

## The Effect of Previous Resident Interactions on the Assessment of Interpersonal and Communication Skills by Teaching Faculty: Are We the Best Evaluators?

Andrew B. Casabianca, MD<sup>1</sup>

Jeffrey S. Berger, MD<sup>2</sup>

Thomas J. Papadimos, MD<sup>3</sup>

Amy Capwell-Burns, Ph.D.<sup>4</sup>

### Original Article

<sup>1</sup> Associate Professor of Anesthesiology, University of Toledo College of Medicine

<sup>2</sup> Associate Professor of Anesthesiology, George Washington University School of Medicine

<sup>3</sup> Professor of Anesthesiology, Ohio State University College of Medicine

<sup>4</sup> Assistant Professor of Communication, University of Toledo

### Abstract

**Background:** Evaluating resident interpersonal and communication skills (ICS) presents a significant challenge. Unlike the In-Training-Exam, an objective measure of knowledge, the evaluation of ICS is subjective. Previous interactions could influence how teaching faculty evaluate this competency leading to inaccurate assessment of resident ICS. Faculty groups from other residencies and non-physicians were enlisted to compare assessments with those by teaching faculty.

**Methods:** A cross-sectional study was conducted comparing how different evaluator groups assessed the ICS of anesthesiology residents. Nine residents participated each in two Standardized Patient (SP) encounters that were video-recorded. The recordings were viewed by eleven evaluators representing four different evaluator groups, one non-blinded teaching faculty group, two blinded anesthesiology faculty groups from separate programs and one blinded non-physician group. They scored each encounter using a modified SEGUE framework evaluation form graded on a Likert scale.

**Results:** The mean score for each resident ICS encounter by evaluator group were as follows: non-blinded teaching faculty (57.89), non-physician group (57.42), and the blinded anesthesiology faculties (53.00) and (53.83) respectively. There was significant difference in how the evaluator groups scored the resident performances ( $p < 0.001$ ). Analysis of ranks showed excellent correlation comparing teaching faculty with the other anesthesiology faculty groups ( $r = 0.764$ ,  $p = 0.017$  and  $r = 0.765$ ,  $p = 0.016$ , respectively). The highest ranked resident overall ranked high across all evaluator groups and the lowest ranked resident was ranked lowest across most evaluator groups.

**Conclusions:** Though potential for biases from previous interactions exist, teaching faculty assessments of resident ICS are similar to the assessments of other anesthesiology faculty evaluator groups.

**Key words:** Communication skills, resident interactions, assessment of interpersonal, best evaluators

**Correspondence to:**

Andrew B. Casabianca, M.D.  
(corresponding author)

University of Toledo Medical  
Center

Room 2195 Hospital Building  
3000 Arlington Avenue

Toledo, Ohio 43614-2598

Tel: 419-383-3556

Fax: 419-383-3550

Email:

[Andrew.Casabianca2@utoledo.edu](mailto:Andrew.Casabianca2@utoledo.edu)

Financial support: Funding for this study was provided by the Department of Anesthesiology of the University of Toledo College of Medicine to the extent of the costs for the standardized patients and materials for reproducing the video-recordings and mailings involved in distributing them to the outside anesthesiology faculty groups.

## Manuscript

### Introduction

There is little doubt that effective Interpersonal and Communication Skills (ICS) are essential attributes for all physician practitioners. There has been increased emphasis on ICS at all levels of training and practice. The Accreditation Council for Graduate Medical Education (ACGME) as part of its Outcomes Project established ICS as one of the core competencies for evaluating resident performance.<sup>1</sup> Medical students as part of the United States Medical Licensing Examination (USMLE) are tested on their ICS as an integral part of the Step 2 Clinical Skills Exam.<sup>2,3</sup> There is a growing trend by hospitals to include ICS evaluations as part of their credentialing and re-credentialing process.<sup>4</sup>

During residency teaching faculty are responsible for the evaluation and teaching of this core competency. Using patient assessments as recommended by the ACGME Toolbox may not be practical. Anesthesiology residents often meet the patient immediately before surgery without the benefit of a long-standing patient-physician relationship. That encounter fades among the multiple other interactions the patient may have with attending staff, other residents, medical students and nursing providers. Relying on teaching faculty assessments of resident ICS also presents a significant challenge. Unlike knowledge which can be measured objectively by test scores such as the In-Training-Exam (ITE) or the Anesthesia Knowledge Test (AKT), the evaluation of ICS is more subjective. Our previous use of video-recorded Standardized Patient (SP) encounters in the format of an Objective Structured Clinical Exam (OSCE), also recommended by the ACGME Toolbox, was encouraging as we found strong agreement between teaching faculty evaluators and the SPs, as well as excellent inter-faculty correlation.<sup>5</sup>

