
Abstract

**Background:** A consistent finding in health care is that interns (junior physicians) are unwilling to “speak up” to attendings or consultants (senior physicians), which involves the related behaviors of challenging or questioning the actions of attendings, seeking guidance about one’s own performance, and reporting one’s errors. Training designed to encourage interns to speak up was designed, conducted, and evaluated at University Hospital Galway, an academic teaching hospital affiliated with the National University of Ireland, Galway, in January 2012–March 2012 and November 2012–January 2013. **Method:** Training, which was based on the Crew Resource Management model, was constructed around filmed stories of attending physicians describing situations in which, when they were interns, their communication and assertiveness skills were challenged, and their reflections on what they could have done better. **Results:** A total of 110 interns attended the training. The feedback from participants was positive. There was a significant increase in knowledge as a result of the training, and some evidence to support a shift in attitudes in the desirable direction relating to the need to speak-up to seniors. No effect of the training was found on behavior. **Conclusions:** The willingness to share examples of poor performance is an important part of building a good safety culture. Listening to the stories of attendings describing and reflecting on a situation in which they struggled is a powerful teaching method for training interns in the communication and assertiveness skills that are critical to their job. However, a sustained change in attitudes and behavior to speaking-up will require a reinforcement of learning with deliberate practice to develop the skills introduced through the training program.
Introduction

Research shows that high-performing teams have a climate of openness and trust, in which team leaders are receptive to alternative views and team members are not afraid to express them.\(^1\) However, a consistent finding in health care research is that status differences between team members are detrimental to open communication.\(^2\)\(^-\)\(^6\)

More than just a willingness to question an attending or consultant (senior) physician, speaking up also relates to an individual’s willingness to inform attendings when they are struggling or stressed or when they have made an error. A number of studies across different domains of health care have shown that junior personnel are unwilling to speak up or challenge the behavior of more senior personnel,\(^2\)-\(^4\),\(^7\),\(^8\) seek guidance and clinical support from supervisors,\(^9\) or report errors.\(^10\) Moreover, attendings seem to be unaware of this reluctance on the part of interns and residents.\(^2\),\(^4\),\(^6\),\(^8\) Unsurprisingly, failure to speak up has been shown to have a negative impact on patient safety and quality of care.\(^5\),\(^7\),\(^11\)-\(^13\) If safety and quality of patient care is dependent on teamwork, communication, and a collaborative work environment, health care organizations must address the problem of behaviors that threaten the performance of the health care team.\(^14\)

The literature on the reasons behind the unwillingness on the part of interns and residents to speak up suggests that the hierarchical structure that exists within health care organizations is largely to blame.\(^2\),\(^15\) Yet, other hierarchical organizations appear to have managed this problem to an extent that is not yet realized in health care.
One particular community that has been successful in implementing a culture of empowerment and openness among those holding junior positions is U.S. Naval aviation. In contrast to physicians, the majority of both junior and senior U.S. Naval aviators agree that junior personnel are not afraid to speak up.\(^4\) However, this has not always been the case. For example, in a survey of 216 military aviation mishaps from the 1980s it was found that as the military rank of the captain of the aircraft increased, so did the likelihood that unassertiveness was a causal factor in the mishap.\(^{16}\) Further evidence of a lack of assertiveness in junior personnel comes from simulator studies conducted in the same decade. It was found that when civilian aircraft captains feigned incapacity during a final landing approach, 25% of the planes “crashed” because the copilot failed to take over control.\(^{17}\)

One of the major changes that occurred in the late 1980s in U.S. naval aviation (and civil aviation) was the introduction of human factors training called Crew Resource Management (CRM).\(^{18}\) The goal of the US Navy’s CRM training is to improve mission effectiveness by minimizing crew-preventable errors, maximizing crew coordination, and optimizing risk management.\(^{19}\) Speaking up is an important component to the training. Every naval aviator must receive CRM training and an evaluation of their CRM skills during an actual or simulated flight once a year.\(^{20}\)

In addition to the acknowledgement of the importance of good teamwork to safe operations, there is also recognition by leadership that sharing examples of poor performance is an important part of building a good safety culture. Discussing individual examples of near misses, also known as “close calls,”\(^{21}\) is an important component of CRM training. For example, *Approach* magazine\(^{22}\) includes stories written by naval aviators describing real-life situations in which a mishap was narrowly avoided, along with reflections on why the
incident occurred. It is important to note that these stories are not anonymous and are available on the internet. This is just one indication of the openness within the aviation culture for exposing errors and learning from them.

