Influence of Surgeon Behavior on Trainee Willingness to Speak Up: A Randomized Controlled Trial

Marco J. Barzallo Salazar, MD Howard Minkoff, MD Jyothshna Bayya, MD Brian Gillett, MD Helen Onoriode, MD Jeremy Weedon, PhD Lisa Altshuler, PhD Nelli Fisher, MD

PII: S1072-7515(14)01469-0
DOI: 10.1016/j.jamcollsurg.2014.07.933
Reference: ACS 7562

To appear in: Journal of the American College of Surgeons

Received Date: 20 March 2014
Revised Date: 8 July 2014
Accepted Date: 18 July 2014


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Influence of Surgeon Behavior on Trainee Willingness to Speak Up: A Randomized Controlled Trial

Marco J Barzallo Salazar, MD¹, Howard Minkoff, MD¹, Jyothishna Bayya, MD¹, Brian Gillett, MD², Helen Onoriode, MD¹, Jeremy Weedon, PhD³, Lisa Altshuler, PhD⁴, Nelli Fisher, MD¹

From the Departments of Obstetrics and Gynecology¹ and Emergency Medicine,² Maimonides Medical Center, Brooklyn, NY; Scientific Computing Center State University of New York Downstate Medical Center, Brooklyn, NY³, and Program for Medical Education Innovations and Research (PrMEIR) NYU School of Medicine, New York, NY⁴

Disclosure Information: Nothing to disclose.

Presented at the 33rd Annual Meeting of the Society for Maternal-Fetal Medicine, San Francisco, CA, February 2013.

Correspondence address:
Nelli Fisher, MD
Maimonides Medical Center, 967 48th Street
Brooklyn, NY 11219
Ph:718-283-7973, Fax:718-283-8468, email: nelli.fisher@gmail.com

Brief title: Interventions to Encourage Trainees to Speak Up
Abstract:

**Background**: To determine if surgeon’s behaviors can encourage or discourage trainees from speaking up when they witness a surgical mistake.

**Study Design**: A randomized clinical trial in which medical students (n=55) were randomized to an “encouraged” (E, n=28) or “discouraged” (D, n=27) group. Participants underwent personality tests to assess decision-making styles, and were then trained on basic tasks (“burn” then “cut”) on a laparoscopic surgery simulator. After randomization students assisted at a simulated laparoscopic salpingectomy. The senior surgeon used either an “encourage” (E) or a “discourage” (D) script (e.g.: E: “Your opinion is important.” D: “Do what I say. Save questions for next time.”). Otherwise, the surgery was conducted identically. Subsequently a surgical mistake was made by the senior surgeon who instructed students to cut without burning. Students were considered to have spoken up if they questioned the instruction and did not cut. Potential personality bias was assessed with two validated personality tests prior to simulation. Data was processed with Mann-Whitney and Fisher exact tests.

**Results**: The students in the encouraged group were significantly more likely to speak up (23/28, 82% vs. 8/27, 30%, P<0.001). There was no statistical difference between the two groups in personality traits, student training level (p=1.0) or gender (p=0.53).

**Conclusions**: A discouraging environment decreases the frequency with which trainees speak up when witnessing a surgical error. The senior surgeon plays an important role in improving intraoperative communication between junior and senior clinicians, and can enhance patient safety.
Introduction:

The publication of an Institute of Medicine (IOM) report in 1999—To Err is Human—signaled the beginning of the patient safety movement in the US (1). That report noted that as many as 100,000 iatrogenic hospital deaths occur annually in the US. Medical errors also occur in operating rooms, with some data suggesting that as many as half of all surgical complications are preventable (2, 3).

