

Value of Debriefing during Simulated Crisis Management

Oral versus Video-assisted Oral Feedback

Georges L. Savoldelli, M.D., M.Ed.,* Viren N. Naik, M.D., M.Ed., F.R.C.P.C.,† Jason Park, M.D., M.Ed.,‡
Hwan S. Joo, M.D., F.R.C.P.C.,§ Roger Chow,|| Stanley J. Hamstra, Ph.D.#

Background: The debriefing process during simulation-based education has been poorly studied despite its educational importance. Videotape feedback is an adjunct that may enhance the impact of the debriefing and in turn maximize learning. The purpose of this study was to investigate the value of the debriefing process during simulation and to compare the educational efficacy of two types of feedback, oral feedback and videotape-assisted oral feedback, against control (no debriefing).

Methods: Forty-two anesthesia residents were enrolled in the study. After completing a pretest scenario, participants were randomly assigned to receive no debriefing, oral feedback, or videotape-assisted oral feedback. The debriefing focused on nontechnical skills performance guided by crisis resource management principles. Participants were then required to manage a posttest scenario. The videotapes of all performances were later reviewed by two blinded independent assessors who rated participants' nontechnical skills using a validated scoring system.

Results: Participants' nontechnical skills did not improve in the control group, whereas the provision of oral feedback, either assisted or not assisted with videotape review, resulted in significant improvement ($P < 0.005$). There was no difference in improvement between oral and video-assisted oral feedback groups.

Conclusions: Exposure to a simulated crisis without constructive debriefing by instructors offers little benefit to trainees. The addition of video review did not offer any advantage over oral feedback alone. Valuable simulation training can therefore be achieved even when video technology is not available.

FULL-SCALE high-fidelity mannequin simulators are increasingly recognized as useful educational adjuncts. Within anesthesia, these tools are used for various training purposes, including simulating rare events, teaching technical skills, or advanced life support algorithms.¹ The simulation room is also an ideal setting for teaching the principles of crisis resource management.² In this environment, the importance of nontechnical skills such as task management, team working, situation awareness,

or decision making, can be safely practiced. A recent study confirmed the instructional value of simulation for acquiring these cognitive and interpersonal skills.³

Simulation-based learning is typically experiential.⁴ The experience is affected by the quality of the scenario, the instructor's expertise, and the feedback process.⁵ The debriefing process following a scenario allows trainees to reflect on their performance as well as receive instructor's feedback. Reviewing one's performance by video may be a useful adjunct to the debriefing process. Among supposed benefits, it is thought to provide an objective record, facilitate instructor's constructive comments, and promote trainee's self-assessment. Videotape feedback has proven useful in other fields outside of medicine and in some areas within medicine, including anesthesia.^{6,7} Although many educators believe in its value, videotape feedback is not systematically used in simulation. In addition, despite the perceived importance of the debriefing process during simulation, only one study has empirically assessed its impact, and the study was inconclusive.⁸

The purpose of this study was to assess the value of the debriefing process during simulation-based education. We compared the changes in nontechnical performance when anesthesia residents received no feedback, instructor oral feedback only, or videotape-aided instructor oral feedback.

Materials and Methods

Participation and Orientation Phase

After Institutional Research Board (St. Michael's Hospital, University of Toronto, Toronto, Ontario, Canada) approval, anesthesia residents in postgraduate years 1, 2, and 4 from the University of Toronto were invited to participate in the study. Informed consent and confidentiality agreements were obtained to ensure that details pertaining to the clinical scenarios would not be disseminated before the end of the study.

Before the simulation sessions, a group orientation session was held for all participants. During this initial 1-h didactic period, the principles of crisis evolution, patient simulation, and anesthesia crisis resource management (ACRM) were discussed.^{9,10} Participants were then familiarized with the Laerdal SimMan[®] simulator mannequin and monitors (Laerdal Medical Canada Ltd., Toronto, Canada), the Datex[®] anesthesia machine (Datex Corporation, St. Laurent, Quebec, Canada), and the mock operating room environment.