However, we were always concerned that the faculty evaluators were not blinded to the residents.. As residents progress along their educational continuum, teaching faculty develop opinions, impressions and evaluations of their performance. These preconceived assessments can potentially influence teaching faculty when it comes time to evaluate resident ICS. Therefore, faculty members' prior resident interactions could create a bias and influence how they evaluate this competency, potentially leading to inaccurate assessment of resident ICS. In the process of evaluating and scoring a resident's present performance, faculty perception of resident previous performance can be affected by experiences with residents inside and outside the operating room. Murphy<sup>6</sup> suggests that raters who develop systematic expectations regarding the performance of a specific resident may find it difficult to accurately evaluate that resident's performance if he or she departs from previous patterns of performance. An *assimilation* effect may occur in which evaluations of present performance will be biased in the direction of previous evaluations or a *contrast* effect in which evaluations of present performance will be biased in a direction opposite to that of previous evaluations. It is also possible that anesthesiology faculty as evaluators may be pursuing different goals<sup>7</sup> in completing their performance appraisals. Some may be overly harsh and use the evaluation process as an opportunity to challenge and motivate residents. Others may be too lenient in an effort to preserve harmony within the residency or the fear that a poor resident evaluation would be reflected in an equally poor faculty evaluation. Additionally, faculty may feel that resident performance is a surrogate measure of their own ability to teach. Finally we must also consider

that faculty may not have the necessary skills, appropriate rating scales or information to accurately evaluate resident performance.

Our study attempts to investigate the impact of prior resident interactions on ICS scores. To investigate this, blinded anesthesiology faculty from other residency programs and non-physicians were asked to participate in this study. We wanted to see how groups free of potential biases from previous interactions would score the resident SP encounters and compare those assessments to the assessments done by the teaching faculty. The goal was to establish whether or not previous resident interactions influenced how non-blinded teaching faculty evaluated resident ICS.

## Methods

A cross-sectional study was conducted to assess the ICS of anesthesiology residents and how previous interactions with teaching faculty would affect the scoring of SP encounters compared to evaluator groups without those interactions. Following Institutional Review Board (IRB) approval, two SPs were scripted for use in the ICS assessment. The format was designed to mimic a typical pre-anesthesia evaluation. The SPs were experienced veterans of prior ICS assessments. They were coached to respond with a chief complaint and history of present illness. The SPs were given specific instructions for answering questions during the review of systems establishing a past medical and surgical history. In addition, each SP was provided with a medication list and drug allergies as well as social, psychiatric and family histories. They were instructed and coached to portray specific behaviors, affects, and mannerisms. Information from the SPs was divided into that which was given freely and information to be given only if asked. Lastly for each encounter the SPs had three prompts/questions that had to be addressed during each interview. The exercise was conducted at the Ruth M. Hillebrand Clinical Skills Center on the Health Science Campus of the University of Toledo.

A total of nine residents participated in this ICS exercise representing all three categorical years in training (CA) [CA-3 (3), CA-2 (1), CA-1 (5)]. Each resident conducted two separate pre-anesthesia evaluations. Each SP encounter was standardized and time-limited to fifteen minutes. The residents were aware that the interviews were with SPs and the encounters were being video recorded for subsequent review.

A total of eleven evaluators independently reviewed the video recordings of the two resident SP encounters leading to a total of twenty-two separate assessments per resident. These evaluators represented four different evaluator groups as follows; one non-blinded anesthesiology group (University of Toledo) representing teaching faculty from the resident's home program (3), two blinded anesthesiology faculty groups (The Ohio State University and George Washington University) from outside and separate residency programs, (3, 3), and a blinded non-physician group (2) to capture the patient perspective. All evaluators scored each resident encounter using an evaluation form based on the SEGUE framework checklist.<sup>8</sup> SEGUE is an acronym that describes the five basic parts of necessary medical communication tasks, Set the stage, Elicit information, Give information, Understand the patient perspective, and End the interview. The evaluation form was modified to make it more relevant for a pre-anesthesia evaluation and the scoring expanded from a checklist to a Likert scale (see Appendix).