Many factors contribute to errors in hospitals, including system and individual failures (4, 5). Several solutions being implemented in hospitals were initially utilized in the aviation industry. Checklists, for example, were introduced after it was found that some plane crashes resulted from the large number of steps needed to assure safety on takeoff, steps that were less likely to be overlooked if they were put on a checklist. Checklists are now being used to reduce central line infections in ICUs (6). The concept of “collapsing the hierarchy” also had its origins in aviation. Studies of cockpit voice recorders have revealed many instances in which the failure of junior officers to “speak up,” even when they were aware of danger, contributed to air crashes (7). It has been suggested that the culture of the cockpit and the culture of the operating room are similar, and that the failure of all team members to “speak up” may contribute to surgical errors. The aviation industry recognized the importance of this issue, and now actively trains pilots to eliminate the atmosphere of intimidation that had been pervasive in cockpits.

There has not been a great deal of published research on surgeons’ attitudes, and the impact of their behavior on communication. Some reports have shown that hierarchy is as important a factor in medicine as it has been in the airline industry. However, one study reported that while most pilots (97%) and intensive care staff (94%) rejected steep hierarchies (in which
senior team members are not open to input from junior members), only 55% of consultant surgeons rejected them (5).

It has also been reported that hesitancy to speak up contributes to 23% of communication errors between junior and senior staff members (8). The American College of Surgeons’ executive leadership noted that clear communication between all members of the operating room is the best safety tool for preventing medical errors and encourages surgeons to take the lead in serving as facilitators in this process (9). The American College of Obstetricians and Gynecologists (ACOG) noted, when discussing the World Health Organization Surgical Safety Checklist, that it is inappropriate to place total reliance on the surgeon to identify the correct surgical site, or to assume that the surgeon should never be questioned. They went on to state that the risk of error may be reduced by involving the entire surgical team and encouraging any member of the team to address any possible error without fear of reprimand (10).

While it seems obvious that when a surgeon is blatantly rude trainees will not speak up, (11), the question of whether a negative message that is “subtle” (i.e., no ad hominem remarks, no raised voice, no menacing body language) can also inhibit communication remains unanswered. Furthermore, the degree of such influence has not been prospectively evaluated. In our study, we looked at how a surgeon’s negative, yet subtle, message affects trainee’s willingness to speak up. To do so, we compared what we considered an appropriate norm, (called the encouraged group), with what we considered a subtle but negative message (called the discouraged group). In this manner we sought to determine whether the senior surgeon could create an atmosphere in the operating room that either promoted or inhibited the willingness of trainees to “speak up” when they perceived a situation that could put a patient at risk.
A secondary goal was to evaluate how differences in medical trainees’ personality traits affected their action in response to a surgical mistake made by a senior surgeon.

**Materials and Methods**

We performed a prospective randomized controlled trial at Maimonides Medical Center after approval by the institutional IRB. Fifty-five medical students who were beginning their obstetrical rotation during 2011 and 2012 participated. The study was performed on two separate days, no more than one week apart (Figure 1).

A focus group of senior obstetrician/gynecologists created scripts based on a synthesis of their experiences in operating rooms that they felt might either encourage or discourage trainees from speaking up during a procedure. A pilot study was then conducted, with eight students randomized to an “encouraged” and nine to a “discouraged” arm. Randomization was achieved using a computer generated list of random numbers and sequentially numbered opaque envelopes with allocation cards designating group assignment.

The pilot study was performed to estimate effect size and thereby to facilitate power calculations. Given the large effect observed (88% of students spoke up in encouraged group and 11% in discouraged group [P<0.001]) a decision was made to introduce personality tests for the remainder of the study in order to determine whether un-assessed baseline personality characteristics drove outcomes. There were no other changes prompted by the pilot.

On day 1, as part of their orientation to the rotation, the 55 participating medical students had two reliable and validated assessments of response style, the General Decision Making Scale (GDMS) (12) and the Self-Construal Scale (SCS) (13) administered. The GDMS assesses how individuals typically make decisions. This measure, which has been shown to reflect a habitual
response style that is not context or problem specific, assesses five decision making styles (resulting in 5 subscales); avoidant (avoids making a decision), dependent (projects responsibility for decision making onto others), rational and intuitive (both assume personal responsibilities, one uses logic and the other hunches) and spontaneous (makes decisions rapidly). The scale consists of 24 items rated on a 5-point a Likert scale. The Self-Construal Scale measures an individual’s self-image, specifically, the relationship between self and others. Two distinct dimensions are tapped in this scale, including interdependence (subscale 1; e.g. reliance on group, respect for authority) and independence (subscale 2; e.g. individualism, direct expression of what one thinks). The scale consists of 24 items using 7-point ratings (1 = strongly disagree, 7=strongly agree). These tests assessed medical students’ cognitive and personality style related to their decision-making strategies within a social context, assertiveness and willingness to speak up, in order to assure that there were no differences between the two groups on these dimensions.

