistic; P = P value for Mann–Whitney U test.

Table 3. Student mean Valle's tutorial assessment item and scale scores.

Year		#1	#2	#3	#4	#5	#6	#7	#8	#9	Group interaction scale (#1–#5)	Active participation scale (#6–#9)
1995	Sum	60	73	87	189	162	106	180	108	96	571	490
($n=44$)	Mean	1.36	1.66	1.98	4.30	3.68	2.41	4.09	2.45	2.18	12.98	11.14
	(SD)	(0.810)	(1.219)	(1.562)	(0.851)	(0.639)	(1.530)	(1.074)	(1.532)	(1.603)	(2.538)	(4.067)
	Range	1–4	1–4	1–5	2–6	2–5	1–6	2–6	1–5	1–5	9–19	5–21
2002	Sum	150	205	217	389	312	261	381	244	222	1273	1108
($n=42$)	Mean	2.14	3.14	3.10	4.76	3.57	3.69	4.79	3.24	3.00	16.71	14.7
	(SD)	(1.389)	(1.523)	(1.605)	(0.932)	(1.085)	(1.297)	(1.094)	(1.411)	(1.593)	(3.330)	(3.744)
	Range	1–4	1–6	1–6	2–6	2–6	1–5	3–6	1–5	1–5	11–25	9–21
MWU		654.0	461.5	608.5	667.0	887.5	471.5	604.0	668.0	675.0	346.0	484.5
P		0.004	<0.001	0.003	0.018	0.723	<0.001	0.004	0.019	0.021	<0.001	<0.001

SD = standard deviation; MWU = Mann–Whitney U statistic; P = P value for Mann–Whitney U test.

both quantitative and qualitative methods. The same rater was used in the qualitative assessment of student performance in all tutorials in order to achieve greater consistency in scoring. The use of two raters for each tape might improve the reliability of the results but it was very difficult to identify another person who was experienced in PBL and had not participated in the two curricula. Other studies on rating of videotaped consultations showed that a single rater could achieve an acceptable reliability (Ram et al. 1999), with generalizability coefficient of 0.82 for rating 12 tapes (Hays et al. 2002). Although there could be personal bias, we tried to minimize this by blinding and the use of a structured assessment form. The bias, if any, would be more likely to affect the absolute scores and should be similar for both groups of students. We aimed to find out the relative difference between PBL and non-PBL students rather than the absolute scores of individual students, therefore the bias of the rater should not have affected the conclusion of the results.

In this study, students of the PBL curriculum, compared with students of the non-PBL, had more active participation and interaction among themselves during small-group tutorials as shown by the observation that they talked more during tutorials. This was further supported by the evaluation with Valle's assessment scales, particularly in the aspects of adjustment of individual roles in the group, helping peers to clarify ideas/opinions, and active participation in discussion.

We tried to evaluate the group function during tutorials and found that tutorials of PBL students had higher mean assessment scores than those of non-PBL students (Table 2). Because of the small sample size of 12 tutorials, the improvement in the tutorial assessment scores did not reach statistical significance except in three items. When the students were assessed as individuals, the improvement reached statistical significance due to a larger sample size of 86 (Table 3). The majority of the mean scores for the 1995 non-PBL students were below 3.0 (lower half of the scale) out of the Likert scales of 1–6. On the contrary, the mean scores for the 2002 PBL students were above 3.0 that could be considered as satisfactory. 'Respect to peers' was the only item in which students did not show any improvement. We had not looked into the cause of this and this could be a topic for further exploration with students. One possible explanation could be that our students were reluctant to challenge or confront each other because it is considered socially improper to do so in the Asian culture.

In Valle's original study all the assessment scales had Cronbach alphas above the group comparison standard of 0.7. The alpha of the group interaction scale (items 1 to 5) was 0.83 but this scale showed sub-optimal internal reliability (Cronbach alpha = 0.40) in our study, which suggested that the items might not measure the same domain. We carried out factor analysis on the scores of these five items and found three components: item 1, items 2 + 3, and items 4 + 5, but the Cronbach alphas were still low (0.50 and 0.53 respectively for the last two components). We could not find in MEDLINE and EMBASE other studies on the psychometrics of Valle's tutorial assessment scales except the original; therefore we cannot be certain whether this finding was related to the instrument or unique to our students. Further research is needed to confirm the validity and reliability of Valle's tutorial assessment scale on group interaction. Until more information on the psychometrics of this scale is available, the results from these five items should be interpreted individually instead of as a summative scale.

The results of our study supported Kember and Gow's hypothesis that teaching method could mould Asian students' learning behaviour (Kember & Gow 1996). The change from a traditional to a PBL curriculum in our medical school seemed to have changed the behaviour and attitudes of Chinese students in tutorials and enabled them to be more active and collaborative.

Although it was possible that the difference in students' performance between the tutorials in 2002 and those in 1995 could be the result of a difference in the training and experience of the tutors in the two years independent of any effect from the curricula, this was unlikely because the same three academic staff were the tutors of more than three quarters of the tutorials in both years and all of them had completed formal PBL and small-group learning training courses at McMaster University before the 1995 tutorials took place. Furthermore our earlier study (Dixon et al. 1997) did not find any difference between tutorials by the tutor who had

more than 15 years of experience and those by the tutors who had 3 to 5 years of experience.

