A Model for the Assessment of Medical Students’ Competency in Medical Ethics

Amanda Favia, The Graduate Center, CUNY–Philosophy
Lily Frank, The Graduate Center, CUNY–Philosophy
Nada Gligorov, Mount Sinai School of Medicine
Steven Birnbaum, The Graduate Center, CUNY–Philosophy
Paul Cummins, The Graduate Center, CUNY–Philosophy
Robert Fallar, Mount Sinai School of Medicine
Kyle Ferguson, The Graduate Center, CUNY–Philosophy
Katherine Mendis, The Graduate Center, CUNY–Philosophy
Erica Friedman, Mount Sinai School of Medicine
Rosamond Rhodes, Mount Sinai School of Medicine

Background: This article focuses on the goals of our medical ethics education program and our formative assessments of students’ competency at various points during this education. Methods: Because of the critical relationship between a program’s goals and the design of an assessment strategy, we provide an overview of the theoretical basis of our curriculum, our program’s objectives, and teaching methods. In order to verify that our students had achieved minimum competency in the objectives of our ethics curriculum, we developed assessments that evaluated their ability to identify and apply ethical principles to clinical cases and to use moral reasoning to resolve dilemmas. We verified the reliability of these assessment instruments by correlating two different Mount Sinai raters’ scores of the same assessments with each other and the validity of these assessments with external reviewers. Results: For interrater reliability, paired raters scored the same student written exercise within 5 points of each other on 119 of the exercises (87% rater consensus). Therefore, we found our assessment tools to be reliable. Regarding validity, all three expert external reviewers agreed that our instruments were well suited for evaluating medical student competency in medical ethics and that they measured what we intended to measure. Conclusions: Our efforts in medical ethics education and competency assessment have produced an integrated model of goals, methodology, curriculum, and competency assessment. The entire model is directed at providing students with the ethical knowledge, skills, and attitudes required of an exemplary physician. We have developed reliable and valid assessment tools that allow us to evaluate the competency of students in medical ethics and to identify students who require remediation, and that are useful for other ethics programs.

Keywords: medical school, ethics education, medical ethics curriculum, competency assessment, clinical moral reasoning

Over the past few decades, courses in medical ethics have become an integral part of the undergraduate medical curriculum in North American and British medical schools (AAMC 1993; Dubois and Burkemper 2002; Hafferty and Franks 1994; Musick 1999; Savulescu et al. 1999). As the content of these courses evolves, the need for developing appropriate methods for assessing competence in medical ethics has become more obvious. Successful assessment of medical ethics education requires demonstration that the learner has acquired the specific skills or competencies defined by the curriculum (Bebeau 1993; Eckles 2005; Goldie et al. 2002; Savelescu et al. 1999; Selt et al. 1993). The importance of assessment as a vital component of the medical ethics curriculum has been broadly endorsed (Arnold and Forrow 1993; Carlin et al. 2011; Goldie 2000; Savelescu et al. 1999). There is, however, ongoing disagreement among medical educators over what the appropriate goals of medical ethics education should be (Novack, Epstein, and Paulsen, 1999). Some maintain that programs should aim at developing virtue, while others focus on stages of moral development,

A. Favia and L. Frank are equal first authors; E. Friedman and R. Rhodes are equal last authors.
Address correspondence to Rosamond Rhodes, PhD, Professor, Medical Education, Mount Sinai School of Medicine, Box #1076, One Gustave Levy Place New York, NY 10029, USA. E-mail: rosamond.rhodes@mssm.edu
legal knowledge, knowledge of ethical theory, professional behavior, or ethical reasoning. Thus, there is disagreement over how competency in medical ethics should be assessed (Arnold and Forrow 1993; Goldie 2000; Herbert, Meslin, and Dunn 1992; Mumford et al. 2008; National Board 2002; Self, Baldwin, and Wolinsky 1992; Smith et al. 1994; Wong and Chung 2003).

This article focuses on the goals of our medical ethics program and how we determine that our students meet minimum competency. Because of the critical relationship between a program’s goals and the design of an assessment strategy, we begin our discussion by providing an overview of the theoretical basis of our curriculum and an explanation of our program’s broad objectives. We briefly describe the competencies for our curriculum, our teaching methods, and the formative evaluation exercises that provide the students and us with information on whether they have achieved the defined competencies. This background will provide the framework for a discussion of our competency assessment. We conclude by presenting what we have learned from our study, exploring challenges that we identified, and offering directions for further efforts in medical ethics education and competency assessment.

THEORETICAL BACKGROUND

The medical ethics curriculum at Icahn School of Medicine at Mount Sinai (ISMMS) is integrated into all four years of undergraduate medical education (UME). Through a 24-year process of critical reflection, design, evaluation, review, and revision, we have developed the view that a medical ethics agenda should be based on professionalism. This perspective informs our positions on the aims and goals of a curriculum; where, when, and how medical ethics should be taught; and how to assess students’ competency.

Our curriculum is informed by the view that medicine is a profession that rests on a set of distinctive moral commitments. Thus, joining the profession entails accepting moral commitments. One thus, joining the profession entails accepting the moral commitments of the profession as one’s own. On this view, medical education is a transformative process in which medical educators help students to become exemplary physicians by learning both the content of and the justifications for their special responsibilities as physicians. To be professionally competent, students must not only become knowledgeable and skilled in the science and techniques of medicine; they must also become compassionate, caring, respectful of autonomy, and committed to upholding the ethics of medicine (Mount Sinai Student Handbook Student Code of Conduct 2012). Students have to mature into professionals with the appropriate dispositions and attitudes (i.e., the character and virtues) essential to being a good physician (Pellegrino 2002; Rhodes and Alfandre 2007; Rhodes et al. 2004). Thus, our approach explicitly embraces the dual aims of developing both competency (knowledge and skills) and virtue (character) (Pellegrino and Thomasma 1993).

