Surgery 165 (2019) 1069–1074

Contents lists available at ScienceDirect

Surgery

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

Quality with quantity? Evaluating interprofessional faculty prebriefs and debriefs for simulation training using video

John T. Paige, MD[a,*], John B. Zamjahn, PhD[b], Raquel B. Carvalho, DDS, PhD[c], Shuang Yang, MS[d], Qingzhao Yu, PhD[d], Deborah D. Garbee, PhD[e], Vladimir J. Kiselov, MD[a], Vadym Rusnak, MD[f], Laura Bonanno, DNP[e], Catherine Callan, MS[d]

[a] Department of Surgery, LSU Health New Orleans School of Medicine, New Orleans, LA
[b] Department of Cardiopulmonary Science, LSU Health New Orleans School of Allied Health Professions, New Orleans, LA
[c] Department of Social Medicine, Federal University of Espirito Santo, Vitoria, ES, Brazil
[d] Department of Biostatistics, LSU Health New Orleans School of Public Health, New Orleans, LA
[e] LSU Health New Orleans School of Nursing, New Orleans, LA
[f] Department of Internal Medicine, Tulane School of Medicine, New Orleans, LA

ARTICLE INFO

Article history:
Accepted 17 January 2019
Available online 11 April 2019

ABSTRACT

Background: High-fidelity simulation-based training is used increasingly for prelicensure student teams. Such sessions rely on faculty who are able to provide quality prebriefing and debriefing to foster learning among participants. We investigated how well faculty conducted prebriefing and debriefing as part of high-fidelity simulation-based training for interprofessional education.
Methods: Two trained observers independently rated 38 video-recorded sessions of combinations of 4 faculty conducting prebriefings and debriefings of prelicensure student teams after high-fidelity simulation-based training. Assessment was undertaken using the Objective Structured Assessment of Debriefing, an 8-item tool using a 5-point Likert scale (1 as minimum and 5 as maximum). Mean scores for each item were calculated. Inter-rater agreement was determined using Cohen's kappa. A one-way between-subjects analysis of variance with post-hoc Tukey's studentized range procedure was conducted to compare the effect of team facilitator grouping on the quality of team performance of each Objective Structured Assessment of Debriefing element during a prebriefing or a debriefing. Trend analyses of teams with 4 or more observations were performed using Kendall's Tau coefficient test and linear regression analyses to identify whether teams showed improvement through time. Statistical significance was set at $P < .05$.
Results: A total of 7 combinations of faculty conducted between 1 to 14 prebriefings or debriefings. In general, faculty combinations performed better during debriefings compared with prebriefings, with only 1 team having 1 mean item score <3.50. Statistically significant differences between faculty combinations in mean item scores was more pronounced during the prebriefings (2 of 3 Objective Structured Assessment of Debriefing items rated) than during debriefings (1 of 8 Objective Structured Assessment of Debriefing items rated). Effect sizes were strong for all differences. Linear regression analysis revealed a statistically significant change through time for the 3 rated prebriefing items and for 7 of the 8 rated debriefing items.
Conclusion: Interprofessional faculty combinations in this study tended to have good quality prebriefings and debriefings. The quality of the prebriefings and debriefings can, however, be influenced by the composition of the facilitator teams, most prominently for prebriefings, and team performance does appear to change through time, especially during the debriefing. Future work will focus on whether the quality of prebriefings and debriefings influences learning by trainees.

Introduction

Quality care in today's dynamic health care environment requires effective teamwork among the ever-increasing number of different professionals who participate in a patient's care. More often than not, however, the tribalism[1] and silo mentality[2]

Presented at the 2018 American College of Surgeons Surgical Simulation Summit.
* Reprint requests: John T. Paige, MD, Department of Surgery, LSU Health New Orleans School of Medicine, 1542 Tulane Ave, Rm 734, New Orleans, LA 70112.
E-mail address: JPaige@lsuhsc.edu (J.T. Paige).

https://doi.org/10.1016/j.surg.2019.01.014
0039-6060/

characteristic of the interactions between professions in the clinical environment impede the reliability of team behavior. For example, both transfer of information and failures in communication occur through the entire continuum of the surgical care of a patient, most commonly during the phases of preprocedural teamwork and postoperative hand over.[3] Not surprisingly, the Institute of Medicine (IOM)[4] and the Joint Commission[5] have focused on promoting better teamwork to address this issue.