A total of seventeen different tasks were graded representing five different ICS areas. Those areas and their corresponding tasks were as follows; opening the interview (1,2), listening skills (3,4,5,6), interview content (7,8,9,10), therapeutic core qualities (11,12,13,14), and closing the interview (15,16,17). The evaluators graded each task strongly agree, agree, disagree, or strongly disagree. Each score was assigned a value of 4, 3, 2, and 1 respectively. The maximum score for each encounter was 68 and minimum score was 17 for a total cumulative score of 136 and cumulative minimum score of 34 for both the SP encounters. All participants received instruction on the use of the evaluation form prior to the commencement of the exercise.

The validity of the evaluation tool (expert opinion) was confirmed by faculty of the University of Toledo School of Communication who also participated in this study as members of the non-physician group. There was a high level of internal consistency (Cronbach's alpha 0.862) for all questions. The assessment tool included all the recommended behaviors necessary for creating a supportive positive climate as described by Burleson<sup>9</sup> and van Ryn.<sup>10</sup> The University of Toledo School of Communication experts were unable to find another assessment tool used in healthcare that was more effective.

The most recent resident In-Training Exam (ITE) scores were also recorded. We wanted to investigate if a resident's acquired knowledge as represented by the ITE influenced how each evaluator group scored the resident performance on the SP encounters.

Statistical analysis was performed using SPSS Statistics 17.0 (SPSS, Inc., Chicago, IL). Non-parametric tests were chosen because of the small sample size, the use of ordinal data and the assumption that the population was not normally distributed. Analysis of the differences in assessments of resident performance and the differences in assessments between the evaluator groups was performed using Kruskal-Wallis one-way analysis of ranks. Comparison of evaluator groups' ranking was done using Mann-Whitney tests. Spearman Correlation Coefficients were used to measure the degree of agreement between the evaluator groups as well as the relationship between resident performance and the ITE scores. A P value of <0.05 was considered significant.

## Results

Table 1 shows the mean scores, standard deviation and range of the composite scores from the two SP encounters for each resident sorted by evaluator group. Analysis of ranks showed a statistically significant difference in individual resident performance ( $p < 0.001$ ).

Table 2 shows the composite scoring, mean, standard deviation and confidence intervals of each resident ICS encounter by evaluator group. There was a statistically significant difference in how the different evaluator groups assessed the residents' performance ( $p < 0.001$ ). The residents' rank order for each evaluator group and mean rank is presented in Table 3. The highest scoring resident overall (resident 6) was ranked the highest across all evaluator groups. The lowest scoring resident (resident 5) was scored lowest or next to lowest by all evaluator groups. It appears that the residents were graded similarly, high or low, across all evaluator groups. This would suggest good correlation of assessment among the different groups of evaluators. We did see more variability in ranking among the mid-range residents.

Agreement of performance of individual residents among different evaluator groups was strong. There was significant correlation between the non-blinded teaching faculty and the two blinded anesthesiology faculty groups ( $r=0.764$ ,  $p=0.017$  and  $r=0.765$ ,  $p=0.016$  respectively). There was also significant correlation between the two blinded anesthesiology faculty assessments ( $r=0.946$ ,  $p=0.001$ ). There appeared to be some agreement between the non-blinded teaching faculty and the non-physicians but this did not approach statistical significance ( $r=0.477$ ,  $p=0.194$ ). The results of the correlation of scoring between evaluator groups are presented in Table 4.

Despite the agreement between evaluator groups there appeared to be a difference in scoring. The mean rank scores showed a statistically significant difference in how the different evaluator groups scored the resident ICS encounters ( $p<0.001$ ) presented in Table 5. The non-blinded anesthesiology faculty and the non-physician group scored the residents similar [119.42, 115.11, ( $p=0.824$ )]. The two blinded anesthesiology faculty groups scored the residents similar to each other [80.19, 82.00, ( $p=0.702$ )] but lower. In each of the individual ICS areas significant differences were noted except in the area of opening the interview. The difference in mean ranked scores in the area of therapeutic core qualities showed the most statistically significant difference ( $p<0.0001$ ) with the scoring as follows, non-blinded teaching faculty [119.45], non-physician group [127.56], and the blinded anesthesiology groups [84.21 and 69.87 respectively]. This is also presented in Table 5.

We did not see a correlation between the assessment of resident performance by the different evaluator groups and the resident ITE scores, Table 6. Those by the non-blinded teaching faculty showed a weak correlation with the ITE scores but this was not statistically significant ( $r=0.653$ ,  $p=0.057$ ). Interestingly the non-physician group showed absolutely no correlation with the resident ITE scores ( $r=0.034$ ,  $p=0.931$ ).