This study suggested that PBL starting from the early years of the curriculum could improve the group learning process. PBL students demonstrated better collaboration in discussion by adjusting their individual roles during group discussion, helping their colleagues to clarify ideas/opinions, and participating more in discussion, as indicated by significantly higher scores in items 3, 6, and 7 of Valle's assessment scales. Though this study did not measure the effectiveness of learning directly, others have shown that learning is more effective if there is collaboration among group members (Paice & Heard 2003). Therefore we believe that students of the PBL curriculum are more likely to have more effective learning than students of non-PBL curriculum.

Limitations

This is a cross-sectional study with a historical-control group instead of a randomised controlled trial, which might not have controlled for all the confounding factors. Although both groups were final-year students with similar clinical exposure, students of seven years apart might have acquired different learning behaviours due to differences in parental attitudes and changes in the secondary school education approach. There could also be cultural changes in both society and the medical school.

This study focused on student participation in small-group learning and did not explore some other important components of PBL such as identifying knowledge gaps, setting learning objectives, independent information search, and sharing new information. The study did not assess the nature and quality of students' talking, an increase in the quantity of participation in small-group tutorial did not necessarily imply that the quality of the content of discussion, was also better.

Conclusion

This study compared the performance in small-group tutorials between students of a PBL curriculum and students of a traditional non-PBL curriculum in a medical school in Asia. We found that PBL students talked significantly more in the tutorials than non PBL students. Qualitative assessment also showed that PBL students were more interactive and participatory in tutorials. PBL starting from the early years of a medical curriculum might enable Asian medical students to be more collaborative and active in small-group learning. Further research is required to confirm the results of this historical case control study and to determine the effect of PBL on the quality of discussion in small group tutorials.

Notes on contributors

All authors are Specialists in Family Medicine and members of the Family Medicine Unit, The University of Hong Kong.

Y. T. WUN was responsible for the rating of tutorials, and a main contributor in data analysis and the writing of this paper.

E. TSE contributed in the planning and supervision of the project, data analysis, and writing of this paper.

T. P. LAM contributed in the planning of the project and writing of this paper.

C. L. K. LAM contributed to the planning of the project, data analysis, and writing of this paper.

References

Distlehorst LH, Dawson E, Robbs RS, Barrows HS. 2005. Problem-based learning outcomes: the glass half-full. Acad Med 80:294–299.

Dixon AS, Lam CLK, Lam TP, Ho R. 1977. Can Hong Kong students cope with small-group, problem-based learning? in: J. Conway, R. Fisher, L. Sheridan-Burns & G. Ryan. (Eds), *Research and Development in Problem-based Learning*, Vol. 4, pp. 119–126, (Newcastle, Australia, The Australian Problem-based Learning Network).

Dolmans DHJM, De Grave W, Wolfhagen IHAP, Van Der Vleuten CPM. 2005. Problem-based learning: future challenges for educational practice and research. Med Educ 39:732–741.

Hays R, Spike N, Gupta TS, Hollis J, Veitch J. 2002. A performance assessment module for experienced general practitioners. Med Educ 36:258–260.

Kember D, Gow L 1996. A challenge to the anecdotal stereotype of the Asian student. Stud High Educ 161:117–128.

Mcparland M, Lorraine NM, Gill L. 2004. The effectiveness of problem-based learning compared to traditional teaching in undergraduate psychiatry. Med Educ 38:859–867.

Paice E., Heard S. 2003. Collaborative learning. Med Educ 37:758–759.

Prince KJA, Van Eijs PWLJ, Boshuizen HPA, Van Der Vleuten CPM., Scherpbier AJJA. 2005. General competencies of problem-based learning (PBL) and non-PBL graduates. Med Educ 39:394–401.

Ram P, Grol R, Rethans JJ, Schouten B, Van Der Vleuten C, Kester A. 1999. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. Med Educ 33:447–454.

Valle R, Petra I, Martinez-Gonzalex A, Jose Antonio Rojas-Ramirez Ja, Morales-Lopez S, Pina-Garza B. 1999. Assessment of student performance in problem-based learning tutorial sessions. Med Educ 33:818–822.

Walton JN, Clark DC, Glick N. 1997. An outcomes assessment of a hybrid-PBL course in treatment planning. J Dental Educ 61:361–367.

Wood DF. 2003. Problem based learning. Brit Med J 326:328–330.

Appendix A: Valle's tutorial assessment rating scales on group interaction and active participation

(Some items were further defined for standardization in this study, as shown in the brackets)

Group interaction scale

1. Accepts suggestions about his/her work
2. Accepts decisions made by the group ('Yes', no counter-decisions, excluding counter-suggestions)
3. Adjusts to different group roles (Leader role: changes discussion topic; helps to define learning objective. Facilitator role: e.g. 'So, . . .')
4. Listens attentively to other members of the group (Eye contact, body language of attentiveness)
5. Shows respect to his/her peers (Acknowledge other students' contribution, no interruption, no personal attack)

Active participation scale

6. Helps his/her peers to clarify ideas (Immediate follow-up statements to clarify another student's statement without tutor's facilitation. Gives further factual knowledge.)
7. Participates in case discussions (Initiates questions/directions/topics of discussion, excluding response to tutor's direct questions. Joins in discussion, e.g., Tutor-Student1-Student2-Student7-Tutor. Suggestive utterance, e.g., 'yeh', 'ah', 'yes'. Exclude behaviours described by Item 8 and Item 9)
8. Shares knowledge with the group (Gives his/her previous readings. Volunteers his/her knowledge)
9. Gives feedback (Silences a disordered discussion, e.g., 'shih'. Reminds others about the discussion direction)