The two principles of speaking up and sharing examples of poor performance formed the basis of a training program for interns in the Republic of Ireland (where, as in the United States and other countries) internship is the first year of practice following completion of an undergraduate medical degree. In this article, we describe a training needs assessment, which was followed by development and evaluation of the training program. The aim of the training program was threefold, as follows:

1. To demonstrate that speaking up is important for patient safety and quality of care
2. To show interns that attending physicians want them to speak up despite any possible discomfort
3. To provide the interns with techniques for speaking up on the basis of strategies used in aviation

**Methods**

**Setting**

The study was carried out with interns from the West North West intern training network in the Republic of Ireland. The training was carried out at University Hospital Galway, an academic teaching hospital affiliated with the National University of Ireland, Galway, during two time periods, January 2012–March 2012 and November 2012–January 2013.
**Training Needs Assessment**

The identification of a need for interns to receive training in speaking up was established through a broader training needs assessment of the nontechnical skills required by interns. This assessment was conducted by the authors from July 2011 until November 2011 at University Hospital Galway. A number of sources of information were assessed, as follows:

1. Content of team training used in health care, aviation, and other high-risk domains
2. Research on teamwork and nontechnical skills in health care professionals, with a particular emphasis on interns or residents

The standardized patient scenarios were obtained from the Program for Medical Education Innovations and Research (PrMEIR), New York University Langone Medical Center, which we adapted.

The assessments proceeded as follows: the intern reads the introductory text that sets the scene for the scenario and introduces the challenging situation. The intern then must proceed to manage the situation presented in the scenario, with actors playing the person with whom they are interacting. The intern was observed by either a physician or psychologist, who completed a rating form specifically designed for each scenario, with between 10 and 20 marker behaviors (for example, "gave the patient an opportunity to emotionally respond," ‘identified preventative strategies’). For each behavior, the observer recorded it as “not done,” “partially done,” or “well done.” Following the completion of the scenario, the observer provided feedback to the intern regarding his or her performance. The overall performance on each of the scenarios is summarized in Figure 1. Only between one third and one half of the interns performed well in these exercises, with particular deficiencies identified in
communication and assertiveness. The needs assessment resulted in the identification of a need for training for interns in the following nontechnical skills: situation awareness, decision making, team working, leadership, communication (including speaking up), managing stress, coping with fatigue, and empathy. However, given the limited time that was available for training, we decided to concentrate on speaking up, a skill that we judged to be particularly important for interns.

**Figure 1.** Only between one third and one half of the interns performed well in the five standardized patient exercises, with particular deficiencies identified in communication and assertiveness.
Design of Training Intervention

Filmed Stories of Challenging Situations. The training strategy was based on CRM training in aviation, in which filmed reenactments of aviation mishaps are commonly used to demonstrate examples of good and poor performance.\(^\text{18}\) We determined that it would be too difficult and costly to produce reenactments of medical mishaps and that recounting of stories would be an acceptable strategy. We made personal requests to eight attending physicians in a broad range of specialties—general surgery, general medicine, pediatrics, and geriatric medicine, as well as medical and surgical subspecialties—to describe particularly challenging situations that occurred when they were interns or residents and to reflect on what they could have done better. Of six attending physicians’ filmed we selected four stories, we selected the four stories with a particularly strong focus on assertiveness and the need for interns to speak up were selected for the training sessions.

Format. Interns’ busy schedules limited the time available for participation in the training program to 90 minutes. To facilitate discussion, we chose a small-group teaching method for the program, which consisted of the following segments:

- A brief verbal and slide presentation introducing the concept of human factors and human error in health care
- A review of the nontechnical skills required of an effective intern
- Instruction in communication techniques adapted from aviation (see Brindley and Reynolds\(^\text{24}\) for a review of these strategies)
- Presentation of the four attending physicians’ film clips followed by discussions of each scenario. The discussions focused on how communication and assertiveness were implicated in events and the challenges associated with speaking up and how to overcome them.
Training Evaluation

The training was evaluated at the first three levels of Kirkpatrick’s training evaluation hierarchy: (1) reactions, that is, whether participants enjoyed the training and whether they found it relevant and useful; (2) learning—which has two components: assessment of knowledge change and assessment of attitude change; and (3) behavior—whether knowledge learned in training actually transferred to behaviors on the job or a similarly simulated environment.

