Assessment of Moral Reasoning Skills in the Orthopaedic Surgery Resident Applicant

Kyle C. Bohm, MD, Tyler Van Heest, BA, Terence J. Gioe, MD, Julie Agel, MA, Todd C. Johnson, MD, and Ann Van Heest, MD

Investigation performed at the Department of Orthopaedic Surgery, University of Minnesota, Minneapolis, Minnesota

Background: Assessing orthopaedic surgery residency applicants is a complex process. One important applicant characteristic not commonly measured during the interview process is moral behavior, a key component of the guidelines of the Accreditation Council for Graduate Medical Education for professionalism competency. We sought to determine whether a relationship exists between the results of residency selection interviews, as measured by match rank order, and moral reasoning skills, as measured by the revised version of the Defining Issues Test (DIT-2).

Methods: The DIT-2, a psychological tool that has been validated in assessing moral reasoning in >30,000 subjects, was administered to orthopaedic surgery residency candidates on a voluntary basis during their interview day. Data were collected annually from four applicant classes (2009 to 2012) at one institution. Each candidate was interviewed by an orthopaedic faculty team and was ranked comparatively with the other interviewing applicants. The five-part aggregate interview method for ranking applicants for the match list included a focused interview station presenting ethical dilemmas as one part of determining rank order.

Results: One hundred and twenty-four fourth-year medical students interviewing for orthopaedic surgery residency were assessed with the DIT-2. The results from the DIT-2 produced a postconventional score, which correlates with various levels of moral development. Each applicant’s postconventional score was then compared with the patient’s overall rank order. Applicants had highly variable moral reasoning skills, with no relationship to the final rank list.

Conclusions: Despite a wide range of DIT-2 scores, the interview process did not differentiate applicants based on moral reasoning skills. It remains unclear whether the DIT-2 could (or should) be used as an additional data point to help stratify orthopaedic residency applicants and to predict their performance.

Peer Review: This article was reviewed by the Editor-in-Chief and one Deputy Editor, and it underwent blinded review by two or more outside experts. It was also reviewed by an expert in methodology and statistics. The Deputy Editor reviewed each revision of the article, and it underwent a final review by the Editor-in-Chief prior to publication. Final corrections and clarifications occurred during one or more exchanges between the author(s) and copyeditors.

Disclosure: None of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of any aspect of this work. None of the authors, or their institution(s), have had any financial relationship, in the thirty-six months prior to submission of this work, with any entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. Also, no author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete Disclosures of Potential Conflicts of Interest submitted by authors are always provided with the online version of the article.
applicants in the areas of intelligence, technical skills, and interpersonal skills. However, this goal remains elusive, as many common objective measures of medical student performance have low predictive value for future resident success.

Residency programs currently grant interviews utilizing limited data points, including United States Medical Licensing Examination (USMLE) scores, medical school transcripts, and letters of recommendation. Although these measures may help quantify an applicant’s intellectual acumen and academic accomplishment, they fail to assess other important applicant characteristics. One important applicant characteristic neglected by these measures is moral behavior, a key component of the guidelines of the Accreditation Council for Graduate Medical Education (ACGME) for professionalism competency. Currently, moral behavior and professionalism are generally assessed through the interview process or other aspects of the application. As the presumed goal for every orthopaedic program is to train clinically competent orthopaedists with a high degree of professionalism, identifying candidates most likely to reach this goal using a quantifiable test of moral reasoning seems attractive. Poor professional behavior has been cited as the most common reason for resident dismissal from a program, suggesting that more attention to professionalism and moral behavior may be warranted. 

Rest et al. formulated a comprehensive theory of moral behavior based on four psychological capacities: moral sensitivity, moral reasoning, moral motivation, and moral implementation. Moral reasoning skills are required when physicians consider conflicting personal opinions, government policies, resource allocation concerns, and the wishes of their patients and patient families. Although moral reasoning is only one component of moral behavior, it is an important and quantifiable measure. Rest et al. classified moral reasoning into three overlapping schemas: the personal interest schema, the maintaining norms schema, and the postconventional schema. These schemas represent moral reasoning strategies, and each individual uses these strategies in varying degrees when reasoning through moral dilemmas. The revised version of the Defining Issues Test (DIT-2) is a reliable and validated instrument using story-based moral dilemmas (Fig. 1) developed by Rest et al. to quantitatively assess the degree to which an individual uses each moral schema. The results from the DIT-2 produce a score that correlates with various levels of moral development.