One approach to improving teamwork is to promote interprofessional education (IPE), which has demonstrated benefits at all levels of practice.[6] An attractive target for IPE is at the prelicensure level to prevent formation of the mental barriers and destructive attitudes that have developed traditionally through the so-called hidden curriculum, as students from different professions learn and train separately. In particular, high fidelity (HF), simulation-based training (SBT) of interprofessional teams of prelicensure students, because pf its immersive, experiential nature, has the potential to embed key competencies and attitudes related to teamwork and collaboration, as demonstrated in the literature.[7]

Key to the success of any prelicensure student SBT activity is the quality of the prebriefing (PB) and debriefing (DB) that faculty conduct for each scenario.[8] Each PD and DB is dependent on multiple variables (eg, the individual or combination of individuals conducting the PB and DB, the learners participating, the environment in which the PB and DB takes place, the time of day, and changing characteristics of the facilitators through time). Until recently, determining such quality remained challenging because of a lack of reliable and valid instruments of debriefing assessment. Fortunately, several tools have now emerged,[9–12] allowing for research looking at aspects of PBs and DBs.[13–15] This report expands on these beginnings by investigating whether differences exist for PBs and DBs among faculty teams after an HF SBT for IPE of prelicensure students and analyzing potential causes for any differences in the quality of PBs and DBs.

Methods

Study design

Retrospective analysis of prospectively collected video recordings of faculty team DBs were conducted as part of HF SBT involving IPE of teams of prelicensure students. The institutional review board approved this work as part of an existing exempt protocol.

Format of simulation-based, interprofessional team training

Faculty PB and DB occurred as part of an HF, SBT IPE activity involving prelicensure health professional students from the Schools of Nursing and Medicine at LSU Health New Orleans Health Sciences Center (LA). The format of this training has been described in depth elsewhere.[7] In brief, third-year medical students in the surgery clerkship and senior undergraduate nursing students on an intensive care rotation underwent the HF SBT using a dual scenario format. Scenarios focused on the initial trauma resuscitation of one of two patients: (1) a victim of a major burn injury with inhalational injury, or (2) a patient with blunt trauma with an unstable pelvic fracture and intra-abdominal hemorrhage from the spleen.

Each SBT session opened with a faculty-led PB in which faculty introduced themselves, oriented learners to the simulated environment, and established the learning objectives and ground rules for the educational intervention. The students then proceeded to participate in the first trauma simulation-based scenario. Immediately after this scenario, faculty led an after-action DB, focusing on Advanced Trauma Life Support (ATLS) protocols for resuscitation of trauma patients and useful, team-based competencies for ensuring highly reliable collaboration and care. Learners then participated in a second scenario with a final, after-action DB. All sessions were video recorded and the files saved on a protected server for later review.

Various combinations of faculty teams consisting of between one to three faculty facilitators (four separate individuals, designated faculty A, B, C, and D) conducted the PB or DB for the SBT sessions. Among the facilitators, two were physicians (one surgeon, one internist), and two were nursing professionals with advanced degrees. Faculty members A and C had experience together conducting joint debriefs. None had had formalized training in any particular style of debriefing. For this project, observers (J.Z. and R.B.) reviewed video of the PB conducted at the beginning of each SBT session and the first after-action DB and evaluated the faculty team overall for quality.

Evaluation instrument

Observers used an electronic version of the Objective Structured Assessment of Debriefing (OSAD)[9] to rate the video-recorded faculty PBs and DBs. The OSAD is an evidence-based, user-informed[16] assessment tool originally designed for use in surgery. Its utility has resulted in its expansion to other areas of health care.[17,18] The OSAD is an 8-item tool incorporating a 5-point Likert-type scale (1 = minimum and 5 = maximum) with associated behavioral markers. The authors (J.P., J.Z., and R.B.) adapted the paper-based version to an electronic format (eOSAD) with immediate uploading of inputs to a database for ease of use and to limit potential data corruption.

Before rating videos, the observers underwent training related to the development and use of the OSAD by one of the authors (J.P.). This training included calibration of scoring through practice rating of debriefing sessions and discussion of differences in scoring with consensus being achieved. On completion of this training, observers rated the videos of the interprofessional faculty teams conducting the PBs and DBs.