As we understand these concepts, developing knowledge and skills goes hand in hand with developing virtue. When students come to understand and accept their professional responsibilities, they also are more disposed to act on the principles that express those commitments. A sincere and abiding commitment to a principle also entails developing an attitude toward actions of a certain sort and being inclined to act in accordance with the principle. In other words, there is no choice between principles and virtues; they are of a piece. To commit oneself to abide by a principle is to commit oneself to develop the character that will make acting from that virtue comfortable and likely. For example, a physician who is committed to truth-telling and beneficence will be a physician who is truthful and caring.

To help our students become exemplary physicians, we must first help them to develop a clear understanding of their distinctive responsibilities. Students must also gain the ethical knowledge and skills that they will need for navigating the common moral dilemmas of clinical practice. To achieve these competencies, students learn the basic principles of medical ethics during the first two years of our curriculum. They learn to recognize them in clinical scenarios, understand how they apply, and appreciate the actions that they require. Toward the end of their second year, they are then taught a replicable method for identifying moral conflicts and resolving dilemmas, a method that we call clinical moral reasoning.

Clinical moral reasoning is a framework that we specifically designed to guide health professionals in thinking through ethical dilemmas that arise in the inpatient and outpatient clinical setting. We developed the template after recognizing that students, house staff, and faculty often experience a complex ethical situation without being able to identify the relevant aspects of the case that could help lead to a solution. Often they do not know where to begin or how to proceed toward a resolution. These challenging situations typically have a moral dilemma at their core, either one involving two important ethical principles pulling in opposite directions (e.g., beneficence directing the physician to spend time with a needy patient and justice directing the physician to provide care for the several other patients that need attention) or a situation in which two mutually exclusive options could follow from a single ethical principle (e.g., beneficence directing the physician to perform surgery to minimize discomfort from an obstruction and beneficence directing the physician to abstain from performing surgery so as to avoid the patient enduring numerous possible complications). Clinical moral reasoning provides a framework that is useful for clarifying ethical dilemmas and prioritizing

---

1. This view is an explicit departure from the most commonly accepted view that medical ethics is common morality applied to clinical practice. Most notably, that standard view has been espoused in all of the several editions of Beauchamp and Childress’s Principles of Medical Ethics (2001) and Gert, Clouser, and Culver’s two editions of Bioethics. Our positive view can be found in Moros and Rhodes (2002), Rhodes (2007), Rhodes and Alfandre (2007), Rhodes et al. (2004), and Swick (2000).
ethics modules (three modules per year). Most of these sessions involve readings, short lectures followed by faculty-led small group case-based discussions, and some short writing assignments. Each preclinical ethics module is built around a different theme. The first five modules provide students with an understanding of the basic concepts of professional responsibility and medical ethics. These sessions focus on the concepts of professionalism, fiduciary responsibility, truth-telling and informed consent, autonomy, surrogate decision making, justice, and research ethics. They provide students with fluency in the vocabulary of ethics and enable them to identify core concepts in an array of clinical circumstances and to grasp what they entail (Table 2).

Education in medical ethics continues in the clinical years with ethics sessions incorporated into each clinical clerkship. Because expertise in both ethical theory and clinical medicine is involved, a member of the clinical department and a member of the ethics faculty co-lead the small group clerkship ethics sessions. These sessions revisit the topics addressed in the first two years in more complicated circumstances and take up additional issues that are particularly relevant to the clerkship. For example, students discuss communicating bad news in Geriatrics, adherence in Family Medicine, truth-telling in Neurology, family meetings in Critical Care, and treatment over objection in Psychiatry. These sessions aim at improving the students’ understanding of key concepts, promoting their ability to employ the concepts in a variety of settings, and developing the habit of discussing ethical issues with peers.

In the Surgery, Obstetrics–Gynecology, and Pediatrics clerkships, students are required to complete an Ethics

2. It is important to note that although students are learning ethical principles, they are not learning the “principism” of common morality as put forth by Beauchamp and Childress and others (see, e.g., Beauchamp and Childress 2001). Instead, our program focuses on an array of principles that govern the medical professions. Thus, we teach more than “the four principles” and emphasize that the principles function as reasons to justify professional action.
Table 2. Basic Principles/Concepts of Medical Ethics Taught in Years 1 and 2

<table>
<thead>
<tr>
<th>ASM Year 1</th>
<th>ASM Year 2</th>
<th>Epidemiology Year 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professionalism</td>
<td>Clinical justice</td>
<td>Importance of research</td>
</tr>
<tr>
<td>Duty to provide care</td>
<td>Assess decisional capacity</td>
<td>Ethical conduct of research</td>
</tr>
<tr>
<td>Beneficence/caring</td>
<td>Justified paternalism</td>
<td></td>
</tr>
<tr>
<td>Minimize harms</td>
<td>Assess surrogate appropriateness</td>
<td></td>
</tr>
<tr>
<td>Nonjudgmental regard</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respect autonomy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trust/fiduciary responsibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidentiality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duty to warn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truth telling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informed consent</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Project for their ethics session. For this assignment, each student identifies a case from the clerkship experience that involves an ethical dilemma. Students then meet in small groups with faculty preceptors to discuss their cases. Students are instructed to explore the ethics literature related to their case, prepare a 10-minute presentation, and lead a discussion of the ethical issues at the heart of their case. The individual feedback that the students receive on their written work is designed to help students further develop their skills in recognizing and analyzing ethical dilemmas that arise in clinical practice and to reinforce the model of clinical moral reasoning. Providing a venue for students to come together and discuss the ethical issues in actual cases that they encounter reinforces the idea that they need to be alert to the ethical issues in medical practice. In addition, as the group solidifies a consensus view around a case resolution, the students also reinforce attitudes of professionalism in themselves.

