Cognitive Moral Development of Medical Students

Desenvolvimento Moral Cognitivo dos Estudantes de Medicina

Donnie J. Self, Ph.D.

"To educate a person in mind and not in morals is to educate a menace to society."

Theodore Roosevelt

"Modern man worships at the temple of science, but science tells him what is possible, not what is right."

Milton Eisenhower

RESUMO

É apresentada uma breve discussão sobre a teoria ética e o processo de tomada de decisões éticas como um substrato para considerarmos uma série de estudos empíricos sobre o desenvolvimento moral cognitivo de estudantes de Medicina. São incluídos critérios para a decisão moral, a distinção teleológica e deontológica, e três métodos para a tomada de decisões. Segue-se uma descrição dos seis estágios na teoria de desenvolvimento moral de Kohlberg e uma breve elaboração da teoria.

ABSTRACT

A brief discussion of ethical theory and the process of ethical decision-making is given as a background for consideration of a series of empirical studies regarding the cognitive moral development of medical students. Included are criteria for a sound moral decision, the teleological and deontological distinctions, and three methods for ethical decision-making. All this is followed by a description of the six stages in Kohlberg’s cognitive moral development theory and a brief elaboration of the theory.

I have been asked to address the moral and ethical development of medical students. I will do that by reporting a series of studies in which I have investigated these issues. But I take my assignment to be somewhat broader than that. In terms of an overview of what I would like to cover, I will share a little of the theory of ethical decision making as well as a summary of Kohlberg’s Cognitive Moral Development Theory and then review several empirical studies involving medical students and other healthcare professionals. I will also elaborate on the instruments used in assessing moral development or moral maturity. These instruments have all been derived from Kohlberg’s cognitive moral development theory which is the intellectual foundation upon which the empirical studies are based.

1 Professor of the Department of Humanities in Medicine – College of Medicine – The Texas A&M University

2 Resumo elaborado pela Comissão científica do Simpósio
Perhaps a brief overview of ethical theory will help us better understand the studies to be reported later. Morality has been defined in many ways from both secular and religious frameworks. But it has been generally agreed that ethics is the study of the principles of human behavior with respect to moral actions. So the process of any ethical or moral decision-making inevitably involves values and choosing among competing values. Of course, not all values are moral values. For example, one can have aesthetic values, economic values, religious values, political values, functional values, and so forth. To say that this is a good pencil is to make a value judgement about it but not a moral judgement about it; whereas to say he is a good husband or she is a good mother is to make, I believe, a moral judgement. What about he/she is a good physician? Is that a moral judgement? What makes a judgement a moral judgement? Or what makes moral behavior moral? Well, first we need to distinguish between morés and morals. Morals are not determined by majority acceptance as morés are. Morés are customs within a society that the majority accepts. For example, gentlemen taking their hats off when indoors or gentlemen opening doors for ladies are socially accepted customs or morés in the United States. But if you don’t take off your hat, you haven’t done something immoral. Rather, morals have to do with interpersonal interactions or interpersonal relationships based on the principle “Always treat persons as an end in themselves.” That is, never use people without taking into consideration their interests. So morality takes into consideration other people’s interests and involves interpersonal relationships. There are potential problems with this interpretation of morality, but not, I think, insolvable ones. The problems might include moral obligations to animals and/or moral obligations to oneself.

Now when you face an ethical dilemma and are making moral judgments there are three issues that always need to be considered. They are the empirical issues, the conceptual issues, and the normative issues.

1. Empirical issues – i.e. what are the facts. For example, with concerns about death, has the heart stopped beating, has all electrical activity in the brain stopped, have the lungs stopped breathing?

2. Conceptual issues – i.e. what does a term mean? For example does the term ‘death’ refer to functions of the heart, lungs, brain, some combination or so forth.

3. Normative issues – i.e. who ought to do what or who should do what?

You see all of these issues come into play when you consider hypothetical case studies and when you are faced with actual decision-making as well. However, when you are actually doing ethical decision-making instead of just talking about it, you need some way to distinguish poor decisions from good decisions. I want to suggest the following criteria for a sound or good ethical decision or moral decision.

**CRITERIA FOR A SOUND ETHICAL OR MORAL DECISION**

You may want to add others to the list or delete some of these.