* Fellow, † Assistant Professor, St. Michael's Anesthesia Research into Teaching (SMART) Simulation Group, Department of Anesthesia, St. Michael's Hospital, and Wilson Centre for Research in Education, University Health Network, University of Toronto. ‡ Fellow, Wilson Centre for Research in Education, University Health Network, University of Toronto. § Assistant Professor, || Patient Simulation Centre Coordinator, SMART Simulation Group, Department of Anesthesia, St. Michael's Hospital, University of Toronto. # Associate Professor, Wilson Centre for Research in Education, University Health Network, and Department of Surgery, University of Toronto.

Received from the Patient Simulation Centre, St. Michael's Hospital, University of Toronto, Toronto, Ontario, Canada. Submitted for publication November 7, 2005. Accepted for publication April 25, 2006. Supported by a Dean's Excellence Fund, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada. Dr. Savoldelli was also supported by the Geneva University Hospitals, the Faculty of Medicine, and by the Eugenio Litta Foundation, Geneva, Switzerland.

Address correspondence to Dr. Naik: Department of Anesthesia, St. Michael's Hospital, University of Toronto, 30 Bond Street, Toronto, Ontario, Canada M5B 1W8. naikv@smh.toronto.on.ca. Individual article reprints may be purchased through the Journal Web site, www.anesthesiology.org.

First- and second-year residents had previous experience with simulation-training during medical school but were complete novices to ACRM principles. Almost all of the fourth-year residents had a single previous remote simulator-based ACRM session, 2 years before the current study.

Study Design and Intervention

This study used a prospective, randomized, controlled, three-arm, repeated-measures study design. On the day of the study, participants attended their sessions individually. Each session consisted of two different scenarios in which the participants played the role of the primary anesthesiologist. Each scenario lasted approximately 8 min. The entire simulation was videotaped, and a graphical display of the patient's vital signs throughout the session was recorded and overlaid on the video footage. Simulation center staff and one investigator functioned as perioperative personnel in each scenario in the scripted roles of surgeon and nurse. A second investigator played the scripted role of a colleague who was available as a second anesthesiologist to help and to perform tasks when directed but did not offer crisis management advice or differential diagnoses.

The two scenarios simulated an intraoperative cardiac arrest. One featured pulseless electrical activity secondary to massive fat embolism, and the other consisted of a severe ventricular arrhythmia due to hyperkalemia. The cause of the arrest, the context, the sequence of event, and the treatment differed from one scenario to another. We chose two cardiac arrest scenarios to minimize the influence of the case or content specificity on the trainees' performance. To control for any sequencing effect, the order of presentation of the two scenarios was randomized for each participant and was equally distributed among the study groups.

After managing the first scenario (pretest), participants were randomized to one of three groups with stratification according to their level of training. Participants in group 1 (control) were asked to manage the second scenario (posttest), without receiving any feedback regarding their first performance. Participants in group 2 (oral feedback) received oral feedback on their performance. During this debriefing, the participants were encouraged to reflect on their performance and on how it could be improved. The process was facilitated by instructors who provided constructive comments. The critique of each performance focused predominantly on nontechnical skills, *i.e.*, cognitive and behavioral skills, and was guided by ACRM training principles.^{9,10} Technical skills were briefly discussed but were not the focus of the debriefing. In group 3 (video-assisted oral feedback), the debriefing was facilitated by reviewing the

videotape of the participant's performance. Selected video segments were chosen to illustrate the instructors' constructive criticism. After reviewing the segment, the instructor would pause the video to encourage subjects to comment and reflect on cognitive and behavioral aspects of their performance, after which the instructor commented on the selected segment. Video segments of little or no educational value were fast forwarded. Similar to group 2, the instructors provided comments pointing out positive aspects of the performance and offered advice on how it could be improved. To reflect our standard practice, the debriefing sessions in groups 2 and 3 were not strictly time-limited. However, the debriefing usually focused on four to six "major critiques" of the participant's nontechnical skills. This was done because in our experience and from informal participants' feedback in previous studies, long individual debriefings (as opposed to group) tended to overload the trainee with information.³ The debriefing was ended when the instructor's comments and trainee's questions were exhausted. All the debriefings were conducted together by two instructors (G.L.S. and V.N.N.) with very few exceptions when one of them was absent. The instructors knew some of the fourth-year residents from previous simulation sessions but not from having worked with them in the operating room.