In sum, the ethics curriculum that we have created is well integrated into the span of undergraduate medical education. The training moves students from basic skills and knowledge to the more complex competencies that our trainees will need in their clinical practice. Throughout, we nurture the habits and attitudes that are essential for medical professionalism and we aim at promoting our students’ competency in medical ethics by helping them to:

- Articulate the ethical commitments of medicine.
- Describe and employ the basic concepts of medical ethics.
- Identify moral issues that arise in medical practice.
- Identify moral conflicts and formulate the dilemma as a question.
- Deliberate about ethical issues with peers.
- Resolve ethical dilemmas with supporting reasons from professional ethics.
- Identify matters that need further investigation.
- Plan the practical steps for implementing their decision, foreseeing future issues and taking them into account.
- Employ the bioethics literature in clarifying concepts and helping to resolve ethical issues.

DESIGNING AN ASSESSMENT OF MEDICAL ETHICS EDUCATION

A successful assessment method should reflect the clearly defined goals and objectives of an educational program. Developing an effective model for assessing student competence took several years, largely because it required fine-tuning our educational goals. While refining and revising an assessment model, we were also adjusting our curriculum to more explicitly focus on the knowledge and skills that we identified as essential tools for every physician. Although our program’s aims include developing the attitudes of professionalism, we focused our assessment instruments on the more feasible evaluation of medical ethics knowledge and skills.

The written assessment tools that we developed allow us to evaluate students’ understanding of ethical concepts, their ability to apply them to clinical cases, and their skills in resolving moral dilemmas. They require students to define the basic concepts of medical ethics and identify the specific concepts that arise in a case. By assessing students’ knowledge, skills, and ability to achieve a well-justified resolution when faced with a moral dilemma, we can test students’ ability to address the messy details of a real ethical dilemma in their future clinical practice.

The oral exercises in our assessment require students to demonstrate their skills in gathering important facts and mapping effective communication with peers, patients, and families. They also encourage students to reflect on their interactions and decisions, adopt appropriate attitudes, recognize the impact of values and beliefs on behavior, and demonstrate professional attitudes of nonjudgmental regard and respect.

To achieve the goals that we set for our assessment exercise, we determined that our assessment should be case based. We also recognized that cases should be similar to those that students will face as physicians and appropriate for a medical student’s level of education. Thus, we had to avoid both overly simplistic, obvious cases and unusual cases that most physicians would never encounter. For example, our cases address basic issues such as duty to provide care and informed consent, rather than whether to separate conjoined twins.

3. The development of the Ethics Project in the Surgery Clerkship is described in greater detail in Gligorov et al. (2009).
We also wanted our assessment to avoid imposing a rigid and narrow view of an ethical resolution and to be open to a range of different approaches to a case. This requirement meant that our assessment tools had to accommodate the multiple ethical issues that can be involved in a single clinical case and allow students to diverge on the issues they find most salient. Even cases that seem straightforward often involve more than one ethical issue. For example, a single case may embody issues related to the duty to provide care, justice, and a conflict of interest. Thus, the conflict could be framed in different ways. We designed our grading rubric and assessment methods to allow for this sort of variation in perspective.

The literature on medical ethics education discusses a variety of assessment methods (Carlin et al. 2011; Goldie 2000; Lynch, Surya, and Eiser 2004; Mattick and Bligh 2006; Patenaude, Niyonsenga, and Fafard 2003; Rezler et al. 1992; Self et al. 1993; Singer et al. 1994). We reviewed and considered a number of the suggested approaches to evaluation, including short-answer tests, moral development tests, reflections, short writing assignments, Objective Structured Clinical Examinations, and standardized patient (SP) encounters. In the end, we realized that none were suited to the assessment of the competencies that we consider critical.

For our unique competency assessment, we developed a structured question format that elicits written essay answers (Table 3). The questions are tailored to assess the specific ethics competencies that we have identified as critical for a physician to master. In making that design choice, we were fully aware that it would be difficult to standardize the evaluation of student answers. Achieving a standardized rating method for the written essay answers has, in fact, taken years of trial and error, critical review, and refinement.

Short writing assignments are part of our ethics curriculum and formative assessment program. Formative evaluation occurs while students are still in the process of learning and at multiple points during their education. The evaluations provide feedback for both the student and the educator (Slavin 1997, 491). Feedback can identify those concepts that a student needs to review and encourage a student to focus attention on developing skills in clinical moral reasoning. In addition, formative assessment reinforces seeing clinical ethics as an ongoing learning process and thus fosters the habit of discussing ethical issues with colleagues. Formative evaluation can also inform educators as to whether their teaching methods are successful and can help them to identify content and skill areas that require further instruction.