A sound decision must be rationally arrived at, not just gut feelings. You can give reasons and justifications, some better and some worse, for choosing one alternative over others. But when asked, you can give a rationale for your actions. One should always be able and willing to defend one’s actions. Science is based on reason and so is ethics.

For a decision to be sound all the relevant data must be fairly considered. The decision must not be biased any more than you can help. Granted we all have our biases, and they show through sometimes. But the data should be objectively considered. Try to put yourself in the opposite position empathetically and see if you still conclude that the decision reached is the best even if you do not like it? Sometimes we are morally obligated to do something we do not like.

A sound decision must be universalizable. You would be willing for all physicians or healthcare providers to do likewise under relevantly similar conditions. This is thought to be the cornerstone of morality by most people. For example, remember the Categorical Imperative of Immanuel Kant or the Judeo-Christian Ethics of the Golden Rule.

A sound decision must be coherent and consistent, i.e., fit in well with other moral beliefs you hold and preferably with the moral beliefs of others but at least with your own beliefs. Philosophically, this is one of the most important criteria to meet.

All this raises theoretical issues regarding whether value judgements make truth claims at all and can be true or false i.e., do value judgements make knowledge claims and have any cognitive content or in contrast are they just purely personal private expressions of opinion and subjective inner feelings, with persuasive force perhaps, but no cognitive content. I realize that this question is not at all settled. But I will set that theoretical discussion aside for another occasion and just assume that we agree that moral judgements are knowledge claims that can be true or false, and that we can get them right or wrong. Otherwise ethics is an exercise in futility.

Try to apply these criteria to the decisions you face in your personal day to day work.

Now with these criteria for a sound ethical decision in mind, perhaps a couple of fundamental philosophical distinctions would be
helpful in regard to ethical decision-making. When faced with an ethical
decision one usually takes one of two general approaches.

1) To choose an act because it has the best consequences, a sort of
greatest good for the greatest number or

2) To choose an act because it is simply the right thing to do by the
very nature of the act. It is in keeping with the rules by which you govern
your life.

In philosophy this is known as the Teleological vs. Deontological
Distinction and sometimes called the Consequence vs. Rule Distinction.

TELEOLOGICAL AND DEONTOLOGICAL
APPROACHES

First, there is the Teleological Approach or Consequence Approach.
The utilitarianism of Jeremy Bentham and John Stuart Mill is the best
known example of a teleological approach advocating the greatest good
for the greatest number. Teleological ethics is consequence oriented.
That is the defining or distinguishing characteristic of all teleological
ethics. The rightness or wrongness and goodness or badness of an act
depends upon the results or consequences. It is end or goal oriented
although that goal might be pleasure, utility, power, knowledge, freedom,
health, or so forth. The production of the goal bestows a right-making
characteristic on the act. It defines ‘right’ and ‘good’ for the teleologist.
The final appeal, directly or indirectly, must be to the production of this
goal or ultimate standard of right or good. So an act is right or good if
and only if it is conductive to the production of the greatest amount of
whatever your goal is, whether it be utility, pleasure, freedom, or anything
else and bad or wrong otherwise if not conductive to your goal.

In contrast there is the Deontological Approach or Rule Approach.
Deontological ethics is not consequence oriented. That is the defining
characteristic of it. It is just the opposite of teleological ethics. An act is
right by virtue of its formal characteristics or by the very nature of the
act, i.e. by virtue of simply being that kind of act. For example, one might
hold that rape is wrong not because of the consequences to the victim but
due to the very nature of the act involving a violation of dignity, autonomy,
and respect. There simply are other considerations in rightness and
goodness besides consequences such considerations as integrity, trust,
dignity, respect, the keeping of a promise, the keeping of a command of
God, and so forth.

The most common example in the philosophical tradition is the
ethics of Immanuel Kant, the 18th century German philosopher, who
formulated the Categorical Imperative of always treating a person as an
end in themselves and never as a means, i.e. never use people, and the
Principle of Universalizability of doing those acts and only those acts
that you would be willing to have become universal law. The most
common example otherwise is the Judeo-Christian tradition of the Golden
Rule and the Decalogue or Ten Commandments.

With this theoretical background let us consider three practical
methods of ethical decision-making.

