Original Investigation

Manual Dexterity Aptitude Testing
A Soap Carving Study

Christopher G. Tang, MD; Raymond L. Hilsinger Jr, MD; Raul M. Cruz, MD; Luke J. Schloegel, MD; Fred M. Byl Jr, MD; Barry M. Rasgon, MD

IMPORTANCE Currently there are few validated metrics for predicting surgical skill among otolaryngology residency applicants.

OBJECTIVE To determine whether manual dexterity aptitude testing in the form of soap carving during otolaryngology residency interviews at Kaiser Permanente Medical Center Oakland predicts surgical skill at the time of graduation from otolaryngology residency programs.

DESIGN, SETTING, AND PARTICIPANTS This study was conducted to determine how applicants with the best and worst soap carvings compared at the time of graduation with respect to various metrics including visuospatial ability and manual dexterity. Over the last 25 years, applicants to the residency program at Kaiser Permanente Oakland were required to carve soap during their residency interview. The 3 best and 3 worst soap carvings from 1990 through 2006 were determined. Of the individuals who carved those soaps, 62 qualified for the study and matriculated into otolaryngology residency programs.

INTERVENTIONS Surveys were sent to the 62 individuals’ residency programs to evaluate those individuals on a 5-point Likert scale in various categories as well as to rank those individuals as being in the top 50% or bottom 50% of their graduating class.

MAIN OUTCOMES AND MEASURES All else being equal, we hypothesized that applicants who had the manual dexterity and visuospatial skills to accurately carve a bar of soap would more likely possess the skills necessary to become a good surgeon.

RESULTS There was no difference between individuals with the best soap carvings and those with the worst soap carvings in all categories: cognitive knowledge, visuospatial ability, manual dexterity, decision making, and overall score (P > .10 for all categories). There was a 95% response rate, with 35 of 37 residency programs responding and 59 of 62 surveys returned.

CONCLUSIONS AND RELEVANCE Manual dexterity aptitude testing in the form of soap carving does not appear to correlate with surgical skill at the time of graduation. Further studies need to be conducted to determine the role of manual dexterity and visuospatial aptitude testing in the otolaryngology application process.
Otolaryngologists, as astute clinicians, need to possess both the mental capabilities and intellect to make sound clinical judgments. As surgeons, manual dexterity and steadiness are paramount in carrying out intricate surgical procedures such as microvascular and otologic surgery. Current metrics during the residency interview process such as US Medical Licensing Examination (USMLE) scores and academic performance in medical school can help identify applicants that excel in basic science and clinical medical knowledge. Other metrics such as research projects, letters of recommendation, and performance during the interview can determine various other aspects of the individual applicant. However, none of these metrics can determine an applicant’s manual dexterity or visuospatial perception.

The importance of assessing surgical aptitude has long been underscored in the otolaryngology community. However, there is still much controversy and disagreement about the inclusion and validity of individual manual dexterity tests in the residency selection process. Despite the controversial nature of manual dexterity aptitude tests, several otolaryngology residency programs, including the residency program at Kaiser Permanente Medical Center Oakland and at Mayo Clinic School of Medicine, continue to use tests of manual dexterity in their applicant selection process.

Several reasons cited by Graham et al as limitations of manual dexterity tests include the paucity of well-designed studies, the use of subjective ratings of surgical skill, and a poor definition of surgical skill. Other studies in the ophthalmology literature have suggested that the inclusion of tests with spatial aptitude may be more helpful than manual dexterity alone. To address several of these issues, we conducted a double-blinded retrospective study looking at the manual dexterity and visuospatial aptitude test given to residency applicants to the Kaiser Permanente Medical Center Oakland Head and Neck Surgery residency program: the soap carving.

**Methods**

An institutional review board exemption was obtained from the Kaiser Foundation Research Institute. In the last 25 years, every resident applicant to the Kaiser Permanente Medical Center Oakland Head and Neck Surgery residency program was required to carve soap at the time of their residency interview. Instructions were given to carve the most proportional, smooth, and accurate representation of the form described in Figure 1. Carvings were evaluated based on adherence to the drawing, distortion, edge definition, and finish. All else being

---

**Figure 1. Soap Carving Instructions**

<table>
<thead>
<tr>
<th>Supplies</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>• One block of Ivory soap — to be cut in half. The soap should be soft, not brittle. If it is, ask for a new bar.</td>
<td>• Split the bar of soap in half.</td>
</tr>
<tr>
<td>• One #15 blade with handle.</td>
<td>• Carve the above figure to approximate scale.</td>
</tr>
<tr>
<td>• One #11 blade with handle.</td>
<td>• Use the second half if needed.</td>
</tr>
<tr>
<td>• One bath towel.</td>
<td>• The towel is to protect your clothing.</td>
</tr>
<tr>
<td></td>
<td>• Clean up your mess. Wrap the chips in the towel and give it to the resident.</td>
</tr>
<tr>
<td></td>
<td>• You should spend about an hour with the carving, not including time out for interview.</td>
</tr>
</tbody>
</table>