Assessment of quality of faculty debriefing

All videos were randomized to account for possible rating bias based on an expected improvement in quality of a team's sequential debriefings. Two authors (V.K. and J.P.) randomized the order of 38 numbered, video-recorded faculty PBs and DBs, which the 2 observers blinded to this randomization (J.Z. and R.B.) then independently rated for overall team quality, using the eOSAD. The number of faculty team combinations participating in PB and DB sessions were identified after observers had rated all teams, using the eOSAD. For the PB session, team ratings were for the first 3 OSAD categories: approach, establishment of the learning environment, and engagement of learners. Teams performing a DB were rated for all 8 categories of OSAD: approach, establishment of learning environment, engagement of learners, reaction, reflection, analysis, diagnosis, and application. The mean scores of the OSAD categories were calculated for each team combination using the OSAD scores of PB or DB sessions in which they participated.

Data analysis

Inter-rater agreement was calculated using Cohen's kappa (k) statistic (SPSS Statistics v 25, SPSS, Inc, Chicago, IL). A k value ≥ 0.61 was considered a substantial agreement.[19] All other statistical analyses were computed using SAS v 9.4 (SAS Institute, Cary, NC). All OSAD scores were expressed as means $\pm$ standard error of the number of team PBs or DBs noted in each table. A one-way between-subjects analysis of variance (ANOVA) with post-hoc Tukey's

Table I
Combinations of faculty team participating in a PB or DB for the HF SBT IPE prelicensure team training exercise

FACULTY TEAM NUMBER	FACULTY TEAM COMPOSITION				TIMES FACILITATED (*N*)	
	Faculty A	Faculty B	Faculty C	Faculty D	Prebrief	Debrief
1	x				2	11
2	x	x	x		9	5
3	x	x			7	3
4		x			4	3
5		x	x		14	11
6	x		x		0	4
7		x		x	2	1
TOTAL	4	5	3	1	38	38

Table II
Descriptive statistics for prebriefing teams

OSAD category	Prebriefing team number	*N*	Mean ± SE
Approach			
	1	2	3.5 ± 0.7
	2	9	4.0 ± 0.8
	3	7	4.3 ± 0.7
	4	4	2.8 ± 0.5
	5	14	3.2 ± 1.0
	6	*	*
	7	2	3.0 ± 0.0
Establishment of learning environment			
	1	2	4.3 ± 0.4
	2	9	4.8 ± 0.4
	3	7	4.5 ± 0.5
	4	4	2.6 ± 1.1
	5	14	3.0 ± 1.0
	6	*	*
	7	2	4.3 ± 0.4
Engagement of learners			
	1	2	4.00 ± 0.0
	2	9	4.7 ± 0.4
	3	7	3.9 ± 0.5
	4	4	3.0 ± 1.1
	5	14	3.4 ± 1.2
	6	*	*
	7	2	4.0 ± 0.7

* Results not applicable because there were no team assignments to this number.

studentized range procedure was conducted to compare the effect of team facilitator grouping on the quality of team performance of each OSAD element during a PB or DP. Trend analyses of teams with four or more observations were performed using the Kendall's Tau coefficient test and linear regression analyses to identify whether teams showed improvement through time. Time was categorized by date and time of PB or DB session, with trend analyses progressing from the earliest conducted PBs or DBs (ie, those done early in the study period) to the most recent (ie, those done later in the study period). Differences were considered statistically significant at $P < .05$.

Results

Faculty teams

A total of 38 video recordings consisted of unique HF, SBT IPE, faculty team-led PB and DBs. A total of 7 separate faculty team combinations led a PB or DB (Table I). Of these combinations, 4 teams conducted 4 or more PBs (teams 2, 3, 4, and 5) or DBs (teams 1, 2, 5, and 6). Individual faculty were members of between 1 to 5 team combinations, with faculty member B being on the most teams ($n = 5$) and faculty member D being on the fewest number of teams ($n = 1$).

Prebrief (PB) analysis

Table II presents mean scores for each of the 3 OSAD categories used to evaluate PBs for each team combination. Scores ranged from a low of 2.6 units for establishment of learning environment for faculty team 4 to a high of 4.8 units for establishment of learning environment for faculty team 2. Only 2 scores across all teams and categories were less than 3. Faculty team 6 did not conduct a PB during the study period.