In addition to the short writing assignments that are returned with feedback, during their first three years of medical education, students complete three more formal formative assessment exercises. These exercises are useful in assessing knowledge of basic concepts and skills in applying a single concept to a case. They are not designed to evaluate competency in resolving a complex ethical dilemma. The first formal assessment occurs as part of the second-year midterm assessment in the ASM-2 course, administered just after students complete the last of the basic concept ethics modules. Students participate in an Objective Structured Clinical Examination (OSCE) in which ethics is one of the stations. It evaluates students’ understanding of a randomly assigned basic concept of medical ethics and their ability to apply the concept to a clinical case. Students are provided with one of several written cases, each designed to highlight a single concept. They are given a brief period to read the case, and, in writing, to identify and explain the relevant ethical concept and describe how it applies to the case.

The second formative assessment occurs at the end of the second year, shortly after the ethics module that introduces them to clinical moral reasoning. After they attend a lecture and work through some cases in a large group, students practice the skill of clinical moral reasoning in small-group sessions that encourage them to explore the ethical issues and resolve moral dilemmas in clinical cases. During the small-group sessions, students receive constructive comments from their peers and preceptor. The medical ethics evaluation that follows is part of the end of second year comprehensive assessment exercises, Compass-1. Compass-1 includes several standardized patient encounters, a required medical writeup, and the ethics assessment. Students are presented with a series of three vignettes that expand upon a clinical case. The questions following each vignette lead students through the steps of clinical moral reasoning. Their written responses demonstrate and

<table>
<thead>
<tr>
<th>Competencies assessed</th>
<th>A. Written exercise</th>
<th>B. Oral exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understands basic concepts</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Applies basic concepts to cases</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Identifies ethical issues in clinical practice</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Accesses the bioethics literature</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Discusses ethical issues with peers</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Identifies moral conflict</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Formulates an ethical question that clarifies the dilemma</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Resolves ethical dilemma supported by reasons</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Plans practical steps, foreseeing future issues</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
Students are graded on a scale of 1–4 for each component of their answers: 1 for an inadequate response, 2 for low competence, 3 for competence, and 4 for high competence. For example, we rate the students on how well they identify and explain ethical principles and concepts and show how they are relevant to the case. Following our rating rubric, when students identify two principles, explain them both, and apply both to the case, they receive a total of 4 points for the answer to the first question.

The third-year comprehensive assessment, Compass-2, is based on the six standardized patient (SP) encounters used to assess students’ history-taking, physical examination, and oral and written communication skills. We recognized that significant social factors were already embedded in these SP cases and therefore chose to use them as the basis for the ethics assessment. Again, this approach encourages students to appreciate that ethical reflection is a constant feature of good clinical practice, rather than enforcing the view that ethics is a discrete domain, separate from medicine and relegated to a team of consulting bioethicists.

In order to assess the full range of ethics competencies, students complete written and oral exercises in their Compass-2 ethics assessment (Figure 1 and Table 3). The morning after their SP encounters, students submit a written ethics exercise. For this component of the assessment, students choose any two of the six SP cases for a written ethical analysis. Students then answer a series of questions that mirror the process of clinical moral reasoning and allow us to evaluate each component using rating tools similar to those used in evaluating the Compass-1 ethics exercises. A standardized rating sheet and rating rubric guide the ethics faculty in grading the exercises. The evaluated written exercise and faculty feedback are returned to students. As we explain in the following, we have found that the written exercises provide material for a meaningful evaluation of students’ minimal competence in applying moral concepts and in clinical moral reasoning.

We have designed the questions on Compass-1 and Compass-2 to parallel the steps of clinical moral reasoning. To capture the students’ competency in applying the model, our assessment instruments are, therefore, closely aligned with the clinical moral reasoning model. For example, in Compass-2 students are first asked to identify two relevant ethical principles and to explain how those principles conflict in the particular case. This is precisely the guidance that ethics is a discrete domain, separate from medicine and relegated to a team of consulting bioethicists.

In order to assess the full range of ethics competencies, students complete written and oral exercises in their Compass-2 ethics assessment (Figure 1 and Table 3). The morning after their SP encounters, students submit a written ethics exercise. For this component of the assessment, students choose any two of the six SP cases for a written ethical analysis. Students then answer a series of questions that mirror the process of clinical moral reasoning and allow us to evaluate each component using rating tools similar to those used in evaluating the Compass-1 ethics exercises. A standardized rating sheet and rating rubric guide the ethics faculty in grading the exercises. The evaluated written exercise and faculty feedback are returned to students. As we explain in the following, we have found that the written exercises provide material for a meaningful evaluation of students’ minimal competence in applying moral concepts and in clinical moral reasoning.

We have designed the questions on Compass-1 and Compass-2 to parallel the steps of clinical moral reasoning. To capture the students’ competency in applying the model, our assessment instruments are, therefore, closely aligned with the clinical moral reasoning model. For example, in Compass-2 students are first asked to identify two relevant ethical principles and to explain how those principles conflict in the particular case. This is precisely the guidance of the first two steps of the clinical moral reasoning model. This parallel structure provides an opportunity for students to practice each step of clinical moral reasoning. At the same time, it allows evaluators to assess a student’s competence at each point in the reasoning process. Evaluators can thus identify precisely what a student has missed and communicate what he or she needs to do to improve. It also allows the faculty to identify students who need extra work in ethics. For remediation, a student who fails to demonstrate minimum competency in the Compass-2 ethics assessment will have an individualized review session with a member of

---

### Table 4. Concepts discussed by students in Compass-2A

<table>
<thead>
<tr>
<th>Principles</th>
<th>DJ</th>
<th>CD</th>
<th>DS</th>
<th>JeTh</th>
<th>FR</th>
<th>JaTu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidentiality</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respect for autonomy</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Nonjudgmental regard</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truth-telling/delivering bad news</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Assess surrogate appropriateness</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duty to provide care</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decisional capacity</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informed consent</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justified paternalism</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beneficence</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimize harm</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trust</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fiduciary responsibility</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsibility to peers and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>institutions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professionality</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

reinforce their skills in identifying ethical conflicts and using clinical moral reasoning to resolve the dilemmas.