Carvings are evaluated for adherence to the drawing, distortion, edge definition, and finish. When completed, please sign or initial bottom of carving and give carving to resident.
equal, it was our assumption that applicants who had the manual dexterity and visuospatial skills to carve an accurate soap would more likely possess the skills necessary to become a good surgeon. To test that hypothesis, we looked at the individuals with the 3 best and the 3 worst soap carvings each year among all the applicants between 1990 and 2006—a total of 16 years (Figure 2). Several individuals, who were blinded to the name of the applicant, selected the best 3 and worst 3 soap carvings of that year in an objective manner. Criteria for a good soap included the following: maintaining proportionality in all dimensions, smooth edges and proper contouring, and being representative of the figure in the drawing.

A total of 96 soap carvings were selected by this method. Of the 96 soap carvings, the individuals who matched into a head and neck residency program were determined, and their residency directors and department chiefs were contacted with a survey to evaluate the individual on a Likert scale of 1 to 5 in 4 categories (cognitive knowledge, visuospatial ability, manual dexterity, decision making/poise), as well as to select whether the individual was in the top 50% of all residents who gradu-

Figure 2. Good and Bad Soap Carvings

Examples of best and worst soap carvings were selected in a blinded fashion from the set of all applicant soap carvings for that year. Each year there were roughly 40 applicants and hence 40 soap carvings to select from. Soap carvings from 1990 through 2006 were examined.

Figure 3. Surgical Skills Survey

Surgical Skills Evaluation Form

<table>
<thead>
<tr>
<th>Identification Number:</th>
<th>Evaluator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Please circle a number 1 through 5 that most represents the head and neck surgeon at the time of graduation:</td>
<td></td>
</tr>
<tr>
<td>Cognitive Knowledge</td>
<td>At times, demonstrated gaps in surgical anatomical knowledge base and preparedness. Below average</td>
</tr>
<tr>
<td>3</td>
<td>Exceptional knowledge base. Very thorough surgical anatomical understanding and preparedness.</td>
</tr>
<tr>
<td>1</td>
<td>Spatial ability as expected for graduating chief resident</td>
</tr>
<tr>
<td>3</td>
<td>Manual dexterity as expected for graduating chief resident</td>
</tr>
<tr>
<td>3</td>
<td>Manual dexterity as expected for graduating chief resident</td>
</tr>
<tr>
<td>3</td>
<td>Decision Making/POise</td>
</tr>
<tr>
<td>3</td>
<td>For the most part, made go-go operative decisions but was considered below average. At times lacked poise necessary</td>
</tr>
<tr>
<td>Bottom 50%</td>
<td>Top 50%</td>
</tr>
<tr>
<td>2) Was this individual at the top 50% of all residents that trained at your program or the bottom 50% of all residents that trained at your program? (circle)</td>
<td></td>
</tr>
</tbody>
</table>

Please mail this form back in the self-addressed stamped envelope included with this survey. All results are confidential and this study is double blinded. The results of this survey are to be used solely for research purposes.

The 3 best and the 3 worst soap carvings among all the residency applicants each year since 1990 were determined. Surveys were submitted to the residencies from which these applicants graduated, specifically to the chief of their department and the residency director. OR indicates operating room.
ated from their program or the bottom 50% of their class (Figure 3). The residency directors and department chiefs were blinded to the soap carvings and did not know if the individual had a “best” soap or a “worst” soap carving. All results were held in strict confidentiality, and each candidate was assigned a unique identification number to make the study double blinded.

Results

Sixteen years’ worth of soap carvings were evaluated, and the best 3 and worst 3 soap carvings of each year were obtained for a total of 96. Of the 96 applicants, 30 ended up going into a specialty other than otolaryngology. The breakdown included the following: 7 went into general surgery, of which 2 individuals did a plastics fellowship and 1 did a thoracic fellowship; 5 went into radiology, of which 1 did a fellowship in neuroradiology; 4, anesthesia; 2, cardiology; 3, emergency medicine; 3, family practice; 2, gastroenterology; 1, preventative medicine/legal medicine; 1, dermatology; 1, nephrology; and 1, hematology/oncology. Two individuals quit medicine altogether. There was no difference in the proportion of good to bad soap carvings in those who matched into otolaryngology and those who did not (P > .05). The remaining 64 individuals all became otolaryngologists. Of these individuals, 2 went to programs that no longer exist (Charles R. Drew University at Martin Luther King Jr Hospital in Los Angeles, California) (Figure 4). The remaining 62 individuals went to 37 different residency programs, of which surveys were sent to both the residency director and chief of the department. Of the 37 residencies contacted, 35 replied, of which 1 decided not to participate because they deemed the study to contain sensitive information; 1 program director/chief only knew 1 of the 2 individuals because the other individual graduated before his time; and 1 program director chief did not know the individual because it was before his time. Two residency programs did not respond. Our survey response rate was 95% (35 of 37 programs).