Reactions. We measured reactions by using a nine-item reactions questionnaire distributed immediately after the training. For five of the items, such as “how interesting did you find the course?,” the participants could respond on a five-point Likert scale (1 [“poor”] to 5 [“excellent”). The remaining four items were open-ended questions, such as “what aspects of the course should there have been more time devoted to?,” focused on the content of the course.

Learning. We assessed knowledge by using a combination of 11 multiple-choice and fill-in-the-blanks questions, based on the content of the training, such as “Give two examples of nontechnical skills.” We measured attitudes using two factors (12 items) from a questionnaire used to assess the attitudes of surgeons,4 U.S. Naval aviators,26 and U.S. Naval Officers of the Deck20—Speaking Up About Stress (for example, “Team members should alert others to their actual, or potential, work overload”) and Speaking Up to Seniors (for example, “Junior team members should not question the decisions made by senior personnel during routine situations”); these measures are available online.27

Behavior. We evaluated behavior using two standardized patient scenarios, both involving breaking bad news, from the PrMEIR. A medical error scenario was concerned with explaining to a patient that during a hernia repair operation the surgeon had initially started the procedure on the wrong side. The scenario required the intern to use the
nontechnical skills of situation awareness, communication, decision making, leadership, teamwork, managing stress, and empathy. An unexpected death scenario was concerned with informing the daughter that her father had died unexpectedly during cardiac surgery. This scenario required the intern to use the nontechnical skills of situation awareness, communication, decision making, team working, leadership, managing stress, and empathy. Although medical error and disclosure issues were discussed in the training, breaking bad news was not explicitly addressed. We wanted to assess whether the concepts addressed in the course generalized to other situations in which communication is important. The interns were observed by a psychologist using a rating form specifically designed for each scenario (see the introduction for more discussion of the ratings). The order of the scenarios was counterbalanced pre- and posttraining.

**Evaluation Design and Analysis**

Evaluations were carried out pre-training, posttraining, and with a control group (Figure 2). Including a control group of interns that had not attended the training allowed a comparison to be made with interns of the same level of experience as the post-training group of interns.
Figure 2. Training was evaluated at the levels of reactions, learning (knowledge and attitude), and behavior for the pre- and posttraining groups and the control group.

The evaluation design was complicated by the fact that not all interns completed both the pre- and posttraining evaluations. To address this issue, we carried out the following comparisons:

- For those interns that completed both the pre- and posttraining attitude and knowledge assessments, we matched their responses and compared their performance using paired-subjects t-tests.
- Using between-subjects analyses of variance (ANOVAs), we compared the interns who completed only the pretraining attitude and knowledge assessment with all interns who completed the posttraining assessment and with the interns in the control group.
• Similarly, because the participants did not complete the same standardized patient exercises at pre-and posttraining, we used between-subjects ANOVAs to compare performance on each exercise.

We measured internal consistency of the attitudes scale with the Cronbach’s alpha statistic. Two psychologists examined the responses to the qualitative questions and used the data to develop a set of categories for classifying the responses.

Results

Number of Participants

A total of 110 interns attended the training in groups of 6 to 12. Although attendance at the training was a mandatory part of the intern curriculum, participation in the evaluation was voluntary. The participants provided consent to participate in the study, and ethical approval for the study was provided by the chairman of the ethics board of Galway University Hospitals.

Reactions

A total of 100 of the 110 interns (90.9% response rate) completed the reactions questionnaire. As shown in Table 1 (page 000), the mean value of was 4.35 (with a maximum of 5). A total of 69 qualitative statements were made concerned with what aspects of the course should be emphasized more (51 statements) or less (18 statements). A total of 39.1% of the participants wanted more film clips of attending physicians to be included, and 17.4% wanted more discussion on the nontechnical skills required by interns. A total of 8.7% of the participants wanted less discussion of non-technical skills required by interns.
A total of 55 responses were made to the question asking what else should be included in the training. A total of 16.3% of the participants suggested film clips of interns discussing a challenging situation, and 5.5% suggested including role-play exercises.