One-way ANOVA, looking for differences between each faculty team mean score for each category evaluated, revealed statistically significant values for all 3 categories (approach, $P = .03$; establishment of learning environment, $P < .0001$; engagement of learners, $P = .02$). Tukey's studentized range test revealed that statistically significant difference in mean scores between certain faculty teams were present for 2 OSAD categories, the establishment of learning environment and the engagement of learners (Table III). These differences involved faculty teams 2, 3, 4, and 5 and demonstrated mean score differences ranging from 1.3 to 2.2 units. Effect sizes were strong for all differences.

For faculty teams conducting 4 or more PBs (teams 2, 3, 4, and 5), Kendall's Tau coefficient test demonstrated 1 statistically significant change through time (positive change for faculty team 2 for approach [Kendall's Tau coefficient = 0.61, $P = .03$]). Linear regression analysis of each PB category did not reveal a significant change through time ($P = .20$, approach; $P = .92$, establishment of learning environment; $P = .94$, engagement of learners). Substantial agreement existed between observer's judgments on faculty teams' overall PB ratings ($n = 114$, k = 0.872 [95% CI, 0.829–0.915], $P < .001$).

Table III
Comparisons that were statistically significant between means using Tukey's studentized range test for prebriefing teams

OSAD category	Tukey's studentized range test—Prebriefing			
	Prebriefing team comparison	Difference between means	95% Confidence limits*	Effect size†
Establishment of learning environment				
	Team 2 versus Team 5	1.8	(0.75–2.80)	4.05
	Team 2 versus Team 4	2.2	(0.71–3.59)	4.89
	Team 3 versus Team 5	1.5	(0.39–2.61)	3.00
	Team 3 versus Team 4	1.9	(0.37–3.38)	3.76
Engagement of learners				
	Team 2 versus Team 5	1.3	(0.17–2.38)	3.63
	Team 2 versus Team 4	1.7	(0.11–3.22)	4.77

* $P < .05$ for all listed comparisons.
† The effect sizes were calculated by dividing the difference of the compared two groups by the bigger standard deviation of the two groups.

Table IV
Descriptive statistics for debriefing teams regarding the following: Approach, establishment of learning environment, engagement of learners, reaction, reflection, analysis, diagnosis, and application

OSAD category	Debriefing team number	*N*	Mean ± SE
Approach			
	1	11	4.7 ± 0.4
	2	5	4.1 ± 0.4
	3	3	4.7 ± 0.3
	4	3	3.8 ± 0.8
	5	11	3.8 ± 0.8
	6	4	4.8 ± 0.5
	7	1	3.00 ± *
Establishment of learning environment			
	1	11	4.1 ± 0.8
	2	5	3.7 ± 1.1
	3	3	3.8 ± 0.8
	4	3	3.7 ± 0.6
	5	11	4.0 ± 0.6
	6	4	4.6 ± 0.5
	7	1	3.5 ± *
Engagement of learners			
	1	11	4.7 ± 0.4
	2	5	4.3 ± 0.5
	3	3	4.7 ± 0.6
	4	3	4.2 ± 1.0
	5	11	4.1 ± 0.7
	6	4	4.8 ± 0.5
	7	1	4.0 ± *
Reaction			
	1	11	4.4 ± 0.7
	2	5	4.0 ± 0.6
	3	3	4.5 ± 0.5
	4	3	3.7 ± 1.0
	5	11	4.11 ± 0.8
	6	4	4.8 ± 0.5
	7	1	4.0 ± *
Reflection			
	1	11	4.7 ± 0.4
	2	5	4.6 ± 0.4
	3	3	4.8 ± 0.3
	4	3	4.3 ± 0.8
	5	11	4.4 ± 0.6
	6	4	4.9 ± 0.3
	7	1	4.0 ± *
Analysis			
	1	11	4.7 ± 0.4
	2	5	4.8 ± 0.8
	3	3	4.8 ± 0.3
	4	3	4.2 ± 0.3
	5	11	4.4 ± 0.7
	6	4	5.00 ± 0.0
	7	1	4.0 ± *
Diagnosis			
	1	11	4.7 ± 0.4
	2	5	4.60 ± 0.6
	3	3	4.50 ± 0.5
	4	3	4.17 ± 1.0
	5	11	4.45 ± 0.7
	6	4	5.00 ± 0.0
	7	1	4.0 ± *
Application			
	1	11	4.6 ± 0.5
	2	5	4.9 ± 0.2
	3	3	4.7 ± 0.3
	4	3	4.1 ± 1.0
	5	11	4.3 ± 1.0
	6	4	4.9 ± 0.3
	7	1	4.0 ± *

* Results not applicable because of only one observation for that group.