After the debriefing, participants in groups 2 and 3 managed the second scenario (posttest). Because this study was built into our current crisis management course, we thought it would be unfair to withhold valuable feedback to some of our trainees. Therefore, after the posttest, participants in groups 1 and 2 received both oral and videotape feedback on their pretest. In addition, all participants received a full debriefing, with videotape on their posttest.

Measurement Instruments and Outcomes Measures

Two evaluators with expertise in simulation and ACRM principles were recruited and trained by the principal investigators to evaluate participants using the Anesthesia Non-Technical Skills (ANTS) scoring system.¹¹ Training of the evaluators consisted of providing them with the background ANTS literature¹¹ and with the User Manual.** In addition, they underwent 4 h of group training using the ANTS to score videotaped performances of simulated crisis. Training videotapes did not involve any study participants. After independently assessing each videotape, scores were compared and discussed. Interrater reliability was not formally assessed before the study but had clearly improved during the training session.

The ANTS system is a behavioral marker system that assesses anesthesiologists' nontechnical skills. It has proven reliability and validity.¹¹ This scoring system is hierarchical and consists of four main skill categories of task management, team working, situation awareness,

** Available at: http://www.abdn.ac.uk/iprc/ants_papers.shtml. Accessed February 9, 2006.

Table 1. Anaesthesia Non-Technical Skills (ANTS) System Categories, Elements, and Rating Scale

| Category | Element |
|---------------------|---|
| Task management | Planning and preparing Prioritizing Providing and maintaining standards |
| Team working | Identifying and utilizing resources Coordinating activities with team members Exchanging information Using authority and assertiveness Assessing capabilities |
| Situation awareness | Supporting others Gathering information Recognizing and understanding |
| Decision making | Anticipating Identifying options Balancing risks and selecting options Reevaluating |

| Rating Label | Description |
|----------------|---|
| 4 = Good | Performance was of a consistently high standard, enhancing patient safety. It could be used as a positive example for others. |
| 3 = Acceptable | Performance was of a satisfactory standard but could be improved. |
| 2 = Marginal | Performance indicated cause for concern. Considerable improvement is needed. |
| 1 = Poor | Performance endangered or potentially endangered patient safety. Serious remediation is required. |
| Not observed | Skill could not be observed in this scenario. |

and decision making. Each category is further subdivided into a number of skill elements (table 1), and each skill element also has a number of different example behaviors for good and poor performance. The ANTS system uses a four-point scale to rate the performance of each skill observed (table 1).

At the conclusion of the intervention phase, the evaluators independently reviewed and rated all videotapes in random order. They were blinded to the participants' randomization, level of training, and scenario's chronological order and had never worked previously with the residents in the clinical setting. For each performance, the behaviors observed were scored at the categorical level and guided by descriptors of the component skill elements. To reflect the overall performance, a total ANTS score was obtained by adding up the four category scores (minimum score 4 and maximum score 16).

Statistical Analysis

Statistical analysis was performed using SPSS 13.0 (SPSS Incorporated, Chicago, IL). Demographic data were analyzed using chi-square test, analysis of variance (ANOVA), and unpaired *t* test as appropriate.

The interrater reliability was assessed using the intraclass correlation coefficient and was measured both for the total ANTS score and at the category level.