**Knowledge**

A total of 92 interns completed the knowledge test before the training (83.6% response rate), and 53 completed the knowledge test after the training (48.1% response rate). The knowledge test was completed by 17 interns from the control group (48.6% response rate). As shown in Table 1 (page 000), there was a significant improvement in performance on the knowledge test after the training as compared with before the training. The posttraining group also had a significantly higher number of correct answers than the control group.

**Attitudes**

A total of 92 interns completed the attitude test before the training (83.6% response rate), and 54 completed the attitude questionnaires after the training (49.0% response rate). The attitude survey was completed by 17 interns from the control group (48.6% response rate). The two factors assessed were: Speaking Up About Stress” and Speaking Up to Seniors. The Cronbach’s alpha for the Speaking Up About Stress factor was 0.60 in the pretraining, 0.58 in the posttraining, and 0.54 in the control group samples. For the Speaking up to Seniors factor, the Cronbach’s alpha was 0.60 in the pretraining sample, 0.67 in the posttraining sample, and 0.50 in the control group samples—reliability scores that are fairly typical for this type of questionnaire.⁴,²⁶
As shown in Table 1, there was no significant effect of training on Speaking up About Stress. For the Speaking Up to Seniors factor, although the repeated measures assessment of the effect of training was not significant, there was a significant effect in the between-subjects assessment. The posttraining group had significantly more positive attitudes toward the need to speak up to seniors than the pretraining group. However, there was not a significant difference between the pre- or posttraining groups and the control group. Although not significant, the effect size of the difference between the control group and posttraining group was a medium to large (Cohen’s d = 0.47), and the size of the difference pre- and post for the repeated measures assessment was medium (Cohen’s d = 0.33).28
<table>
<thead>
<tr>
<th>Evaluation (n)</th>
<th>Pretraining Mean (SD)</th>
<th>Posttraining Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reactions (100)</td>
<td>-</td>
<td>4.35 (0.45)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Repeated Measures**

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Pretraining Mean (SD)</th>
<th>Posttraining Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge (45)</td>
<td>43.83 (16.3)</td>
<td>63.84 (14.44)</td>
<td>-</td>
<td>t = 6.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(df = 44, p &lt; .05)</td>
<td></td>
</tr>
<tr>
<td>Attitude</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Speaking Up About Stress (46)</td>
<td>3.87 (0.41)</td>
<td>3.88 (0.44)</td>
<td>-</td>
<td>t = 0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(df = 45, ns)</td>
<td></td>
</tr>
<tr>
<td>Speaking Up to Seniors (46)</td>
<td>2.97 (0.74)</td>
<td>3.20 (0.68)</td>
<td>-</td>
<td>t = 1.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(df = 45, ns)</td>
<td></td>
</tr>
</tbody>
</table>

**Between Subjects**

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Pretraining Mean (SD)</th>
<th>Posttraining Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge (47-pre, 66-post, 17-control)</td>
<td>44.68 (16.61)</td>
<td>59.22 (14.81)</td>
<td>43.85 (9.22)</td>
<td>F(2, 127) = 15.91, p &lt; .05‡</td>
</tr>
<tr>
<td>Attitude</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Speaking Up About Stress (46-pre, 67-post, § 17-control)</td>
<td>3.91 (0.53)</td>
<td>3.88 (0.42)</td>
<td>3.81 (0.42)</td>
<td>F(2, 127) = 0.30, ns</td>
</tr>
<tr>
<td>Speaking Up to Seniors (46-pre, 67-post, § 17-control)</td>
<td>2.79 (0.68)</td>
<td>3.27 (0.64)</td>
<td>2.98 (0.69)</td>
<td>F(2, 127) = 7.2, p &lt; .05</td>
</tr>
<tr>
<td>Behavior</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Admitting Error</td>
<td>23.45 (1.51)</td>
<td>21.04 (2.85)</td>
<td>21.07 (3.69)</td>
<td>F(2, 45) = 2.9, ns</td>
</tr>
<tr>
<td>Breaking Bad News</td>
<td>27.63 (2.79)</td>
<td>27.65 (3.33)</td>
<td>25.37 (3.07)</td>
<td>F(2, 44) = 1.8, ns</td>
</tr>
</tbody>
</table>