The ethics cases for Compass-1 are based on a medical case discussed earlier in the second year of the Art and Science of Medicine course. Again, this approach reflects our view that ethics education should be horizontally integrated into the medical school curriculum and assessments. Incorporating a medical case into the ethics assessment is intended to remind students that ethical elements are involved in every patient interaction. As Professor Loretta Kopelman pointed out when she reviewed our assessment tools, this strategy “reinforces the idea that ethics is part of the routine of everyday medicine and that the discerning clinician should look for the ‘signs and symptoms’ related to moral and social problems as well as others.”

For each case in the Compass-1 ethics assessment, students identify at least two ethical principles and then briefly explain each principle and how it is relevant to the case. Next, students identify an ethical conflict that the case raises for the doctor and express that conflict as a question. Then students either explain which of the conflicting principles should be given priority and why, or what further information they would need in order to resolve the dilemma and how that information would help. Finally, students must formulate a plan of action, including suggestions about significant future ethical issues and the practical steps that they would take in implementing their decision.
Part A - Instructions for Written Assignment

Part A Instructions: Consider the six cases of your Morchand Center assessment exercise. Each case raises several ethical issues for the physician involved.

Choose two of the cases to discuss. Answer the following questions (a-e) with respect to each of your selected cases. Feel free to consult the bioethics literature to inform your answers, but your answers should focus specifically on the questions below. Taken together, your responses to (a-e) should constitute an ethical analysis of the case and demonstrate your competence in moral reasoning.

1a. Which patient will you discuss? (Choose one of the six listed above.)

1b. Which ethical principles/concepts are relevant to the case? (List at least two. Briefly explain each and explain how the principle/concept is relevant to the case.)

1c. What is the ethical conflict that this case raises for the doctor? (Notice which of your listed principles conflict, and formulate this conflict as a question for the doctor.)

1d. How should this conflict be resolved? (Which of the conflicting principles should be given priority in this case? Why? Explain.) OR What further information would you need to learn about the case to resolve the dilemma? How would the information help you to resolve the ethical conflict?

1e. Given your analysis of the conflict, how would you implement your plan? (Explain your implementation of the plan. Which significant ethical issues do you foresee down the road? Which factors are especially relevant, and why? What practical steps should you consider? Why?)

Part B - Instructions for Oral Presentation

Part B Instructions: After your Morchand Center exercises, you will be assigned a principle/concept to address in the context of a specific case as the basis for leading your small group discussion. (Feel free to use this assignment as one of your cases for Part A.) You are to consult the medical ethics literature, identify at least one article or chapter to help inform your thinking on the issue. Present the case in your small group discussion on the second day of the Compass-2 assessment exercises. Your presentation should be no longer than 8 minutes. For the rest of your time you will lead a discussion of the case. Be sure to include the following in your presentation:

1. Summarize the medical aspects that are relevant to the ethical issue raised by the case.
2. Explain the ethical concept that you were assigned and how it relates to the case.
3. Discuss the article/chapter that you selected and explain how it has informed your understanding of an ethical dilemma raised by the case.
4. Discuss what you think should be done and why, or explain what more you would need to learn about the case in order to reach a decision, why that information would be important, and how you might find it.
5. Prepare at least five questions to use in leading your peers in a discussion about what should be done in this situation and why. Also, if you have encountered a case during your 3rd-year that vividly illustrates this issue, feel free to present it in your discussion.

Figure 1. Instructions for Compass-2 A & B (2011)—medical ethics assessment.
Medical Students’ Competency in Medical Ethics

literature, explain how their assigned concept applies to the case, and pose discussion questions focused on resolving the ethical dilemma. The oral presentations and discussions occur in small groups of five or six students overseen by a faculty member who serves as preceptor and evaluates each student’s performance. Students receive written feedback on both their presentations and their participation in their group’s discussions.

We have found that it is not feasible to use the oral presentations for an assessment of competency. The oral presentations are, however, extremely valuable for achieving other formative goals. They provide an opportunity for reviewing basic concepts. The sessions also allow students to express their views and to legitimize their ethical concerns. We want to foster group discussion in the oral sessions because expressing aloud commitments to the principles of medical ethics serves to encourage attitudes of professionalism (Pellegrino 2002). These case-based conversations among peers also prepare the students for the discussions that they are likely to have on the floors during their residencies. We do grade students on their performance in the oral sessions because grading encourages students to participate and perform well, and because grading rewards outstanding effort. Preceptor feedback is returned to students as a formative assessment of their specific strengths and weaknesses (Figure 2).