Our primary outcome measure was the pretest to posttest change in the total ANTS score (posttest minus pretest score). The mean change in score was treated as the dependent variable and analyzed using a one-way between-subjects ANOVA with the study group as the

independent variable. Significant results were then analyzed using a Student-Newman-Keuls test for *post hoc* comparisons. A similar ANOVA was performed to assess mean change in ANTS score at the category level (posttest minus pretest score).

The association between the level of training and the pretest scores was assessed using a one-way between-subjects ANOVA with the level of training as the independent variable. To determine whether the level of training influenced the amount of learning or interacted with the type of feedback received, a two-way between-subjects ANOVA of the change in total ANTS score was performed, with group and level of training as independent factors. A two-tailed *P* value of less than 0.05 was considered significant for all analyses.

Sample size was calculated *a priori*. In the field of psychology and education, an effect size of greater than 1 SD is considered large and acceptable for a given teaching intervention.¹² We agreed that an effect size of 1.1 would be required to demonstrate a practically significant difference. Therefore, assuming an effect size of 1.1 and a power of 0.8 for ANOVA with three groups, we calculated a total sample size of 39 subjects ($\alpha = 0.05$ two-tailed). Allowing for attrition, we recruited 14 subjects per group.

Results

Demographics and Pretest Results

A total of 42 residents in postgraduate training years 1 ($n = 15$), 2 ($n = 15$), and 4 ($n = 12$) participated in and

Table 2. Participants' Demographics and Pretest Performance

| | Control (n = 14) | Oral Feedback (n = 14) | Video-assisted Oral Feedback (n = 14) | Significance |
|--|------------------|------------------------|---------------------------------------|-----------------------------|
| No. male/No. female | 8/6 | 10/4 | 7/7 | $\chi^2 [2] = 1.38, P = NS$ |
| Mean age, yr* | 29.8 ± 0.81 | 28.9 ± 0.88 | 28.3 ± 0.83 | $F_{2,39} = 0.86, P = NS$ |
| No. per postgraduate training year 1/2/4 | 5/5/4 | 6/5/3 | 4/5/5 | $\chi^2 [2] = 0.92, P = NS$ |
| No. of previous simulator sessions, (including non-ACRM sessions)* | 1.3 ± 0.38 | 1.2 ± 0.39 | 1.0 ± 0.28 | $F_{2,39} = 0.20, P = NS$ |
| ACLS course completed, yes/no | 13/1 | 13/1 | 14/0 | $\chi^2 [2] = 1.1, P = NS$ |
| Debriefing time, min* | NA | 19.7 ± 0.48 | 19.9 ± 0.45 | $t [26] = -0.14, P = NS$ |
| Pretest ANTS score (min 4, max 16)* | 9.7 ± 0.70 | 8.5 ± 0.49 | 9.4 ± 0.54 | $F_{2,39} = 1.12, P = NS$ |

* Values are mean ± SE.

ACLS = Advanced Cardiac Life Support; ACRM = Anesthesia Crisis Resource Management; ANTS = Anaesthesia Non-Technical Skills; NA = not applicable; NS = not significant.

completed the study. The demographics and the mean pretest ANTS score for each group are summarized in table 2. There was no significant difference among the three study groups.

Interrater Reliability

The overall interrater reliability for the total ANTS score was acceptable: intraclass correlation coefficient (single rater) = 0.64 ($P < 0.001$). At the category level, across the four categories, interrater reliability was acceptable: intraclass correlation coefficient (single rater) = 0.58 ($P < 0.001$).

Primary Outcome: Pretest to Posttest Change in Total ANTS Score

Change scores were calculated as total ANTS score at posttest minus total ANTS score at pretest. Analysis of variance revealed that this change score was significantly different between the three groups ($F_{2,39} = 6.10, P < 0.005$). *Post hoc* comparisons revealed that compared with controls (−1%), improvement was greater among participants who received feedback, either oral (+15%) or video-assisted oral feedback (+11%) (fig. 1). The amount of improvement in total ANTS score was not

significantly different between the oral and video-assisted oral feedback groups. Within-group comparisons of the mean pretest and posttest total ANTS score are shown in figure 2.