1. The Public Defense Rule of Thumb is the first line of defense or
assessment. Would you be willing to publicly defend and endorse your
action? Is it okay for everybody to know about it? If not, it is probably
wrong, probably but not necessarily. No doubt you can think of
exceptions. For example, perhaps certain sexual behavior or any illegal
behavior that is not morally wrong of course that depends a lot on your
belief about the relationship between legality and morality. You might
include, for example, such acts as not using seat belts while driving,
smoking pot or marijuana, or in human medicine performing active
euthanasia, and so forth.

2. Then there is Universalizability Approach or Golden Rule. “Do
only that which you would be willing for anyone else to do under
relevantly similar circumstances,” or perhaps better known as “Do unto
others as you would have them do unto you.” But this is not well
formulated. Maybe it should read “Do unto others as others would have
you do unto them.” But there are problems with this approach also. A
counter example is that you might be willing for everyone else to
discriminate against certain groups just like you do! But that does not
make it morally right even if you are willing to universalize it.

3. Lastly, there is Social Perspective Taking, Role Exchange, or Role
Switching approach, which Lawrence Kohlberg referred to as playing
Moral Musical Chairs. Look at the issue from the perspective of each
person in the dilemma and do that which is the most advantage for the
least advantaged person in the dilemma. It is like an Ideal Observer
Theory. Do that which you would want done if you did not know which
person in the situation you were going to be, that is, act from behind
the veil of ignorance. All of this is based on John Rawl’s conception and
explication of the concept of justice found in his book Theory of Justice
– a modern philosophical classic.

Principle of Justice – the duty or obligation to allocate social burdens
and benefits fairly among all persons. But what constitutes fair allocation
or distribution is often very difficult to determine. This is always an
issue in the question of allocation of limited resources, whether it be on
a macro level or a micro level.

Now intellectual thought in the field of the psychology of moral
development has been significantly influenced for the past several decades
by Lawrence Kohlberg’s cognitive moral development theory, which he
acknowledged was derived from Piaget and Dewey. It notes that cognitive
development is a necessary but not sufficient condition for moral
development.
The goals of cognitive moral development theory are: (1) to improve the quality of thought about moral and social issues, which implies that there are better and worse ways to think about moral issues and that not just any values are as good as any other values, (2) to aid the understanding of the basic documents upon which the United States society is built, namely the Declaration of Independence and the Constitution, (3) to integrate moral thought and action so that moral behavior and action will become more consistent with moral reasoning, as research has shown, and (4) to alleviate the problems associated with the individualism of values, such as narrow-mindedness, prejudice, intolerance, lack of adaptability to change, and so forth.

Based on decades of quantitatively reproducible research, Kohlberg’s theory provides three levels of moral development known as preconventional morality, conventional morality, and postconventional or principled morality. Each level contains two stages, which are described as follows.

First there is Preconventional morality. Stage one, from birth to 6-8 years of age with overlap, is an authority-punishment—obedience stage or perspective of heteronomous morality or external morality where what is considered to be right is whatever the authority figures say to do, and the reason for doing it is to avoid punishment. Morality is determined by forces, circumstances, and authority outside of oneself. For example, Nazi soldiers at Nuremberg, religion, and medical education all have many external claims.

Also in Preconventional morality, stage two (9-14 years of age) is a self enhancement, and egocentric, instrumental exchange in which what is considered right is whatever meets one’s own needs, but with a sense of fairness in terms of equal exchange between parties in agreement. Basically, it is a “what’s in it for me” attitude, or “I’ll scratch your back if you’ll scratch mine” approach to morality. Many marriages operate at this level, and only 20-25% of adults in society ever get beyond this stage – a frightening but believable notion when you consider how many people operate from a “what’s in it for me” perspective.

Stage three which is the teenage years of group influence orientation moves into the conventional level of mutual interpersonal expectations, peer relationships, and interpersonal conformity where what is considered to be right is what is expected by people close and important to you. “Being good” and living up to expectations of people that are important to you becomes primary in the roles one occupies. Showing good motives, concern for others, and developing relationships of trust, loyalty, respect or gratitude flourish in stage three.

Stage four, which is also conventional morality and develops in the late teens and 20’s, involves a societal maintenance and conscience orientation or a law & order approach in which one fulfills one’s agreed upon duties and contributes to the welfare of the whole group, institution, or society. Right is defined in terms of that which maintains a smoothly running society and avoids the breakdown of the system. It involves upholding the rules as agreed in the U.S. Constitution as best for all in the long run. It develops a “what if everybody did it” approach to morality.