*SD, standard deviation; df, degrees of freedom; ns, nonsignificant.
†Paired-subjects t-tests.
‡Between-subjects analyses of variance (ANOVAs).
§The posttraining figure is derived from adding the number of interns who completed the posttraining knowledge and attitude questionnaires in the repeated-measures design to the number of interns who only completed the posttraining questionnaires.
Behavior

Because of the logistical difficulties of assessing the behavior, a random selection of 50 interns in the pre- and posttraining groups and 35 from the control group were asked to participate, A total of 30 interns (60% response rate) completed one of the standardized patient exercises before the training, 43 (86% response rate) after the training, and 22 interns from the control group (62.9% response rate). As can be seen in Table 1, no significant differences were found as a result of the training for either exercise.

Discussion

In a training program to promote speaking up among interns, which was based on a training needs assessment and evidence from the literature, the reactions to the training were generally positive. The mean value of 4.35 on the reactions questionnaire compares favorably with the mean of 4.18 for reactions to training reported in a meta-analysis of CRM training effectiveness.23

The attending physicians’ stories were particularly well received, with almost half of the interns suggesting that more of them should be included in the training. The stories not only served as a method for demonstrating why speaking up is important for patient safety and quality of care but also demonstrated a willingness on the part of attending physicians to discuss a situation in which they had not performed as well as they would have liked. Unlike in aviation, this is something that persons in junior positions are not used to hearing.

A significant effect of the training on knowledge was found between the posttraining group and the pretraining and control groups. Knowledge gain is not commonly reported in evaluations of this type of training. For example, only 4 of 16 CRM evaluation studies
identified by O’Connor et al\textsuperscript{29} reported an assessment of the knowledge of participants. Yet, an assessment of knowledge is important; no changes in behavior or attitudes can be expected if no new knowledge has been learned.

The training was not found to have a significant effect on speaking up about stress. Despite discussions on these issues in the training, it would appear that a 90-minute training course was insufficient to result in changes in attitudes towards speaking up and compensating for stress in other team members. This finding is similar to results reported for physicians\textsuperscript{4,30} and aviators alike.\textsuperscript{20}

Evidence to support a shift in attitudes towards the need to speak up to seniors was found in the significant difference pre- and posttraining in the between-subjects comparison and a medium-effect size in the difference for the repeated measures and the control-versus-posttraining comparison.

The fact that the training did not have an effect on the interns’ performance in the two standardized-patient exercises may reflect the fact that was far shorter than the two or three days typical of CRM training.\textsuperscript{26} Also, the interns were not given the opportunity to practice and receive feedback on their performance as part of the training. A sustained and large change in attitudes and behavior will require reinforcement of learning through practice and feedback. Another limitation of the study is the lower than desirable response rate to the posttraining questionnaires, which reflects the difficulty in obtaining responses from busy interns moving between specialties and hospitals.
Conclusion

This willingness to share examples of poor performance, an important part of building a safety culture, is, in contrast to aviation, infrequent in health care, where it tends to be done behind closed doors with a select group of individuals—and is certainly not recorded. The personal stories described by the attending physicians in this study illustrated the detrimental effects that failing to speak up can have on patient safety and quality of care. Moreover, the reflections of the attending physicians on their performance in these scenarios demonstrated that they want interns to speak up when they have concerns. Simple stories can be a powerful training tool to ensure the same mistakes are not repeated and reinforce a culture where openness is valued.

This research was partially supported through funding by the Health Services Executive (HSE), and the Irish Social Sciences Platform (ISSP) under the Program for Research in Third Level Institutions, administered by the Higher Education Authority and co-funded under the European Regional Development Fund (ERDF). The authors thank Dr Sondra Zabar, Director of Program for Medical Education Innovations and Research (PrMEIR), New York University Langone Medical Center, New York City, for allowing the use of the standardized patient scenarios.
References