Secondary Outcome Measures: Pretest to Posttest Change in ANTS Score at the Category Level

For each of the four ANTS categories, there was a significant difference in the mean changes in score (posttest minus pretest) between the three groups: task management ($F_{2,39} = 4.42, P < 0.05$), team working ($F_{2,39} = 6.65, P < 0.005$), situation awareness ($F_{2,39} = 3.56, P < 0.05$), and decision making ($F_{2,39} = 3.94, P < 0.05$).

Post hoc comparisons are detailed in figure 3. In three categories, the improvement in the oral feedback group was significantly greater compared with the control group; it did not reach significance in the situation awareness category. The difference between the control and video-assisted oral feedback groups reached significance in two categories: task management and team working. The oral and video-assisted oral feedback groups did not differ statistically for any of the four ANTS categories.

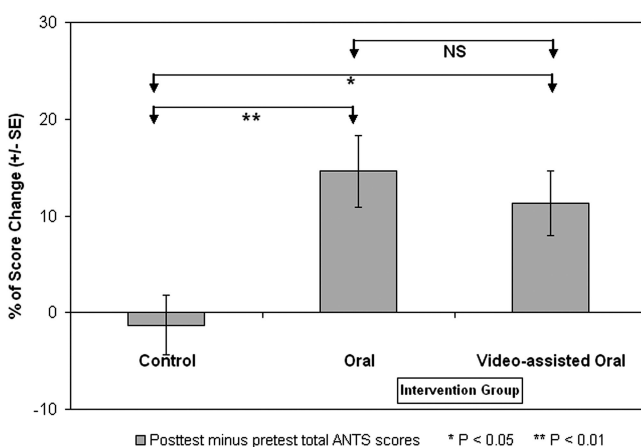


Fig. 1. Average pretest to posttest changes in nontechnical skills performances. ANTS = Anesthesia Non-Technical Skills; NS = not significant.

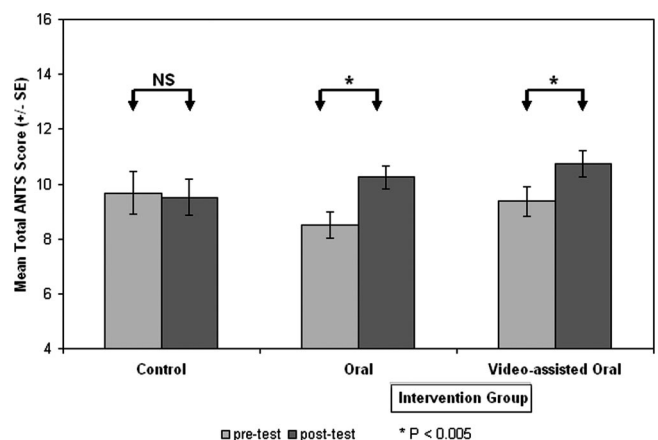


Fig. 2. Comparison of pretest and posttest total Anesthesia Non-Technical Skills (ANTS) scores. NS = not significant.