Faculty preceptors who evaluate students’ small-group performance typically give all of the students a high rating because the group discussions are almost always excellent. Students report that they find the small-group sessions

<p>| Compass-2B, 2011: Medical Ethics Oral Presentation Assessment |</p>
<table>
<thead>
<tr>
<th>EVALUATION SHEET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student’s Name:</td>
</tr>
<tr>
<td>Preceptor:</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

**PRESENTATION & LEADING THE DISCUSSION**

Explains the assigned concept and how it relates to the case.  
*Comments:*

Identifies a relevant reading and explains how it sheds light on the case.  
*Comments:*

Asks thoughtful questions and shows facility in leading the discussion.  
*Comments:*

**ADDITIONAL REMARKS:**

**DISCUSSION PARTICIPATION**

Shows respect to peers and tolerance of other reasonable views.  
*Comments:*

Contributes to the discussion.  
*Comments:*

**ADDITIONAL REMARKS:**

Figure 2. Compass-2B evaluation sheet.
interesting and valuable as a review of important concepts. Faculty members report that they are tremendously impressed by the quality of the students' performance and by how much the students have advanced in knowledge and moral reflection since their first-year ethics sessions.

**METHODS AND RESULTS**

To provide a useful assessment of our students' competency in medical ethics we aimed at creating reliable and valid ethics assessment tools:

1. For the instruments to be reliable, we had to show that we could achieve interrater agreement in our ratings of the same exams (Gwet 2010).
2. For the instruments to be valid, we had to show that the exercises actually assess knowledge and skills that demonstrate competency in medical ethics. This required that experts in medical ethics education review our materials and provide their opinions on whether our instruments actually measure competency in medical ethics.
3. To determine whether our assessment tools might be useful outside of our program, we wanted to see whether medical ethics educators who were not part of our team could use the tools and achieve similar results in grading the same student written exercises.

**Reliability**

Our methodology for demonstrating interrater agreement involved eight raters from the ethics faculty, each evaluating nearly 40 student exercises. All told, the raters evaluated the performance of 137 students who each completed a written ethics exercise based on two SP cases. Two raters individually evaluated each student's performance, with the evaluation of a student's performance taking approximately 20 minutes. According to our rating template, the highest possible total score was 32. We compared the eight individual raters' evaluations of mutually scored exercises. For our study, the raters were randomly assigned to each two-case student assessment exercise, and two raters rated each exercise. We calculated an intraclass correlation (1, 1) as discussed in Schrout and Fleiss (1979). The calculated correlation for the exercise is .22952 (Figure 3).

Our results show that although our raters did not always give identical scores, scores were sufficiently close to suggest agreement (Figure 4). We assigned a 5-point range to the rating categories of high competency (28–32), competency (23–27), and low competency (18–22), and a 17-point range for inadequate scores (17 and below). Paired raters scored the same student within 5 points of each other on 119 of the exercises (87% rater consensus). Raters scored more than 5 points apart on 18 pair-rated exercises (12%); however, their disagreement was no more than a difference of 10 points, which was equivalent to one letter grade category difference. Put in terms of academic letter grading (A–F) of essay exams in the humanities, these results can be understood as showing our raters agree on a letter grade 87% of the time, and disagree by only one letter grade on only 12% of the exercises.

Using the higher score for each Compass-2 exercise in 2011, the overall scores allowed us to distinguish levels of competency in the students' performance (Figure 5). Of the 137 students who were evaluated, 59 students scored in the “competence” range, 30 in the “high competence” range, 40 in the “low competence” range, and 8 in the “inadequate” range. From discussions with the students who required remediation, a few explained their poor performance as a time problem related to conflicting clinical obligations, some had failed to invest adequate effort, and others had difficulty understanding moral dilemmas and how to prioritize conflicting principles.

It was also instructive to see that our evaluation tools were flexible and accommodated students emphasizing different issues in their analyses of the same case. For example, the SP case CD of an alcoholic with abdominal pain was analyzed by some students in terms of a conflict between confidentiality and respect for autonomy, while other students focused on the conflict between nonjudgmental regard and justice (Figure 3). This result suggests that our model for clinical moral reasoning is broadly applicable. It also suggests that our evaluation tools could be used by other medical school programs with SP cases of their own design.

**Validity**

In order to validate our assessment tools, we enlisted the help of three experts with many years of experience in medical education and teaching medical ethics to medical students. Professor Leonard Fleck of Michigan State University, Professor Loretta Kopelman of Georgetown University, and Professor Laurence McCullough of Baylor College of Medicine agreed to serve as reviewers of our ethics
Figure 3. Compass 2 medical ethics assessment: range of total score by rater. (Color figure available online).

<table>
<thead>
<tr>
<th>Rater</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R6</th>
<th>R7</th>
<th>R8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>24</td>
<td>23</td>
<td>25</td>
<td>23.5</td>
<td>21</td>
<td>23</td>
<td>22.5</td>
<td>23</td>
</tr>
<tr>
<td>75th Percentile</td>
<td>27</td>
<td>24</td>
<td>28</td>
<td>26</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>25th Percentile</td>
<td>21</td>
<td>20</td>
<td>22</td>
<td>21</td>
<td>15</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Mean</td>
<td>24.4</td>
<td>22.1</td>
<td>24.8</td>
<td>23.4</td>
<td>20.8</td>
<td>21.9</td>
<td>22.5</td>
<td>22.6</td>
</tr>
</tbody>
</table>

R5 mean significantly lower than R1 and R3.