Debrief (DB) analysis

Table IV presents mean scores for each of the 8 OSAD categories used to evaluate DBs for each team combination. Faculty team 7 conducted only 1 DB. Of those faculty teams facilitating more than 1 DB, mean scores ranged from a low of 3.7 units for establishment of the learning environment and reaction for faculty team 4 to a high of 5.0 units for analysis and diagnosis for faculty team 6. All mean scores were above 3.5 units across all 8 OSAD categories, except for faculty team 7. The engagement of learners, reflection, analysis, diagnosis, and application categories all had faculty team mean scores greater than or equal to 4.0 units.

One-way ANOVA, looking for differences between the mean score of each faculty team for each category evaluated, revealed differences for only 1 of the 8 OSAD categories (approach, $P = .004$). Tukey's studentized range test revealed that a difference in mean scores was present between faculty team 1 compared with faculty team 5 for approach (difference of means = 0.86; 95% CI 0.10–1.63; effect size = 2.15).

For faculty teams conducting 4 or more DBs (teams 1, 2, 5, and 6), Kendall's Tau coefficient test demonstrated positive changes through time for faculty team 1 in 3 categories (reaction, Kendall's Tau coefficient = 0.62, $P = .01$; analysis, Kendall's Tau coefficient = 0.57, $P = .03$; diagnosis, Kendall's Tau coefficient = 0.58, $P = .02$). In addition, a positive change through time was present for faculty team 5 in 1 category (analysis, Kendall's Tau coefficient = 0.60, $P = .02$). A negative change was present for 1 category for faculty team 2 (reflection, Kendall's Tau coefficient = −0.89, $P = .04$). Of note, linear regression analysis of each DB category revealed statistically significant changes through time for 7 of the 8 OSAD categories (Table V). Substantial agreement existed between observer's judgments on the overall DB ratings of faculty teams ($n = 304$, k = 0.665 [95% CI, 0.601 to 0.729], $P < .001$) and combined PB and DP ratings ($n = 418$, k = 0.781 [95% CI, 0.744 to 0.818], $P < .001$).

Discussion

Although medical educators have long recognized the importance of effective DB as a critical component for the success of SBT,[20,21] little work to date has focused on determining the quality of PBs and DBs within such SBT activities.[13–15,22] Furthermore, such work has for the most part used the Debriefing Assessment for Simulation in Healthcare assessment tool,[10] concentrating on aspects of learner engagement[22] or comparisons between learners' and facilitators' perceptions of DB quality[13] or resident versus faculty DB.[14] By comparison, this current work assessed the quality of the DB using a different instrument: the OSAD.[9] In addition, our work focused solely on faculty teams conducting PBs and DBs as part of an HF SBT for IPE of prelicensure students, providing unique insight into the current state of PBs and DBs in actual teaching practice.

Among the 7 different faculty teams facilitating the PBs and DBs for the HF, SBT sessions, the OSAD ratings revealed several interesting findings. First, the PBs demonstrated the most variability among the team combinations, with all 3 items (ie, approach, establishment of learning environment, and engagement of learners) being statistically significant. These differences in quality were especially marked for the item of establishment of learning environment between faculty teams 2 and 3 versus teams 4 and 5, in which the former teams outperformed the latter ones by more than 1.5 units. In addition, differences were noted between team 2 versus teams 4 and 5 related to the scores for engagement of learners in which team 2 outscored teams 4 and 5 by more than 1.2 units. Given that effect sizes for these differences all were between 3.5 and 5.0, such variability could have had an impact on learning.

This observation is especially true when taking into account that a critical feature of establishing a learning environment at the beginning of a SBT session is to create the psychologic safety that is essential for learners to be receptive to new concepts and ideas.[8] Of

Table V
Linear regression results for debriefing teams concerning approach, establishment of learning environment, engagement of learners, reaction, reflection, analyses, diagnosis, and application