Later on day 1 the medical students underwent training on a virtual reality laparoscopic simulator (Lapsim), and were trained to perform two basic tasks: “burn tissue,” using electrocautery, and “cut tissue,” with scissors. They were also taught the “burn and cut” sequence, and were instructed that safety required them to always burn before cutting during a simulated laparoscopic salpingectomy. They were told that a large amount of blood loss occurs when a “cut” is used before a “burn.” This lesson was complemented by a visual display of profuse bleeding when the cutting occurred without first burning.

On day 2, the students were tested to ensure that the basic tasks had been learned, including the burn and cut sequence, prior to initiation of the intervention. This ensured that all the students had sufficient knowledge to be able to recognize a mistake that was scripted to occur while they
assisted on a subsequent laparoscopy. Then trainees were randomized to two groups: “encouraged to speak up,” and “discouraged from speaking up.” The students were blinded to the focus of the study. They were informed that the research projects’ aim was to assess trainees in laparoscopic simulation. The same senior surgeon worked with both groups.

In the “discouraged” group, what was considered an inappropriate environment was created in the following manner: the student who was asked to scrub overheard a nearby phone ringing, which was picked up by the surgeon who was going to conduct the simulation. The student heard the surgeon answering the phone, and stating, “I can’t talk right now. I have to do this student thing. They never told me about it, it’s not on my schedule. I’ll talk to you later.” As the student entered the simulated operating room, the same surgeon meeting the student at the door, said: “Let’s get this done. Just do what I say. Save your questions for next time.” At no time did the surgeon raise her voice, and during the procedure they made no negative comments. During the remainder of the simulation, the surgeon merely provided one word directions: “Burn” or “Cut”.

For students in the “encouraged” group an “appropriate” environment was created. No phone rang. The surgeon met the student at the door, shook hands with the student and said: “Hi Mr./Mrs. X, I am Dr. Y. We are going to do exactly what you have practiced. You are going to be my assistant and my extra pair of eyes. Everyone is human and can make mistakes. Your opinion is important. So speak up if something doesn’t look right to you.” During the remainder of the simulation, the surgeon did not talk to the student, except to provide one word directions: “Burn” or “Cut.”

In both scenarios the senior surgeon performed a laparoscopic salpingectomy on the laparoscopic simulator with the medical student’s assistance. The junior surgeon (medical student) had the scissors. During the surgery the senior surgeon ordered the student to cut the
tissue each time after burning. At one point during the surgery, a scripted surgical mistake was
made by the senior surgeon; she gave an order to cut the tissue even though the “burn” step had
been omitted. The surgeon was instructed to repeat the “burn” order until the student responded.
The student was considered to have spoken up if s/he refused to cut the tissue, clearly stating
that a burn is needed before a cut, or if the student refused to cut the tissue and asked the surgeon
about the need to burn before cutting. The surgeon then was instructed to agree that the burn is
needed.
Collected data was de-identified. The students were immediately debriefed after the simulation
and reassured that the simulation performance was not part of their evaluation for the rotation,
and that instructors involved in evaluation of the rotation did not see the results of their
performance.
Based on findings from our pilot data of the effect of the senior surgeon’s attitude on the
willingness of junior team members to speak up, we wanted to be able to demonstrate a 30%
increase in speaking up when the surgeon conveyed support as opposed to discouragement.
Given these considerations and with an alpha of 0.05 and a power of 80%, 21 students were
required per group. While we only required 21 students per group, we recruited slightly more in
order to include all the students who were in the cohort of students at our institution at the time
we reached our required sample size. Comparisons of groups for the action of speaking up and
for nominal demographic parameters were assessed with Fisher's exact test. Comparison of the
mean age between groups was performed with the unpaired t-test. Comparison of median values
for the assessed psychometric domains between groups was performed using the Mann-Whitney
test. Logistic regression was used to assess if the encouraged and the discouraged groups were
similar in their personality profiles, whether those who spoke up versus those who did not speak
up (irrespective of the encouraged/discouraged group to which they were assigned) were different in their personality profiles, and finally if taking the personality tests had, unto itself, in any way moderated the study arm effect.