Of the 59 applicant surveys returned, 53 were evaluated, 4 were not evaluated because the institution did not want to participate in this study, and 2 were not known by the residency director/chief. The response rate was 95% (59 of 62 surveys received), and the evaluation rate was 90% (53 of 59 surveys received).

Of the 53 surveys that were evaluated, 38 individuals were ranked at the top 50% of all residents, and 15 were ranked at the bottom 50% of all residents. There appeared to be a recall bias in that most program directors and chiefs remembered their past residents in a positive light. Because both variables were dichotomous, a χ² test was performed, yielding a P value of .41, which suggests there is no relationship between quality of soap carving and class ranking (Figure 5). Therefore applicants with a good soap carving were not more likely to be in the top 50% of their class surgically compared with those with bad soap carvings.

The mean scores for cognitive knowledge, visuospatial ability, manual dexterity, and decision making were 4.09, 4.00, 3.88, and 4.16, respectively, for individuals with good soap carvings. For individuals who had bad soap carvings, the mean scores were 3.84, 3.73, 3.57, and 3.61, respectively (Figure 6). A t test was performed, yielding high P values suggesting that there is no relationship between quality of soap carving and resident qualities. In other words, applicants with good soap carvings did not have better cognitive knowledge, visuospatial ability, manual dexterity, or decision making compared with
applicants with bad soap carvings. There was a trend toward significance with respect to visuospatial ability \((P = .10)\), but significance was not achieved, possibly because of the small sample size of applicants who were in the “bottom 50%.”

**Discussion**

Since the early 1970s, dental programs have evaluated applicants’ manual dexterity through wax and chalk carvings.\(^5\)\(^6\) Although a study had shown a correlation between manual dexterity and both USMLE scores and class rank, many surgeons would still like some form of manual dexterity testing that would consistently predict surgical skill.\(^7\) Several efforts have been undertaken to develop such an aptitude test.

Carlson et al\(^3\) published a study in 2010 describing the manual aptitude testing performed at Mayo Clinic in Rochester, Minnesota, for otolaryngology residency applicants. At the Mayo Clinic otolaryngology training program, after orientation, applicants were seated at a microsurgical training station and allotted 20 minutes to suture an incision using 10-0 nylon suture on a latex practice card. Their performance was graded using a 1 to 5 scoring system for the following categories: microscope use, respect for tissue, instrument handling, knot tying and suture control, skills acquisition, and attitude toward the exercise. In their study, they did not correlate it with surgical skill at the time of graduation. However, they found value in that it allowed them to have a direct assessment tool that may prove useful in identifying outliers, both high and low, to aid in final applicant ranking.\(^3\)

Various metrics during the residency interview process such as USMLE scores and academic performance in medical school can help identify applicants that excel in various facets of residency. For example, an applicant’s research profile can reveal to the program whether they want to pursue an academic path. Letters of recommendation may provide a glimpse into what kind of intern and resident they are likely to be. The interview process reveals an applicant’s personality, values, and critical thinking ability. However, none of these metrics can determine an applicant’s manual dexterity or visuospatial perception.

The soap carving was set up by our program over 25 years ago in an attempt to determine an applicant’s manual dexterity and visual spatial perception. Most of the 40 soap bars carved each year appeared similar in quality, but each year there would be 2 or 3 soap bars that were exquisite carved and 2 or 3 that were very poorly carved. There were even a few instances when applicants would cut themselves while carving.

A blinded retrospective review was performed by Schloegel et al\(^8\) in 2012 at Kaiser Permanente Oakland look-
ing at the soap carvings of the residents who had graduated from the Kaiser Permanente Medical Center Oakland Head and Neck Surgery residency program. In their study, no correlation was found between the quality of the soap carved as an applicant and the surgical skill of the resident at the time of graduation. The authors believed that this study did not show the validity of soap carving as an evaluation tool because all of the soap carvings were of high quality and there was not enough variability. The authors hypothesized that the true predictive value of soap carving as a test of manual dexterity rested in identifying outliers. Those applicants with either the highest-quality or the lowest-quality soap carving would correlate with ultimate surgical skill or lack thereof.