Figure 4. Interrater reliability: agreement on level of competency demonstrated on Compass-2A (n = 137) papers graded by two raters.
assessment. Our methodology involved providing our reviewers with the Compass-2 SP cases and our Compass-2 ethics assessment materials, including the exercises; the instructions that the students received; our grading rating sheets (Figure 6); and our evaluation guide rubrics (Figure 7). We wanted them to determine whether our tools are “practical, pertinent, and related to the purpose of the test” and whether they measure what we state that they measure, which would show the validity of the instruments (Mosier 1947, 192). Specifically, we asked these experts to answer two questions:

1. Does Compass-2A evaluate skills in identifying and analyzing ethical dilemmas that arise in clinical practice?
2. Can the Compass-2A exercise be useful in distinguishing students with high competence from students with inadequate understanding of ethical concepts and inadequate skills in clinical moral reasoning?

All three expert reviewers agreed that our instruments were well suited for evaluating medical student competence in medical ethics. In addition, the expert reviewers agreed that the evaluation tools and instructions for faculty

For example, Professor Fleck wrote that his answer to both of the questions we posed “is a strong yes.” According to Professor McCullough, “This is an excellent set of pedagogical materials” that is “nicely complementary with the goal of producing students who will be able to think through ethical challenges in clinical practice in a disciplined, reliable fashion that can be assessed during learning.” Professor Kopelman remarked that “you have developed impressive tools to evaluate your students’ performance and to refine your curriculum.” Specifically, with respect to the first question, Professor McCullough responded that “the core skill that this assessment is designed to measure is the attainment of practical skills of normative ethics reasoning about cases using ethical principles.” McCullough also commented that our assessment questions “reflect well-accepted approaches to normative ethical reasoning in the clinical setting.” In answer to the second question, Professor Fleck stated that “your instrument should allow you to judge reasonably well which students have a well developed capacity to identify morally relevant considerations with regard to a concrete clinical ethics problem from those who fail to identify many considerations which they ought to recognize as having moral relevance.” Professor Kopelman concurred, writing that the “Compass-2A exercise should be very useful in distinguishing students with high

Figure 5. Compass-2A student grade distribution, n = 137 students.
Competency from students with inadequate understanding of ethical concepts.”

7. Professor McCullough remarked that the Student Instructions seemed “clear and fair” and that the Rater Evaluation Guide was a “detailed . . . clear, easy-to-use evaluation sheet for faculty.” Professor Fleck pointed out how difficult it is to achieve interrater agreement in essay-type evaluations, saying, “I like the way you have broken down each item into what are supposed to be identifiable sub-parts which should result in various raters making the same judgment.”

8. For example, Professor Kopelman wondered whether students have sufficient time in their busy schedules to complete the assignment. Both Professors Fleck and Kopelman suggested that we supplement our assessment by asking students to consider objections or weaknesses of their proposed resolution. Also, Professor reviewers are helpful for demonstrating face validity and content validity. This information does not, however, support interrater agreement.

Interrater Agreement With Outside Raters

Our next step was to see whether raters outside of our group could use our tools and reach similar rating of assignments. We asked an assistant to select two anonymized student exercises for which our paired raters had close agreement in their ratings, examples that were neither rated with the very highest nor the very lowest scores. We sent those anonymized examples to our three experts and an

Kopelman was skeptical about assessing individual students’ performance on the oral component of the exam. Their overall opinions can be summed up in Professor Fleck’s response: “What you have in your evaluation instrument is the best effort I have seen thus far in trying to assure interrater reliability in judging ethics essays from medical students.”
All items of an answer must be explicitly stated. Rate the answer to each question separately. Do not infer meaning. Rate only what is actually written.

Keep in mind that the rating sheets provide feedback to students. Make your comments clear and appropriate. Every report should include at least one written comment.

A comment is required for every question score of “1” or “2”.

Scoring: 4 = high competence; 3 = competence; 2 = low competence; 1 = inadequate

b. Rating Rubric
Identifies and explains ethical principles/concepts and shows how they relate to the case and how they capture the conflict.
Elements: identifies relevant principles (2); explains relevant principles (2); relates relevant principles to case (2). [When the same principle is given two different names or when an irrelevant principle is mentioned or explained, the student should not be allotted credit.]
4 = all 6 elements
3 = 4-5 elements
2 = 2-3 elements
1 = 1 element

c. Rating Rubric
Identifies a relevant ethical conflict. This analysis should be compatible with the answer to b. If not, provide a comment, but score according to the rubric.
Elements: (a) relevant conflict noted (1); (b) expresses the conflict in the form of a question (1); (c) both relevant principles implied (1 point) or both relevant principles implied & at least one named (2 points)
4 = all 4 elements
3 = 3 elements
2 = 2 elements
1 = 0-1 element

d. Rating Rubric
Explains how the conflict should be resolved and gives reasons to support the position.
Elements: (a) identifies the conflict or explains how additional information would be used to resolve the conflict; (b) identifies facts about the case that are relevant to the decision; (c) explains why they are relevant, (d) shows excellent insight and/or sensitivity to the specific nuances of the case.
4 = all 4 elements
3 = 3 elements
2 = 2 elements
1 = 0-1 element

e. Rating Rubric
Explains implementation of a plan.
Elements: (a) provides an appropriate plan; (b) plan is clear and detailed; (c) shows excellent insight and/or sensitivity to the specific nuances of the case; (d) additionally, identifies future issues.
4 = all 4 elements
3 = 3 elements
2 = 2 elements
1 = 0-1 element

Figure 7. Compass-2A 2011 rater evaluation guide.

additional six experts in medical school ethics education. Although all of our ethics faculty raters and the three expert reviewers are trained in philosophy, some of the outside raters have other academic backgrounds. These experts used our assessment tools to rate the two student exams. Each of their ratings provided us with scores on a total of four case analyses, two per student exercise.