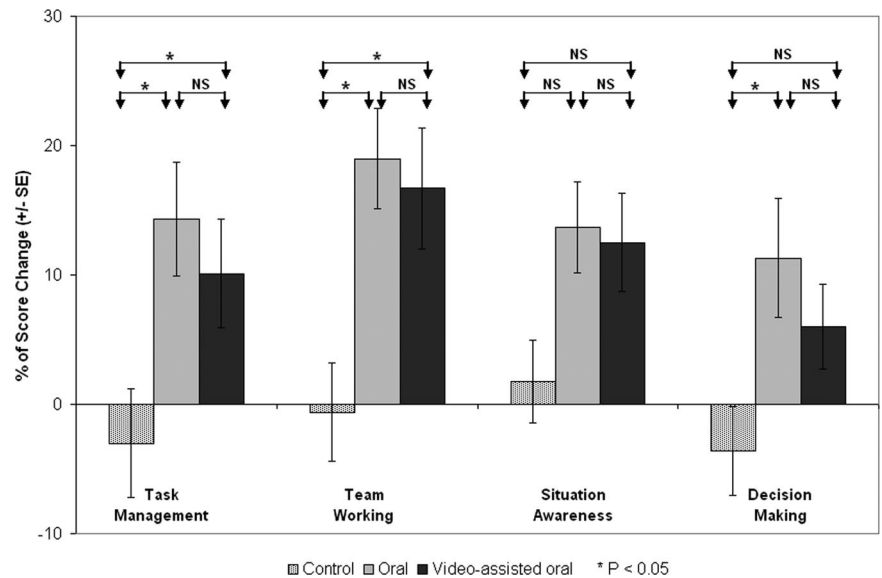


Fig. 3. Average pretest to posttest changes in nontechnical skills categories. NS = not significant.

The Effect of Level of Training

Mean pretest total ANTS score significantly differed among first-, second-, and fourth-year residents ($F_{2,39} = 5.45$, $P < 0.01$). *Post hoc* analysis revealed that fourth-year residents' mean scores were significantly higher compared with those of first-year residents (10.6 ± 0.5 vs. 8.0 ± 0.4 , $P < 0.01$) but did not significantly differ from those of second-year residents (10.6 ± 0.5 vs. 9.25 ± 0.7 , $P =$ not significant). First- and second-year residents did not differ.

The results of the two-factor ANOVA on total ANTS change scores showed a significant main effect for the group factor ($F_{2,33} = 5.0$, $P < 0.05$), no significant main effect for the level of training factor ($F_{2,33} = 0.1$, $P =$ not significant), and no significant interaction between the two ($F_{4,33} = 0.8$, $P =$ not significant). This suggests that although experienced residents scored higher at baseline testing, subsequent learning depended on the provision of feedback but was not influenced by the level of training.

Discussion

The current study investigated the effect of two different debriefing modalities when teaching nontechnical skills using simulation. Participants' performances did not improve in the absence of debriefing, whereas the provision of constructive feedback on the initial performance by skilled instructors resulted in significant improvement. The addition of video review did not offer any advantage over oral feedback alone.

Our results confirm that from an educational standpoint, exposure to a simulated crisis without debriefing seems to offer little benefit to learners. Trainees' self-reflection along with instructors' feedback during a debriefing session seems to be required during simulation-based education.

Videotape feedback is currently regarded as a valuable component of simulation-based education.⁵ However, to our knowledge, only one study has previously investigated its value.⁸ Byrne *et al.*⁸ compared the effect of two types of feedback: a brief explanation of the simulated case with or without a videotape review of the performance. Their results showed that the performances did not improve between the pretest and the posttest regardless of the feedback received. The authors attributed their disconcerting findings to the large variability of their outcome measures (time and chart completion errors) and the use of different content domains (clinical subject areas) for the pretest and the posttest scenarios, rather than an ineffectiveness of the feedback. It should also be noted that in their design, the instructor's role was limited to the provision of a short explanation of the crisis and did not include constructive feedback on the performance. Therefore, their limited feedback may not have been effective.

The current study differed from the study of Byrne *et al.* in several ways. First, the focus of our training was on "nontechnical skills." Therefore, we used the ANTS system as an outcome measure. This measurement tool has been shown to be reliable, valid,¹¹ and able to capture performance improvement during subsequent simulation sessions.³ Second, our debriefing intervention was specific, was detailed, and included trainees' self-assessment and reflection along with the provision of constructive feedback (with or without videotape feedback). This type of debriefing is in accordance with the educational theory of experiential learning and reflects the current practice at many centers.⁴ Finally, we included a control group and used pretest and posttest scenarios that were different but shared similar content domains.