Personal interest strategies focus on what each stakeholder has to gain or lose in the moral dilemma without worrying about broader societal implications. Individuals who score high in this schema justify their decisions by focusing on the personal consequences of their actions.

Maintaining norms strategies focus on upholding social order and respecting established hierarchical role structures. Individuals who score high on this schema focus on following rules, laws, procedures, and guidelines, as they are deemed necessary for maintaining structure and curbing self-interested behavior.

Postconventional strategies focus on the concept that moral obligations are based on shared moral ideals. Importantly, these moral ideals are considered open to scrutiny. Individuals with high postconventional scores justify their actions by appealing to moral ideals that are shareable, reciprocal in nature, and logically coherent. Rest et al. defined any individual postconventional score of >50 points (range, 0 to 95 points) as thinking primarily at the principled level of moral reasoning.

According to Rest et al., as individuals develop increasingly sophisticated reasoning strategies, there will be an anticipated decrease in the use of the personal interest and maintaining norms schemas and an increase in the use of the postconventional schema. More sophisticated moral reasoning skills (as evidenced by a postconventional score of >50 points) are associated with more complex decision-making processes and less paternalistic attitudes among residents, resulting in a more sophisticated attitude toward patient care. Such skills correlate with clinical performance in numerous medical studies, as well as in dentistry, nursing, and law. Low moral reasoning scores preclude high clinical scores, and residents with the highest levels of moral reasoning are seldom judged as poor clinical performers. Low postconventional scores as measured by the DIT-2 correlate with increased malpractice claims among practicing orthopaedic surgeons. Orthopaedic surgeons with postconventional scores on the DIT-2 of <50 points had claims rates of >0.40 claim per year compared with <0.20 claim per year for practicing orthopaedic surgeons with postconventional scores of >50 points.

The purpose of this study was to assess whether a relationship exists between the rank order of residency applicants after a highly structured aggregate interview residency selection process and moral reasoning skills, as measured by the DIT-2. Furthermore, we assess the relationship between USMLE scores and moral reasoning skills. We hypothesized that moral reasoning as measured by the DIT-2 would show a direct relationship with our rank list.

Materials and Methods
The participating institution granted institutional review board approval for this study. Data were collected from 2009 to 2012 from four separate orthopaedic residency applicant classes interviewing at one training institution. Residency applicants responded to the DIT-2, a validated measure of moral reasoning, on a voluntary basis and to the institution’s standard ranking process. The study was explained to the residency applicants, including the fact that participation and results were not used for ranking; written consent from participants was obtained; and time was granted as part of the interview day to complete the DIT-2. Of the 208 residency applicants interviewed, 124 fourth-year residents completed the DIT-2 for a total response rate of 60%; of those, four tests were purged, leaving 120 tests for analysis. DIT-2 scores were not incorporated into the ranking process. USMLE scores were acquired from the residency application.

The DIT-2
The DIT-2 is a reproducible, quantitative, multiple-choice test used to assess moral reasoning. Participants are given five dilemmas to evaluate and to decide on the best courses of action. Each dilemma is followed by a set of questions that enable the participants to reveal the logic behind their moral reasoning. The answer set also includes meaningless questions, which allows the purging of scores if the examinee answers in a way that indicates that their responses are random. On the basis of these responses, participants receive scores...
for each of the three moral schemas: personal interest, maintaining norms, and postconventional. Higher scores indicate a greater proportionate importance given to that schema when making moral judgments. It is important to distinguish that this test does not assess moral values, but rather how one reasons through moral dilemmas. The test takes approximately twenty to thirty minutes to complete and has been validated by more than 44,000 subjects.