The present study was designed to identify the outliers and then follow up what happened to these applicants who did not match into the Kaiser Permanente residency program. Therefore, all 800 soap carvings that had been kept for over 16 years were reviewed, and the 3 best and the 3 worst soap carvings among all the residency applicants each year from 1990 through 2006 were obtained. More recent soap carvings were not evaluated because those individuals would still be in residency. Surveys were submitted to the residencies from which these applicants graduated. Several individuals, who were blinded to the name of the applicant, selected the best 3 and worst 3 soap carvings of that year in an objective manner by determining which soaps were most proportional, smooth, and representative of the figure in the instructions and those that were the least out of all the soap carvings that year. A total of 96 soap carvings were selected by this method (48 best and 48 worst soap carvings of all residency applicants).

The goal of the present study was to determine the head and neck surgeon's surgical skill at the time of graduation from their residency training program. Although the initial hypothesis was that the soap carving test would be a predictor of surgical skill, the actual results showed no correlation between the quality of soap bars carved and the surgical skill of the individual. In addition, there was no correlation between the quality of soap bars carved and whether that applicant matched into a head and neck residency.

A study by Wanzel et al.9 examined the influence of visuospatial ability and manual dexterity on surgical performance across 3 levels of expertise: dental students, surgical residents, and staff surgeons. Wanzel et al.9 concluded that among novices, visuospatial ability was associated with improved performance on spatially complex surgical procedures, but not in advanced trainees and experts, suggesting that practice and surgical experience may supplant the influence of visuospatial ability over time. Similarly to Wanzel et al.,9 it is possible that all applicants advanced equally throughout their residency training regardless of the quality of their soap carvings at the time of residency interviews. Moreover, these data may suggest that practice and experience during residency training can compensate for any initial discrepancies in visuospatial ability and manual dexterity.

On the other hand, in the soap carving study there was a trend toward significance with respect to visuospatial ability (P = .10), but significance was not achieved, possibly because of the small sample size of applicants who were in the “bottom 50%.” There is a possibility that innate manual dexterity can affect an individual's surgical ability. There may be individuals who have tremors or other physical manifestations that make operating difficult but can still be safe surgeons. One comment returned along with the survey stated that “This resident was exceptional regarding knowledge and decision making. Secondary to a tremor, he had difficulty under the microscope. Very safe surgeon.”

There are 3 possible areas of bias in the study. First, although the soap carvings were judged in a blinded fashion for maintaining proportionality in all dimensions, smooth edges and proper contouring, and being representative of the figure in the instruction, the judges are separate individuals. Certain individuals may have a different concept of what is a better soap carving compared with other individuals, even based on these strict criteria. To mitigate this type of bias, we only looked at the top 3 and worst 3 of more than 40 soap carvings each year. All judges unanimously agreed on which soaps were the worst and best 3 of that year.

A second area of bias may be attributed to a positive recall bias from the residency directors and department chiefs. Of the 53 surveys that were evaluated, 38 individuals were ranked at the top 50% of all residents and 15 were ranked at the bottom 50% of all residents. This positive recall bias may suggest that residency applicants can become good surgeons despite having a poor soap carving and possibly that manual dexterity and visuospatial aptitude can be taught and learned during residency, similar to what was suggested by Wanzel et al.9

Finally, a third area of bias may simply be that residency directors and staff surgeons have a hard time recalling how their residents were as surgeons 2 decades ago. The study was performed only up to 2006 because those applicants who were interviewed in 2007 were still in residency in 2011, when the surveys were mailed out. One comment from a returned survey stated “It is difficult to accurately evaluate residents 6 years after they completed training.” Moreover, at one of the institutions, the program director/chief of the department only knew 1 of the 2 surgeons listed in the survey because the other one attended the residency program before his time.

Conclusions
At the Kaiser Permanente Medical Center Oakland, soap carving at the time of residency interviews does not correlate with surgical skill at the time of graduation. As Krespi et al. suggested roughly 30 years ago, the search for a convenient, rapid, manual dexterity examination that could be used at the time of residency applicant's interviews is still elusive. Although our screening test showed that there was no correlation between soap carving quality and surgical skill at the time of graduation, there is still a large need for the development of such a tool. Further studies need to be performed to validate such tests.
Conflict of Interest Disclosures: None reported.

Previous Presentation: This study was presented as a poster presentation at the 2012 Annual American Academy of Otolaryngology-Head and Neck Surgery Meeting and Oto Expo; September 9-12, 2013; Washington DC.

Additional Contributions: Mary H. Tang, BA, assisted with data collection.

REFERENCES