In the current study, participants who received a debriefing session did show an improvement when com-

pared with controls. Surprisingly, the improvement tended to be lower in the video-assisted oral feedback group than in the oral feedback group. This trend is illustrated by the pattern of changes in the total ANTS scores and in the changes in the ANTS scores at the category level. It is possible that the review of the videotape may have interfered with the instructors' feedback by changing the content of the feedback. However, this seems unlikely because our instructors discussed similar aspects of the performances during the debriefings regardless of the use of videotape review. A more likely explanation is that the videotape review in addition to verbal instructor feedback may have caused "information overload." Trainees may have been distracted by the video, thus paying less attention to the instructors' constructive comments and criticisms.

As mentioned previously, the outcome measure is critical when assessing educational outcomes. The ANTS scoring system has been extensively described and discussed in previous literature.^{3,11} It has some limitations, such as modest reliability and an imperfect distinction between nontechnical skills and pure medical knowledge in some of the elements of the scale. However, it offers the major advantages of being a useful instrument for the review and the scoring of videotaped performance of simulated anesthesia crisis management. The moderate interrater reliability that we observed illustrates the difficulty of consistently assessing nontechnical skills. Nevertheless, this level of reliability compares favorably with previous studies of behavioral performances.^{3,11,13} We also found that senior residents obtained higher ANTS scores than their junior counterparts during the pretest, even though the variability in performances was observed within a given level of training. The fact that fourth-year residents had a previous remote ACRM training session may partly explain the differences in scores at baseline. However, it is probable that the skills they have learned 2 yr ago have decayed with time. Therefore, a more likely explanation is that greater clinical experience and increased exposure to attending anesthesiologists in the clinical setting resulted in better anesthesia nontechnical skills when compared with very junior residents. This important finding provides new information regarding the ANTS system and further supports its construct validity.

Our results may also have important implications for crises that occur during real case management. This study suggests that oral debriefs of such events may be effective at changing behaviors without the use of video. Unfortunately, systematic debriefing of real crises does not occur because of many organizational barriers. We therefore offer the recommendation that initiatives to facilitate crisis debriefing should be considered at every institution.

The moderate amount of improvement observed in the intervention groups deserves some comments. It is dif-

ficult to devise an 8-min scenario that covers and highlights all of the nontechnical skills enclosed in the ANTS. The two cardiac-arrest scenarios used in the study emphasized issues relating to the four ANTS categories. Many nontechnical skills were common between the two scenarios. However, some may have been more important in one or the other scenario. During the debriefing of the pretest, no formal attempt was made to discuss skills deemed useful to score high in the posttest but which were not highlighted in the pretest. This was not done to avoid merely "teaching for the test." In addition, this would have been an additional teaching intervention different from the one we were interested in studying. Nevertheless, it is likely that such a strategy, or a repeated exposure to several scenarios, may have boosted the performance in the posttest and would have translated into a larger amount of improvement.

Our study has some limitations. We only tested the participants immediately after the intervention and did not evaluate the retention of their nontechnical skills. As mentioned previously, we provided all the participants with full debriefing after their posttest. This ensured a high participation rate in our study but prevented us from measuring the long-term impact of our intervention. This limitation is important, because one could argue that the benefit of videotape feedback may take some time to become apparent and may only have appeared after repeated sessions. Previous studies in medical education have demonstrated improved efficiency with videotape feedback when participants have repeated opportunities to review their performance.^{6,7} Another limitation is that we did not strictly control the duration of the debriefing time. Given the similarity in our debriefing times, it is probable that the video-assisted oral feedback group received less actual instruction given the time for videotape review. Controlling the amount of actual time given for instruction *versus* watching the videotape may have helped to clarify this issue. Finally, the provision of an effective debriefing is a difficult task that is more of an art than an exact science, because every instructor and educator possesses his or her own style and set of skills. Therefore, the generalizability of our results to other centers, to nontechnical skills outside the context of intraoperative cardiac arrest, or to a team-oriented training approach is unknown.