When we compared the ratings of the outside raters with our own, we found that some of the ratings by outside experts were very close to ours; others were not (Figure 8).
Six of the nine outside raters were in close agreement with the ISMMS raters on the evaluation of two cases; four were in close agreement with the ISMMS raters on the evaluation of three cases. Three outside raters trained in philosophy were in close agreement with the ISMMS raters on the evaluation of all four cases. The possibility of rating agreement indicates that our assessment tools may be useful in other medical ethics programs.

We attribute the discrepancies between the ratings by ISMMS raters and outside raters to several factors. ISMMS raters have spent years working together to fine-tune our assessment tools and curriculum. We also spend several hours together before grading the Compass-1 and Compass-2 exams and prepare by rating exams from the previous year, comparing our scores, and discussing the rationales for the scores that we assigned. In these sessions, we discuss the cases and relevant principles, and thereby develop a common view of what would count as good answers. Through this process we come to see the cases in the same way and develop a shared view of what student answers should include.

Our outside expert raters had none of the shared experience that we have created over years of working together. Divergent readings of the cases could also explain significant differences in ratings. Furthermore, the meanings of our grading sheet and rating rubric are very clear to us, but our tools may not convey everything that one needs to know in order to understand our expectations for student competency and to consistently grade the exercises. In addition, because many concepts in medical ethics are interpreted in somewhat different ways, the ethics faculty raters at ISMMS may also understand the concepts and principles of medical ethics, such as “autonomy,” in a somewhat different way than some of the outside raters do. All of this suggests that there may be a learning curve for developing reliability with our assessment tools. We may also need to refine our rating rubric to more fully explain and communicate the instructions.

**DISCUSSION**

**Plans for the Future**

At this point we have demonstrated interrater reliability in our grading and shown the validity of our methods for assessing competency in medical ethics education. Nevertheless, we plan to further refine our assessment tools and improve rater training in applying the scoring rubric. These efforts will focus on increasing the reliability of our own ratings and making our instruments more useful for the evaluation of ethics education at other medical schools.

Our evaluation tools allow us to identify the areas where students require additional work. For example, student performance on the most recent round of exercises revealed that they needed to improve on their ability to identify and resolve moral dilemmas. To address this need, this year we initiated an additional formative training exercise in three of the students’ third-year clerkship ethics sessions. Students now submit written ethical analyses of their cases by answering a series of questions that replicate the first steps of the clinical moral reasoning template. Students receive personalized feedback on those targeted written exercises. We plan to compare the performance on this year’s Compass-2 assessment to the overall performance of last year’s students to see whether this intervention has helped to address those specific educational needs. We also plan to examine this year’s exercises to see whether we can identify other areas where educational intervention could help to improve students’ competency in medical ethics.

**Summary**

Our efforts in medical ethics education and competency assessment have produced an integrated model of goals, methodology, curriculum, and competency assessment. The entire model is directed at providing students with the ethical knowledge, skills, and attitudes required of an exemplary physician. The curriculum helps students to understand and apply key concepts of medical ethics, to independently identify ethical issues, to analyze them in terms of moral conflicts, and to come to justified resolutions. The curriculum is structured developmentally to build from basic to more complex knowledge and skills, and it culminates in a valid and reliable assessment exercise.

In sum, we have developed assessment tools that allow us to evaluate medical student competency in medical ethics and to identify those students who require remediation. The
exercises mirror the process of clinical moral reasoning and reinforce the skills and habits of critical analysis. Such exercises are also sufficiently flexible to reflect the complexity of clinical cases.

ACKNOWLEDGMENTS

We appreciate the contributions of James Hitt and Michael Nair-Collins who worked on the development of our rating instruments while they served as Ethics Fellows at Icahn School of Medicine at Mount Sinai.

AUTHOR CONTRIBUTIONS

All of those listed as authors made substantive intellectual contributions to this study as well as drafting and revising the article. Each author approved the final version. Amanda Favia and Lily Frank were the lead authors on this paper. They wrote the first draft of the paper and worked on subsequent revisions. Erica Friedman and Rosamond Rhodes initiated the design and implementation of the assessment project and this study related to the project. Both contributed significantly to the numerous drafts and revisions. Amanda Favia, Lily Frank, Nada Gligorov, Steven Birnbaum, Paul Cummins, Kyle Ferguson, and Katherine Mendis all contributed to the refinements and evolution of the project and they all contributed bits and pieces to the writing and editing. Steven Birnbaum also created and/or refined our graphic tables and figures. Robert Fallar provided guidance on the study designs that we used to achieve inter-rater agreement, validity, and reliability. He provided data analysis, reviewed the paper sections related to data collection and analysis, calculated the intra-class correlations, and wrote portions of the section on methods and results.

FUNDING

The John Conley Foundation supported work on this project.

COMPETING INTERESTS

None declared.

ETHICAL APPROVAL

An Exemption Letter was obtained from the Mount Sinai Institutional Review Board. This project involved educational research on the performance of all students, no recording of identifiable student information, and no interaction with students beyond their regular comprehensive assessment exercises. Thus, it was exempt from informed consent requirements. Students were, however, informed that their exercises were being used in a study aimed at assessing our evaluation tools.

REFERENCES