Question	Variable	Linear regression—Debriefing				
		Estimate	Standard error	DF	T value	*P* value
Approach	Time	0.02	0.009	30	2.10	0.04
Establishment of learning environment	Time	0.006	0.01	30	0.47	0.64
Engagement of learners	Time	0.03	0.008	30	4.21	0.0002
Reaction	Time	0.04	0.01	30	3.88	0.0005
Reflection	Time	0.02	0.008	30	2.71	0.01
Analysis	Time	0.03	0.006	30	4.26	0.0002
Diagnosis	Time	0.03	0.008	30	3.78	0.0007
Application	Time	0.03	0.01	30	3.21	0.003

DF, degrees of freedom.

note, team 2 consisted of three faculty members, suggesting that having multiple cofacilitators might help with establishing such psychologic safety. In addition, having multiple cofacilitators might assist with obtaining early learner involvement that promotes self-assessment, which is key to conducting successful DBs.[23] The fact that team 4 consisted of only one individual, faculty member B, provides additional evidence that having cofacilitators may confer a benefit for conducting PBs. Finally, the fact that two of the three faculty members on team 2 (faculty members A and C) had experience together conducting joint DBs may have helped this team in setting the stage for learning.

Clearly, differences in team performance did exist during PBs. These differences, however, were not necessarily only faculty dependent. For example, faculty member B was a facilitator on teams 2, 3, 4, and 5, and faculty member C participated on teams 2 and 5. In addition, team 6 consisted of faculty members who had experience with joint DBs, but this team did not outperform other teams. Thus, differences were likely attributable to a combination of issues, such as the interplay between facilitators and that interaction with learners. Thus, cofacilitators could potentially complement the style of one another to improve the quality of the PB, or they could have competing agendas, degrading the quality of the PB.[24] The differences could also be attributable to variables that are facilitator independent: the learners participating in the session, the environment where the PBs and DBs occur, or the time of day in which the PBs or DBs are taking place. In the current study, the learners' relative unfamiliarity with PBs and DBs may have had an impact. Also, the locale of the PBs and DBs was within the simulation room itself. Finally, the afternoon sessions, after a morning of teaching, may have contributed to the receptivity of the learners.

Whereas the PBs demonstrated variability in the quality of individual teams' performances, the DBs were remarkably consistent related to the OSAD items. Of eight evaluated, only the item of approach had a statistically significant variability between team ratings. This variability, however, did not translate into the team-specific differences revealed with certain PB items. More noteworthy was the high quality of DBs (ie, scores $\geq$4 units) that all seven teams demonstrated for the majority of the OSAD items: engagement of learners, reaction (except for one team), reflection, analysis, diagnosis, and application. All these items, except for quality of engagement of learners, correspond to one of the three recognized phases of a DB conversational structure: reaction or defusing (ie, reaction), analysis or discovering (ie, reflection, analysis, diagnosis), and summary or deepening (ie, application).[23] The faculty teams, therefore, were quite good at performing the core elements of a DB, which help to guide learners to new insights, knowledge, skills, and attitudes.

Another interesting finding was related to the teams' performances through time. Six of the seven faculty teams conducted four or more PBs or DBs, providing an opportunity to examine whether the quality of particular aspects of their PBs and DBs improved, degraded, or remained constant as they facilitated sessions throughout the duration of the course sessions. Kendall's Tau analysis identified specific teams with significant trends through time for particular OSAD items: positive changes in DB reaction, analysis, and diagnosis for team 1; a positive change for PB approach and a negative change for DB reflection for team 2; and a positive change for DB analysis for team 5. Thus, overall, individual teams tended to improve specific aspects of their performance and quality of PBs and DBs as they facilitated more sessions. The exception was team 2, which demonstrated a decrease in the DB reflection score through time. This team was the only team having three faculty members. Although it performed quite well compared with other teams, the number of cofacilitators during the reflection phase may have led to interactions recognized as detrimental when more than one person conducting the debriefing is present (ie, lack of knowledge of learning objectives, differing personal agendas, inadequate use of expertise, hijacking of debrief, domination of discussion, concentration on only one learner group, and open disagreement[24]). We found it encouraging that team improvements focused again on core elements related to the three conversational components of debriefing: reaction, analysis, and diagnosis.

Linear regression analysis revealed another encouraging finding among teams conducting four or more DBs. Statistically significant positive changes through time were present for these teams for every OSAD item except establishment of learning environment. Thus, as teams facilitated more sessions throughout the duration of the course, their overall quality of DB seemed to improve, giving credence to the well-known expression, "Practice makes perfect." Faculty teams, just like the SBT participants themselves, appear to learn and adjust as they progress from one DB to the next. Such an improvement through time may also have resulted from the fact that the faculty members were rotating through various teams, giving them the opportunity to see and learn techniques from each other. This trend was not present in the PBs, which may explain the differences found between scores and teams. By their nature, the PBs had to be shorter in time and more structured in content, and this may account for the lack of improvement. Efforts directed at faculty development targeting the PB might provide a better return on investment.