Results
Of the 55 medical students who participated 28 were randomized to the “encouraged” group and 27 were randomized to the “discouraged” group. The students in the encouraged group were significantly more likely to speak up when witnessing an error during the simulated laparoscopic gynecological surgery (23[82%] out of 28 [100%] vs. 8[30%] out of 27 [100%]; p<0.001). In order to determine whether the process of completing the personality profile per se could influence subsequent actions (“priming”) we compared the pattern of responses of students who completed those questionnaires (the 55 students in the actual study) with the responses of the students that did not (the 17 students in the pilot study). In the cohort without personality tests (N=17), among the encouraged (N=8) seven (88%) spoke up and 1 (12%) did not speak up; and among the discouraged group (N=9) one student (11%) spoke up and eight (89%) did not speak up (p<0.001). In the cohort with personality tests (N=55), among the encouraged group (N=28), 23 (82%) spoke up and 5 (18%) did not speak up; among the discouraged group (N=27) eight (30%) spoke up and 19 (70%) did not speak up (p<0.001). While the magnitude of the intervention effect appeared slightly greater when personality tests were not administered, there was no statistically significant difference.

The demographic profiles and assessed personality tests of the medical students are listed in Table 1. There were no significant differences in demographic characteristics between groups.
There was also no difference between the encouraged and discouraged groups in personality profiles.

In addition, we did not find any difference in the personality profiles between those who spoke up and those who did not speak up, regardless of their intervention arm (encourage or discourage arm; Table 2).

Among those with personality assessments, logistic regression was used to determine whether the personality scales exerted interactive or confounding effects on the study arm effect. No interaction of any of the 7 personality sub-scales with study arm was detected ($p>0.10$ in all cases).

**Discussion**

In this randomized control trial we have found that the senior surgeon can play a critical role in establishing an environment that either fosters or inhibits their junior colleagues’ willingness to “speak-up.” An environment that we considered appropriate, in which students were treated as team members and explicitly asked to participate, resulted, as could be anticipated, in a high rate of speaking up. What was surprising was the dramatic effect of a surgeon appearing harried, and signaling that questions would not be appreciated—a not unrealistic scenario in some operating rooms—on rates of speaking up. Given the increasing evidence that a flattened hierarchy, in which the whole team feels empowered to speak up, contributes to a safer environment this is a potentially important finding, demonstrating how little it takes to create an inviting or intimidating environment.

It has been shown that disruptive behavior in perioperative settings and on obstetrical services has a direct effect on communication (14,15). In turn, poor communication is a major contributor to up to 60% of adverse events in patient care (16). Intimidation can also lead to an
unwillingness to speak up, and as such is an important barrier to communication and care (14,15,17). An unwillingness to speak up, or speaking up in a way that is ambiguous and downplays the significance of an emergency (mitigation), has been described as a key explanation for plane crashes (7). In his book “Outliers,” Malcolm Gladwell (7) noted that during black box conversations between senior and junior pilots, even in the setting of a threat to personal safety, junior crew members were hesitant to communicate clearly and assertively. Similarly, in the perioperative setting junior surgeons may be hesitant to speak up even when a patient’s safety is at issue.