Attempting to beat the test by faking to gain a higher postconventional score is not successful. Rest et al. defined any individual postconventional score of >50 points (range, 0 to 95 points) as thinking primarily at the principled level of moral reasoning.

Applicant Ranking Process
The ranking process used in this training program has been consistently used for the last ten years, has predicted future resident performance, and has been shared with the wider orthopaedic community via a peer-reviewed publication.

All of the applicants each year are initially screened according to USMLE scores, academic performance, research experience, volunteer experience, background diversity, letters of recommendation, and orthopaedic sub-internship scores. Fifty-two applicants are selected to be interviewed. Each interviewing applicant rotates through five dedicated interview stations with two faculty members at each station. Each station is assigned to assess one of five aspects of the applicant’s background: (1) motivation and work ethic, (2) patient care and motor skills, (3) career goals, (4) ethical behavior and professionalism, and (5) academic performance. The ethical behavior and professionalism station currently includes two faculty members questioning applicants about scenarios involving moral dilemmas or challenging situations. Following the interview period, there is a group faculty discussion regarding each residency applicant. Each interviewer then submits a ranked list of the applicants. The lists are aggregated to formulate the final ranked list as utilized in this study.

As part of this study, the faculty members in the ethical behavior and professionalism station separately scored applicants on a scale of 1 to 5 points, in which 1 point denotes worst and 5 points denote best, regarding their assessment of the applicant’s response to moral dilemmas. However, all applicants received scores of 4 or 5 points; because there was little to no discernment between applicants, this information was not used for further analysis.
Statistical Methods
DIT-2 scores, USMLE scores, and applicant ranks were available for 120 residency applications. Applicants were stratified into groups of high rank (1 to 12 points), mid-rank (13 to 36 points), and low rank (≥37 points) for analysis. Stratification was based on our program’s prior match experience, in that high-ranked applicants often match at our institution, mid-ranked applicants may rank at our institution or other institutions, and low-ranked applicants often do not match. One-way analysis of variance (ANOVA) tests with a Bonferroni correction (p = 0.05) for post hoc analysis was used to assess the relationships between rank groups and DIT-2 scores. The Pearson product-moment correlation coefficients were computed to assess the relationship between USMLE scores and DIT-2 scores.

Results
The overall mean postconventional score (and standard deviation) was 51.9 ± 15.09 points (range, 0 to 86 points). The mean postconventional score (and standard deviation) was 53.3 ± 15.32 points for high-ranked applicants, 50.9 ± 14.58 points for mid-ranked applicants, and 52.3 ± 16.08 points for low-ranked applicants. There were no significant differences among the three groups (p = 0.776) (Fig. 2). Of the 120 medical students in this study, twenty-eight had a postconventional score of <40 points (Fig. 2). The mean maintaining norms score (and standard deviation) was 14.1 ± 9.36 points (range, 0 to 50 points). The mean maintaining norms score (and standard deviation) was 29.0 ± 14.07 points for high-ranked applicants, 30.0 ± 13.70 points for mid-ranked applicants, and 34.0 ± 15.84 points for low-ranked applicants. There were no significant differences among the three groups (p = 0.327) (Fig. 2). The overall mean personal interest score (and standard deviation) was 14.1 ± 9.36 points (range, 0 to 50 points). The mean personal interest score (and standard deviation) was 14.1 ± 10.75 points for high-ranked applicants, 16.2 ± 9.45 points for mid-ranked applicants, and 10.6 ± 6.67 points for low-ranked applicants. The differences among groups were significant (p = 0.022) (Fig. 2). Post hoc analysis was conducted and found a significant difference between low-ranked and mid-ranked applicants (p = 0.018). There were no significant differences between high-ranked and mid-ranked applicants (p = 0.925) or between high-ranked and low-ranked applicants (p = 0.394).