Limitations do exist in our study. First, the differences in the quality of the debriefing may be more faculty dependent than it otherwise appears to be. In addition, there were only 4 faculty evaluated in the study. Each faculty member had varying degrees of experience in performing the IPE PBs and DBs for the HF SBT of prelicensure student teams. Determining the influence of this variability in experience was difficult because baseline scores of individual faculty members were absent. Furthermore, faculty teams consisted of 1 to 3 members and represented 7 different combinations of faculty members. Such variability in combinations

and size make differentiating which influences were more influential difficult. Finally, faculty members did not receive formative feedback related to their performance because the videos were evaluated after the sessions were completed. Having faculty rate themselves and each other immediately after each DB would be a useful way to encourage self-reflection and potentially improve skills.

Two future directions for this work are apparent. One avenue relates to faculty development in debriefing. The use of the OSAD helped to identify specific items related to the PBs and DBs that were different among teams. This knowledge can promote initiatives of faculty development to train teams with particular weaknesses to hone these skills. Such focused, team/facilitator-specific training would address a key need in SBT faculty development.[25] In addition, peer coaches or facilitators themselves could use the OSAD for formative assessment to help identify and improve components of PBs and DBs. Peer coaching is now an important adjunct to improving the quality of PBs and DBs.[26] Determining the most effective techniques for training and coaching of the facilitators would provide key insights and increase the efficiency of training. A second avenue focuses on better defining the relationship between the quality of PBs and DBs with learning. Determining which, if any, specific OSAD items correspond with the greatest gain in learning among SBT participants would be invaluable in targeting those aspects of SBT sessions on which to concentrate to optimize learning.

In conclusion, effective debriefing is essential for learning in HF SBT. To date, the quality of such debriefing in actual teaching practice is not well known. This work demonstrated that among seven faculty teams facilitating IPE involving the HF SBT for prelicensure student teams, the overall quality of PBs and DBs is rather good. Differences between various team performances, when they existed, occurred during the PB period of these sessions. We were encouraged to find that faculty teams tended to improve the quality of their DBs through time. Future directions include determining the relationship between the quality of components of PBs and DBs and participant learning.

Acknowledgments

Louisiana State University Health New Orleans Health Sciences Center is the institution from which this work originated. The videos of the faculty prebriefings and debriefings were initially collected as part of an in-house 2011 to 2012 Educational Enhancement Grant from the LSU Health New Orleans Academy for the Advancement of Educational Scholarship for a project on team training of interprofessional students.

Disclosure

Dr. Paige receives royalties from the following: Oxford University Press as co-editor for *Simulation in Radiology*, wound research grant support from Acell, Inc. as principal investigator (PI), instrument development research grant support from the Southern Group on Education Affairs (SGEA) as PI, software development research grant support from the LSU Board of Regents (BOR) as co-PI, and a recently received training grant support from the Health Resources and Services Agency (HRSA) as lead faculty.

Drs. Garbee, Bonanno, Yu, and Kiselov also received research support for the SGEA grant and the HRSA grant, of which Dr. Bonanno was the Director.

Drs. Rusnak and Kiselov also received support for the LSU BOR grant.

Drs. Zamjahn and Carvalho, as well as Shuang Yang and Catherine Callan, report no conflicts of interest, financial or otherwise.

Funding

The original team training of interprofessional students project was funded by an in-house 2011 to 2012 Educational Enhancement Grant from LSU Health New Orleans Academy for Advancement of Educational Scholarship.