The Institute for Safe Medication Practices (ISMP) defined intimidation as “any overt or covert interaction between health care professionals that results in either intended or unintended reluctance to speak up”(11). In an anonymous survey of surgery, orthopedic and gynecology residents and attending physicians, Belyansky et al had found that only 55% of residents felt they were encouraged to speak up. However, all the attending physicians thought that they encouraged residents to speak up (18). Dull and Fox surveyed surgeons about their perception of intimidation in a perioperative setting. They reported that 68% of surgeons did not consider reluctance to answer questions as intimidation, 55 % thought impatience with questions did not constitute intimidation and 59% of surgeons did not think that saying, “Just give me what I ordered” was intimidation (19). These behaviors are very similar to those that were scripted in our “discouraged” group. We found that when these messages are given, even when the senior surgeon does not behave abusively and isn’t visibly angry, students do not speak up despite seeing a surgeon do something that they know is inappropriate.

Merely providing personal examples to surgeons of what constitutes intimidation does not lead to change (17). In order improve the culture of the health care setting, a great deal of
effort has been expended to introduce Crew Resource Management or a TEAMSTEPS communication education, with the goal of improving communication, combating mitigation and increasing patient safety. We have demonstrated that a simple first step in creating an “encouraging environment” can be achieved by a senior surgeon simply saying a few sentences before a case, encouraging junior team members to speak up. Additionally, as part of medical training, junior team members may need skills development not just about procedures (e.g. “burn then cut”) but also about how to question those in authority.

One of the strengths of our study was the randomized control single-blinded design. While many studies assessing attitudes and perceptions and the effects of hesitancy to speak up on patient safety are surveys, we utilized a simulated operating room environment and a laparoscopic simulator to more nearly approximate a clinical environment. Another strength was that we showed that there were no differences in personality profiles between our encouraged and discouraged groups, as well as between those who spoke up and did not speak up, thereby eliminating an important potential confounder. This suggests that situational factors are more potent than individual differences in personality style or decision-making strategies in promoting or inhibiting behavior within the operating room environment.

We also considered whether we “primed” the students because we utilized a personality test prior to the experiment. In other words it is possible that being asked, for example, about predilections to respect authority could make one more likely, or more hesitant, to do so in an experimental environment. Although it appears that among students in the discouraged groups those who received personality tests were more likely to speak up (30%) than students who did not (11%), the opposite trend, albeit much more subtly, was observed among those in the encourage groups: without personality tests 88% spoke up, with personality tests 82% spoke up.
However, neither trend was significant, and if there were a priming effect, both the encouraged and discouraged groups should have demonstrated an increase in the rate of speaking up after administration of personality tests. In addition, to minimize the possibility of priming precautions were taken to administer personality tests as part of the orientation package, by a person not related to the study, at a different time of the day than when the Lap Sim training was performed.

The study also had limitations. The subjects were medical students who were meeting the “senior surgeon” for the first time, and there was only the student and the surgeon in the room. There may be a difference in how a senior surgeon’s behavior affects residents who have a longstanding relationship with the surgeon. However, results from the airline industry do not support the supposition that intimidation disappears with familiarity. Another limitation was that students from throughout their first clinical year were included in our study, albeit all were at the very beginning of their ob/gyn clerkship. While we did not specifically look at the difference between novice clinical rotators and more experienced students, within each level of experience (i.e., at the start of the rotation students at the same level) students were randomized to encouraged and discouraged groups. The randomized nature of the study thus controlled for students’ experience and reduced potential bias.

Another limitation was that the study was not double blinded. The same surgeon-actor played the role of an “encouraging” and “discouraging” surgeon during the few minutes of greeting. While efforts were made to avoid any variation in the environment other than the scripted words, slight variations (such as body language or tone of voice) may have occurred. However, we did pilot the surgeon’s performance, and filmed several “rehearsals” in order to make sure that the approach was as uniform as possible.
We conclude that the senior surgeon plays an important role in improving communication with junior trainees in the operating room. Very simple messages—either positive or negative--can have powerful effects on the rate at which junior members of the surgical team are willing to point out potential surgical mistakes. Therefore an encouraging environment should be created by senior staff and shared with all team members prior to performing a procedure. We believe that this will increase all team members’ willingness to speak up when concerns appear. This should increase patient safety in the operating room and improve communication between team members. Applying these encouraging behaviors inside the operating room may increase patient safety without increasing costs, and should improve the spirit of the surgical team.