Discussion
In this study, the interview process, as reflected in the rank order of applicants, did not correlate with the results of a widely used, validated test of moral reasoning. There are several possible explanations for this finding. The DIT-2 scores gathered in this study were not used in any fashion during the application process and were, in fact, unknown at the time of the applicant’s residency match. The interviewers cannot adequately assess an applicant’s moral reasoning skills during the brief interview period or through reviewing other aspects of the application. The lack of relationship suggests that the interview technique of posing ethical dilemmas to the applicant may not be as useful as previously thought.

Another explanation for this finding could be that applicants with lower moral reasoning scores have desirable achievements in other areas of the application that overshadow their low moral reasoning scores. This possibility is supported by our findings that USMLE scores do not correlate with moral reasoning scores. Clearly, applicants with low moral reasoning scores can perform very well on objective medical testing. This, or other scholastic achievements, might make the applicant desirable regardless of his or her DIT-2 score. However, given some prior evidence of a relationship between higher moral reasoning scores and resident clinical performance, as well as decreased malpractice claims among attending physicians, orthopaedic surgery programs may benefit from placing more emphasis on moral reasoning.

The lack of a correlation between USMLE scores and DIT-2 scores also highlights the fact that those who score the best on objective medical testing do not necessarily possess the...
highest moral reasoning abilities and vice versa. Moral reasoning reflects a different ability than standardized medical test-taking. The DIT-2 may allow programs to evaluate applicants via a distinctly separate data point, one that has not been widely utilized by orthopaedic residency programs.

There is a growing body of evidence suggesting that moral reasoning skills do not progress throughout medical education and may actually decline or, at best, plateau. Moral reasoning scores do not improve during medical training, the individual with low moral reasoning skills could be identified early with use of the DIT-2 so that intervention could be initiated.

Moral reasoning skills among applicants were quite variable, a surprising result given the highly educated nature of our cohort and the similarity of their training. Postconventional scores ranged from the level of a junior high student to a moral philosopher. It may be important to identify those applicants with unusually low moral reasoning scores even if they are otherwise attractive candidates. In our cohort, applicants ranked in the top ten had a postconventional score ranging from 0 to 76 points. Furthermore, of the 120 medical students in this study, 23% had a postconventional score of <40 points. Because a previous study has shown an increased rate of malpractice claims in practicing orthopaedic surgeons with a postconventional score of <40 points, these low moral reasoning scores serve as an alert for further inquiry or remediation.

There were obvious limitations to this study. It has yet to be determined if higher moral reasoning skills, based on the DIT-2 score, correlate directly with orthopaedic residency performance. Such a correlation would allow an additional data point for resident selection and, equally importantly, would identify at-risk applicants with low moral reasoning skills. Additionally, only 124 of 208 medical student applicants elected to complete the DIT-2, which may introduce a selection bias. However, a response rate of 60% is considered high for survey-type research.

It is important to realize that these relationships are susceptible to false positives and false negatives. It is possible that residency applicants and surgeons with low postconventional scores may not make moral reasoning errors that impact patient care, nor is a high postconventional score a guarantee of faultless moral reasoning. For this reason, further research and discussion into the sensitivity and specificity of the DIT-2 as a predictor of clinical performance are warranted. Lastly, in our statistical analysis of the DIT-2 results, it is acknowledged that our sample size may be insufficient to detect differences among rank groups.

The strengths of this study included the use of a psychological instrument (DIT-2) that has been validated through a large sample size (>30,000 subjects). Not only does the DIT-2 have adequate internal consistency and high test-retest reliability, but it also has been used extensively in higher education, including medicine, veterinary medicine, dentistry, law, and nursing. Also, to our knowledge, our sample population is the largest studied to date in the orthopaedic literature.

In conclusion, the highly competitive nature of the orthopaedic residency selection process requires effective stratification of applicants. Many programs wish to assess moral behavior and professionalism through the interview process. Our study has shown that there is a wide range in moral reasoning skills among orthopaedic residency applicants according to the DIT-2 and that moral reasoning skills are not correlated with either the standardized applicant ranking system at our institution or USMLE scores. More research and discussion in the orthopaedic community are required to determine the value of the DIT-2 as a potential tool for further stratifying applicants on the basis of moral reasoning.

References