References

1. Gillespie BM, Chaboyer W, Wallis M, Fenwick C. Why isn't 'time out' being implemented? An exploratory study. *Qual Saf Health Care*. 2010;19:103–106.
2. Bleakley A. You are who I say you are: The rhetorical construction of identity in the operating theatre. *J Workplace Learning*. 2006;18:414–425.
3. Nagpal K, Vats A, Ahmed K, Vincent C, Moorthy K. An evaluation of information through the continuum of surgical care. *Ann Surg*. 2010;252:402–407.
4. Greiner AC, Knebel E, eds. *Health professions education: A bridge to quality*. Washington, DC: National Academies Press; 2003.
5. Chatman IJ. *Medical team training: Strategies for improving patient care and communication*. Oakbrook Terrace, IL: Joint Commission Resources; 2008.
6. Paige JT, Garbee DD, Brown KM, Rojas JD. Using simulation in interprofessional education. *Surg Clin N Am*. 2015;95:751–766.
7. Paige JT, Garbee DD, Yu Q, Rusnak V. TTIPS (Team Training of Inter-Professional Students) for Improving Teamwork. *BMJ Simul Technol Enhanc Learn*. 2017;3:127–134.
8. Paige JT, Arora S, Fernandez G, Seymour N. Debriefing 101: Training faculty to promote learning in simulation-based training. *Am J Surg*. 2015;209:126–131.
9. Arora S, Ahmed M, Paige J, et al. Objective structured assessment of debriefing: Bringing science to the art of debriefing in surgery. *Ann Surg*. 2012;256,982–988.
10. Brett-Fleegler M, Rudolph J, Eppich W, et al. Debriefing assessment for simulation in healthcare: Development and psychometric properties. *Simul Healthc*. 2012;7:288–294.
11. Kolbe M, Weiss M, Grote G, et al. TeamGAINS: A tool for structured debriefings for simulation-based team trainings. *BMJ Qual Saf*. 2013;22:541–553.
12. Saylor JL, Wainwright SF, Herge EA, Pohlig RT. Peer-assessment debriefing instrument (PADI): Assessing faculty effectiveness in simulation education. *J Allied Health*. 2016;45:e27–e30.
13. Durand C, Secheresse T, Leconte M. The use of the debriefing assessment for simulation in healthcare (DASH) in a simulation-based team learning program for newborn resuscitation in the delivery room. *Arch Pediatr*. 2017;24:1197–1204.
14. Adams T, Newton C, Patel H, Sulistio M, Tomlinson A, Lee W. Resident versus faculty member simulation debriefing. *Clin Teach*. 2018;15:462–466.
15. Bullard MJ, Leuck JA, Howley LD. Unifying interdisciplinary education: designing and implementing an intern simulation educational curriculum to increase confidence in critical care from PGY1 to PGY2. *BMC Res Notes*. 2017;10:563.
16. Ahmed M, Sevdalis N, Paige J, Paragi-Gururaja R, Nestel D, Arora S. Identifying best practice guidelines for debriefing in surgery: A tri-continental study. *Am J Surg*. 2012;203:523–529.
17. Runnacles J, Thomas L, Sevdalis N, Kneebone R, Arora S. Development of a tool to improve performance debriefing and learning: The paediatric objective structured assessment of debriefing (OSAD) tool. *Postgrad Med J*. 2014;90:613–621.
18. Runnacles J, Thomas L, Korndorffer J, et al. Validation evidence of the paediatric Objective Structured Assessment of Debriefing (OSAD) tool. *BMJ Simul Technol Enhanc Learn*. 2016;2:61–67.
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
20. Issenberg SB, McGaghie WC, Petrusa ER, et al. Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. *Med Teach*. 2005;27:10–28.
21. McGaghie WC, Issenberg SB, Petrusa ER, et al. A critical review of simulation-based medical education research: 2003–2009. *Med Educ*. 2010;44:50–63.
22. Roh YS, Jang KI. Survey of factors influencing learner engagement with simulation debriefing among nursing students. *Nurs Health Sci*. 2017;19:485–491.
23. Sawyer T, Eppich W, Brett-Flegger M, Grant V, Cheng A. More than one way to debrief: A critical review of healthcare simulation debriefing methods. *Simul Healthc*. 2016;11:209–217.
24. Cheng A, Palaganas J, Eppich W, Rudolph J, Robinson T, Grant V. Co-debriefing for simulation-based education: a primer for facilitators. *Simul Healthc*. 2015;10:69–75.
25. Cheng A, Grant V, Dieckmann P, Arora S, Robinson T, Eppich W. Faculty development in simulation programs: Five issues for the future of debriefing training. *Simul Healthc*. 2015;10:217–222.
26. Cheng A, Grant V, Huffman J, et al. Coaching the debriefing: Peer coaching to improve debriefing quality in simulation programs. *Simul Healthc*. 2017;12:319–325.