Further studies could help to determine if specific scripts, when added to pre-operative checklists, can improve communication during surgery.
References:


4. JCAHO Sentinel Alert Issue 30, July 21, 2004


Table 1. Demographic Profiles And Personality Tests of “Encouraged” and “Discouraged” Participants

<table>
<thead>
<tr>
<th></th>
<th>Encourage (n=28)</th>
<th>Discourage (n=27)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of 2nd year students</td>
<td>13</td>
<td>9</td>
<td>0.4</td>
</tr>
<tr>
<td>No. of 3rd year students</td>
<td>15</td>
<td>18</td>
<td>0.4</td>
</tr>
<tr>
<td>Mean age, y</td>
<td>26</td>
<td>25</td>
<td>0.2</td>
</tr>
<tr>
<td>Males, n</td>
<td>12</td>
<td>12</td>
<td>1.0</td>
</tr>
<tr>
<td>Females, n</td>
<td>16</td>
<td>15</td>
<td>1.0</td>
</tr>
<tr>
<td>DMS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rational subscale, median</td>
<td>4.3</td>
<td>4.0</td>
<td>0.37</td>
</tr>
<tr>
<td>Intuitive subscale</td>
<td>3.8</td>
<td>3.8</td>
<td>0.08</td>
</tr>
<tr>
<td>Dependent subscale</td>
<td>4.0</td>
<td>3.8</td>
<td>0.52</td>
</tr>
<tr>
<td>Avoidant subscale</td>
<td>3.8</td>
<td>3.0</td>
<td>0.69</td>
</tr>
<tr>
<td>Spontaneous subscale</td>
<td>3.2</td>
<td>3.0</td>
<td>0.55</td>
</tr>
<tr>
<td>Self-construal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent subscale</td>
<td>5.3</td>
<td>5.3</td>
<td>0.57</td>
</tr>
<tr>
<td>Interdependent subscale</td>
<td>5.5</td>
<td>5.5</td>
<td>0.83</td>
</tr>
</tbody>
</table>

DMS, decision making scale.
Table 2. Personality profiles of students who spoke up and those who did not speak up

<table>
<thead>
<tr>
<th>Decision making scale</th>
<th>Speak up Yes, n=31</th>
<th>Speak up No, n=24</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rational subscale, median</td>
<td>4.3</td>
<td>4.3</td>
<td>0.40</td>
</tr>
<tr>
<td>Intuitive subscale</td>
<td>3.8</td>
<td>3.8</td>
<td>0.61</td>
</tr>
<tr>
<td>Dependent subscale</td>
<td>4.0</td>
<td>3.8</td>
<td>0.76</td>
</tr>
<tr>
<td>Avoidant subscale</td>
<td>3.6</td>
<td>3.2</td>
<td>0.99</td>
</tr>
<tr>
<td>Spontaneous subscale</td>
<td>3.0</td>
<td>3.3</td>
<td>0.86</td>
</tr>
<tr>
<td>Self-construal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent subscale</td>
<td>5.3</td>
<td>5.5</td>
<td>0.49</td>
</tr>
<tr>
<td>Interdependent subscale</td>
<td>5.5</td>
<td>5.5</td>
<td>0.55</td>
</tr>
</tbody>
</table>
Figure 1. Flow of participants through the study.
Precis

The behavior of the senior surgeon can create an environment that encourages or discourages trainees from speaking up when they see a potential surgical mistake.
Pilot Study: Medical Students
N=17

Medical Students
N=55

Simulation Training

Simulation Assessment

Randomization

Encourage to speak up
N=28

Discourage to speak up
N=27

Simulated surgery with mistake

Debriefing

Personality Tests
N=55