## Discussion

Despite the challenges facing educator evaluators we were encouraged to find that there was general agreement among the different categories of evaluators for each resident. A resident who was ranked high by the teaching faculty was similarly ranked high by the other evaluator groups. This also held true for the lowest ranked residents who were ranked low with most of the evaluator groups. Since it is one of the responsibilities of teaching faculty to evaluate this core competency, the correlation among the anesthesiology faculty groups was especially encouraging. Our results are consistent with the findings of Joshi<sup>11</sup> and Berger<sup>12</sup> who used multiple groups including faculty, nurses, allied health professionals and patients to evaluate the ICS of residents. Using a multi-rater survey they found general agreement among the different categories of evaluators.

Although there is strong agreement of resident ranking among the different evaluator groups, there still appeared to be a difference in how each group viewed the resident performances. The non-blinded teaching anesthesiology faculty and the non-physician group scored the resident ICS encounters similarly. The two blinded anesthesiology faculty groups also had similar, but much

lower encounter scoring. The difference in scoring between the evaluator groups, especially the anesthesiology faculty warrants further investigation with detailed ICS analysis.

The Kalamazoo II report<sup>13</sup> tells us that ICS is an integrated competence with two distinct parts. Interpersonal skills are considered relational and process oriented, i.e., the effect communication has on another person, and communication skills are the performance of specific tasks and behaviors (opening the interview and listening skills). The two skills are inherently related. Interpersonal skills build on basic communication skills. They have been described as the “humanistic qualities” or in our study the “therapeutic core qualities” we strive to create and sustain in a relationship. The end goal being to establish a sense of shared thoughts and feelings with the patient regarding their care. In fact this is one of the six identified goals reported by Kluft and Roizen<sup>14</sup> in a pre-anesthesia evaluation. Communication skills may not be accurately assessed if interpersonal skills are lacking in a patient-physician relationship. Interpersonal skills are more relation dependent than communication tasks and this may be what is influenced the most by previous resident interactions. The blinded anesthesiology faculty groups and non-physicians see only a brief video recorded snapshot of resident ICS and may be forced to focus on other factors such as knowledge, content or the communication tasks themselves.

In our study we used the ITE scores as a representation of resident knowledge. If knowledge were a factor we might have expected to see residents with higher ITE scores achieving higher scores by the evaluator groups on the ICS encounters. The anesthesiology faculty groups who are also required to assess the medical knowledge of residents, another core competency, may subconsciously be influenced by the content and medical knowledge exchanged during the encounter. However, we did not see a correlation between the ITE scores, absolute or scaled, and the assessments done by the anesthesiology faculty. The non-physicians, who were communication experts without a medical background, should not be influenced by resident knowledge and in fact this is what we saw. There was absolutely no correlation between resident ICS performance and their ITE scores, as scored by the non-physician group. It does not appear that resident knowledge and the medical information exchanged is responsible for the difference in scoring between the evaluator groups.

Therapeutic core qualities are the part of ICS that are reflected by our interpersonal skills. They are subjective and possibly most influenced by previous resident interactions. We did see a significant difference in this area between the various evaluator groups. The blinded anesthesiology faculty members' lack of prior direct resident interactions could explain why these groups scored the residents lower. Repeated resident interactions with the non-blinded teaching faculty seem to foster a positive relationship and may there be the reason for the much higher scores assigned by non-blinded faculty. The non-physician group who were non-medical but faculty in communication would be expected to respond to a more “pure” (i.e. not influenced by prior encounters) display of ICS in their assessments as they were possibly more attuned to how a patient may perceive these encounters. The fact that they also scored the residents higher and similar to the teaching faculty was encouraging. Familiarity with residents may be an advantage and allow non-blinded teaching faculty to assess the ICS of residents similar to the patient's perspective. This should be one of the major goals of the ICS assessment.

Limitations of this study include the small sample size of both residents and evaluator groups. Accordingly, any conclusions based on the information presented must be interpreted cautiously. More evaluators within each group would have increased the reliability of the ratings for each evaluator category. Though we were able to identify both high performing and poor performing residents, it is unclear what we can say about the residents in the middle where there seemed to be more variability in scoring. We also did not account for previous exposure with SPs so differences in experience may have led to a disparity in scoring by evaluators and evaluator groups. In addition, different anesthesiology residency programs may have different criteria for selection leading to different resident populations and expectations of performance by the teaching faculty. Further investigation could involve having residents from the blinded anesthesiology faculty programs participate in the SP encounters and exchange those video recordings between institutions for evaluation. This would increase the total resident numbers lending more validity to the study. It would also be interesting to see if the agreement of resident performance among the different evaluator groups and the scoring differences on the encounters would remain with a larger sample size.