In conclusion, our study emphasizes the importance of providing feedback during simulation-based education. We have demonstrated that constructive feedback provided by skilled instructors is effective, but we did not observe extra benefit from adding a videotape review to the debriefing. These findings highlight the role of reflection and debriefing during simulation-based education. They suggest that effective teaching of nontechnical skills, pertaining to the management of intraoperative cardiac arrest, can be achieved even when video technology is not available, *e.g.*, when financial resources are

limited. Nonetheless, because of its theoretical advantages, it is still possible that videotape feedback may be a valuable adjunct to debriefing during simulation-based education. Videotape feedback may prove more useful for experienced faculty who misinterpret their performance and require a video record to facilitate change. Further research, both quantitative to measure the impact and qualitative to gain insight on how videotape feedback influences the learner, would guide optimal use and enrich understanding of this adjunct to simulation-based education.

The authors thank the anesthesiology residents of the University of Toronto (Toronto, Ontario, Canada) for their participation in this study.

References

1. Loyd GE: Simulation in anesthesiology, *Practical Health Care Simulations*. Edited by Loyd GE, Lake CL, Greenberg R. Philadelphia, Elsevier Mosby, 2004, pp 245-81
2. Gaba DM, Howard SK, Fish KJ, Smith BE, Sowb YA: Simulation-Based Training in Anesthesia Crisis Resource Management (ACRM): A decade of experience. *Simul Gaming* 2001; 32:175-93
3. Yee B, Naik VN, Joo HS, Savoldelli GL, Chung DY, Houston PL, Karatzoglou BJ, Hamstra SJ: Nontechnical skills in anesthesia crisis management with repeated exposure to simulation-based education. *ANESTHESIOLOGY* 2005; 103:241-8
4. Glavin R, Maran N: An introduction to simulation in anaesthesia, *Clinical Teaching: A Guide to Teaching Practical Anaesthesia*. Edited by Greaves DJ, Dodds C, Kumar CM, Mets B. Lisse, The Netherlands, Swets & Zeitlinger, 2003, pp 197-205
5. Mort TC, Donahue SP: Debriefing: The Basics, *Simulators in Critical Care and Beyond*. Edited by Dunn WF. Des Plaines, Illinois, Society of Critical Care Medicine, 2004, pp 76-83
6. Birnbach DJ, Santos AC, Bourlier RA, Meadows WE, Datta S, Stein DJ, Kuroda MM, Thys DM: The effectiveness of video technology as an adjunct to teach and evaluate epidural anesthesia performance skills. *ANESTHESIOLOGY* 2002; 96:5-9
7. Scherer LA, Chang MC, Meredith JW, Battistella FD: Videotape review leads to rapid and sustained learning. *Am J Surg* 2003; 185:516-20
8. Byrne AJ, Sellen AJ, Jones JG, Aitkenhead AR, Hussain S, Gilder F, Smith HL, Ribes P: Effect of videotape feedback on anaesthetists' performance while managing simulated anaesthetic crises: A multicentre study. *Anaesthesia* 2002; 57: 176-9
9. Gaba DM, Fish K, Howard SK: Basic principles of crisis management in anesthesiology, *Crisis Management in Anesthesiology*. Edited by Gaba DM, Fish K, Howard SK. Philadelphia, Churchill Livingstone, 1994, pp 5-47
10. Howard SK, Gaba DM, Fish KJ, Yang G, Sarnquist FH: Anesthesia crisis resource management training: Teaching anesthesiologists to handle critical incidents. *Aviat Space Environ Med* 1992; 63:763-70
11. Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R: Anaesthetists' Non-Technical Skills (ANTS): Evaluation of a behavioural marker system. *Br J Anaesth* 2003; 90:580-8
12. Cohen J: *Statistical Power Analysis for the Behavioral Sciences*. New York, Academic Press, 1977, pp 24-7
13. Gaba DM, Howard SK, Flanagan B, Smith BE, Fish KJ, Botney R: Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *ANESTHESIOLOGY* 1998; 89:8-18