Brief Report

Frame-of-reference training for simulation-based intraoperative communication assessment

CrossMark

Aimee K. Gardner, Ph.D.[a,*], Michael A. Russo, M.D.[b],
Ibrahim I. Jabbour, M.D.[c], Matthew Kosemund, E.M.T.[a],
Daniel J. Scott, M.D.[a]

[a]Department of Surgery, University of Texas Southwestern Medical Center, 5323 Harry Hines
Boulevard, Dallas, TX 75390, USA; [b]Smart Dimensions Weight Loss, Fountain Valley, CA; [c]Department
of Surgery, Phoenix, AZ

Abstract
BACKGROUND: The purpose of this study was to examine the impact of frame-of-reference (FOR) training on assessments of intraoperative communication skills and identify areas of need to inform curricular efforts.
METHODS: Simulation instructors (M.D., Ph.D., Research Fellow, Simulation Technician) underwent a 2-hour FOR training session with the operating room communication instrument. They then independently rated communication skills of 19 PGY1s who participated in a team-based simulation. Residents completed self-assessments via video review of the scenario. Intraclass correlation coefficients were used to examine inter-rater reliability. Relationships between trained raters and resident scores were assessed with Pearson correlation coefficients and paired sample t tests.
RESULTS: Inter-reliability after FOR training was .91. The correlation between trained rater scores and resident evaluations was nonsignificant. Residents significantly underestimated their intraoperative communication skills ($P < .05$). Use of names, closed loop communication, and sharing information with team members demonstrated consistently low ratings among all residents.
CONCLUSIONS: These findings reveal that a number of individuals can be trained to reliably rate resident intraoperative communication performance and that residents tend to under-rate their communication skills. © 2016 Elsevier Inc. All rights reserved.

Communication breakdowns are a common threat to patient safety. Not only have communication errors been highlighted as a leading cause of adverse events across the entire health care system, but they have also emerged particularly in surgical care pathways and in the operating room.[1] In fact, in a study examining reasons for malpractice claims of surgical patients, Greenberg et al[2] noted that 30% of communication breakdowns occur in the intraoperative environment, with the attending surgeon the most frequently involved participant in breakdowns.

The aforementioned outcomes research coincides with national efforts to better understand needs of surgical training. For example, in a recent study led by the American College of Surgeons Accredited Education

Institutes Curriculum Committee, national stakeholders were interviewed to elicit their opinions and recommendations regarding the future of training in surgery. From this work, one major theme emerged: surgical educators need to focus training efforts on enhancing communication skills among our trainees.[3] Unfortunately, although, there are incomplete data to best understand specifically where communication gaps lay, how they can be assessed, and the efficacy of curricula intended to improve communication skills. This is despite national mandates to evaluate and document these competencies.[4] Thus, there is ample opportunity to guide initiatives intended to assess and improve communication skills in a standardized fashion among surgical trainees.

To further understand and examine these issues, we designed a study with 3 overarching goals in mind. First, we wanted to examine the impact of frame-of-reference (FOR) training[5] on evaluations from raters with a variety of backgrounds, with hopes that providing raters with a common FOR for using a communication tool would eliminate individual differences and result in more reliable ratings. Additionally, we wanted to examine if residents' self-assessments of their communication skills were systematically different than those of trained raters. And last, we wanted to identify any gaps that may exist in these communication ratings to better inform what additional curricula may need to be implemented.

Methods

Simulation training

PGY1 residents (n = 19; 12 categorical and 7 preliminary; 58% male) participated in an institutional review board–approved simulation scenario adapted from the American College of Surgeons/Association of Program Directors in Surgery team-based skills curriculum (laparoscopic trouble-shooting module),[6] in which residents had to troubleshoot a technical malfunction, respond to a de-saturating patient after insufflation for a laparoscopic cholecystectomy, and lead a particularly unhelpful confederate team (operating room nurse, anesthesiologist) through the Advanced Cardiac Life Support algorithm.

All performances were video recorded and assessed with the Operating Room Communication Assessment (ORCA) tool, which was initially validated for this scenario.[7] This instrument was developed as part of a previous body of work regarding training and assessment of teamwork skills in an operating room environment. The tool was created by a multidisciplinary group of experts, including surgeons, anesthesiologists, and nurses and includes 16 domains that aim to assess communication, leadership, and team interactions. This group felt that although instruments to evaluate teamwork exist, none specifically and thoroughly addressed operating room communication. Our preliminary work involving 24 participants demonstrated significant

differences between novices and experts using this instrument for the laparoscopic troubleshooting simulation.[7] After a brief orientation to the ORCA tool, trainees were provided with their own videos and rated themselves after the simulation was completed.

FOR training

Before trainee participation, 4 members of the simulation staff (Surgeon, Ph.D., PGY2 Research Fellow, Simulation Technician) participated in a 2-hour FOR training session as this method has been shown to improve performance appraisal rating accuracy.[8–10] Training began with all participants watching five 6- to 10-min intraoperative scenarios and independently completing a 16-item ORCA evaluation tool (Appendix 1). Raters were instructed to take notes while watching the videos to indicate why each rating was chosen. After this initial rating session, raters then went through each item on the rating scale and discussed definitions and scale anchors. Raters discussed ratee behaviors that exemplified different performance levels for each scale within the context of the scenario previously observed, until they reached an overall consensus on expectations to achieve each level of all 16 items. The goal of the training was to create a common performance theory (ie, FOR) among raters such that they would agree on the appropriate performance dimension and effectiveness level of different behaviors. As a key component of FOR training is the discussion of dimensions and the behaviors indicative of each dimension, a group setting was necessary. After training, trained simulation team raters watched the remaining 14 videos and rated them independently.

All analyses were conducted with SPSS, version 22.0 (IBM, Chicago, IL) and a significance level of P less than .05 was chosen. Inter-rater reliability of communication ratings among trained raters was assessed with intraclass correlation coefficients (ICCs). Correlation coefficients and independent-sample t tests were used to compare differences between trainee and rater averages.

Results

Results from the FOR training are displayed in Fig. 1 as composite values of the ORCA tool. As shown, the reliability of ratings across faculty raters before FOR was .32 (95% confidence interval, −.06, .42, ns). After training, however, ICC values improved to .91 (95% confidence interval, .72, .97, $P < .001$), making the use of mean ratings of performance acceptable.

Table 1 illustrates mean values for both resident self-ratings and trained simulation team ratings. The lowest ratings for both groups were in the dimensions of use of names (PGY mean = 2.31 ± 1.38; faculty mean = 1.18 ±.40), anticipation (PGY mean = 2.46 ± .88; faculty mean = 2.93 ± .62), and verifying acknowledgment with the team