In summary, data from this three-institution study indicate that previous interactions with residents do not adversely affect the teaching faculty's ability to evaluate their interpersonal and communication skills using an evaluation form based on the SEGUE method of assessing communication. .

## References

1. Tetzlaff, JE. Assessment of competency in anesthesiology. *Anesthesiology* 2007; 106:812-25
2. Association of American Medical Colleges: Medical School Objectives Project (1999) *Contemporary Issues in Medicine: Communication in Medicine*. AAMC, Washington, D.C.
3. Klass D, De Champlian A, Fletcher E, King A, Macmillan M. Development of a performance-based test of clinical skills for the United States Licensing Examination. *Federal Bulletin* 1998;85:177-85
4. Rider EA, Keefer CH. Communication skills competencies: definitions and a teaching toolbox. *Med Educ*. 2006;40:624-29
5. Casabianca AB, Papadimos TJ, Bhatt SB. The use of standardized patients to evaluate interpersonal and communication skills of anesthesiology residents: a pilot study *JEPM* 2008;10(2):1-22
6. Murphy KR, Balzer WK, Lockhart MC. Effects of previous performance on evaluations of present performance. *J Applied Psych*. 1985;70(1):72-84
7. Murphy KR, Balzer WK, Lockhart MC. Effects of previous performance on evaluations of present performance. *J Applied Psych*. 1985;70(1):72-84
8. Makoul G. The SEGUE Framework for teaching and assessing communication skills. *Pat Edu Couns*. 2001;45(1):23-34
9. Burleson BR, MacGeorge EL. Supportive Communication. In Knapp M, Daly J, eds. *Handbook of Interpersonal Communication*. Thousand Oaks, CA: Sage Publications; 2002:374-422
10. Van Ryn M, Heaney CA. Developing effective helping relationships in health education practice. *Health Educ Behavior* 1997;24(6):683-702
11. Joshi R, Ling FW, Jaeger J. Assessment of a 360-degree instrument to evaluate residents' competency in interpersonal and communication skills. *Acad Med*. 2004;79:458-463
12. Berger JS, Pan E, Thomas J. A randomized, controlled crossover study to discern the value of a 360-degree versus traditional, faculty-only evaluation for performance improvement of anesthesiology residents. *J Educ Periop Med*. 2010;12(1):1-13

13. Duffy DF, Gordon GH, Whelan G, Cole-Kelly K, Frankel R. Assessing competence in communication and interpersonal skills: The Kalamazoo II report. *Acad Med.* 2004;79(6):495-507
14. Klafta JM, Roizen MF. Current understanding of patients' attitudes toward and preparation for anesthesia: a review. *Anesth Analg.* 1996;83:1314-1321

**Appendix**

**Anesthesiology ICS Assessment**

Resident:

Date:

	When conducting the preoperative evaluation, the resident	Strongly Agree	Agree	Disagree	Strongly Disagree
1.	Introduced self in a respectful manner and used proper name of patient	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	Verified purpose of visit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	Seated self in an appropriate manner and distance in relation to the patient	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.	Maintained appropriate eye contact	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.	Did not interrupt unnecessarily	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.	Appeared attentive and interested	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.	Used open-ended questions followed by closed-ended questions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8.	Used vocabulary consistent with patient background, avoided jargon	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.	Obtained information in a systematic, orderly process	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.	Was non-judgmental	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.	Provided reassurance and guidance if necessary	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12.	Showed a courteous attitude toward the patient	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13.	Showed a compassionate attitude toward the patient	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14.	Explored patient's concerns or perspectives regarding the problem	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15.	Asked if the patient had questions or anything to add at end of interview	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.	Summarized pertinent information to clarify for patient and interviewer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17.	Informed the patient the interview had concluded and what would happen next	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Evaluator _____ Total Score _____				

Opening the Interview (1,2)  
Listening Skills (3,4,5,6)  
Interview Content (7,8,9,10)  
Therapeutic Core Qualities (11,12,13,14)  
Closing the Interview (15,16,17)