Stage five, which develops in adulthood in the late 20s and beyond, enters into postconventional morality, or principled level moral reasoning, emphasizing individual moral rights such as life and liberty and equality, but endorsing a social contract which protects all people’s rights with a contractual commitment freely entered upon to serve the greatest good for the greatest number. It cultivates values such as those found in the United States Declaration of Independence with the notion that these values ante-date government. So if government violates these values one has a moral obligation to abolish that government and establish one that will respect the moral rights of the people. It is impartial toward the individual. Most values are relative to the group but some are non-relative. It is based upon a rational calculation of the best welfare of all humankind. And only 10% of the population reach this stage.

Lastly, stage six, which is also postconventional morality, is based on a commitment to universal ethical principles of justice, equality, autonomy, and respect for the dignity of all human beings as individual persons. Although laws and social agreements are usually valid because they are based on these principles, when laws violate these principles one acts in accordance with the moral principles. Right is considered to be whatever is required by a personal commitment to these self-chosen universal principles of justice, equality, autonomy, and respect for the dignity of all persons. It is a form of autonomous morality with right and wrong determined from within. Examples of people demonstrating this level of moral development would include Gandhi, M.L. King Jr., Abraham Lincoln, Socrates, Jesus, and Buddha.

A number of interesting claims made by and about the theory have been studied extensively. Indeed now hundreds of Kohlberg studies have been conducted. Generally the research studies have shown that one can understand moral reasoning at one’s own stage and all stages below one’s own stage, sometimes at one stage higher but almost never at two stages higher than one’s own stage. This was demonstrated in an interesting faking study by McGeorge. He asked subjects to take the moral reasoning test twice – once under normal conditions to establish their normal scores and then again with half of them instructed to fake stages lower than their own and the other half instructed to fake stages higher than their own. Attempts at faking lower stages were successful while attempts at faking higher stages were unsuccessful. Even students who took Kohlberg’s course and studied the theory did not improve their scores just by knowing the theory behind it.

However, studies by Mosha Blatt and others have shown that stage change can be stimulated by structured dilemma discussions which
create cognitive dissonance or mental conflict by pitting arguments at one stage against arguments at a different stage. An ethics course was divided into two groups with one half getting the traditional lecture material on the various "isms" or ethical schools of thought while the other half participated in structured dilemma discussions. At the end of the course the traditional group made no significant increases in their moral stage development, while in the dilemma discussion group the majority of the students had increased their stage development by at least 1 stage. Also in many of the studies it has been found that when one sees one stage higher in operation, one tends to be attracted toward it.

It has been claimed and found to be the case that the stages form an invariant upward sequence with no significant regression. This means that everyone begins at birth at stage one and proceeds developmentally although one can stop in any stage and not proceed further. Thus, typically or ideally moral development occurs sequentially. But because development can be stopped at any stage, any of several alternative patterns of moral development are possible. For example, one could develop normally until stage two and then remain there for the rest of one's life. Or development could be normal up to stage four and then stop.

But what one could not have is a pattern of development which involved regressions from once attained higher stages. Each stage presupposes attainment of the previous stage and is necessary for progression to the next stage. Studies have generally shown regression of no more than 6%, which is well within the 10% testing error limits established by test-retest studies under normal conditions.

The validity of Kohlberg's system has been well established cross-culturally and under a wide variety of socio-economic situations in scores of studies in twenty-six cultures in the East and West from both Northern and Southern hemispheres. According to the theory, people proceed through these stages as they mature. The sequence is invariant, although the rate and end stage reached vary with the individual. It is important to understand that only the type of justification provided or the logic of the thinking or reasoning used is considered in assigning a stage score, not a particular set of values or moral beliefs. This distinction between moral reasoning and moral values is often misunderstood and falsely used in criticizing moral reasoning research. It is extremely important to be clear about this distinction. What is being tested is only the person's capacity for moral reasoning, or a process, and not the person's particular set of moral beliefs or moral values, which represents their morality. For example, one could be scored as Stage 4 while holding either conservative or liberal values. Indeed, whether the person holds conservative or liberal values is immaterial to the reasoning capacity for supporting whatever values happen to be held. It is not what one believes but why one believes it that is being analyzed. A given score in no way judges the worth of a person or in no way indicates whether they are a "good person" or not. Also important to keep in mind and be clear about is the distinction between moral reason and moral action. It must be remembered that a low assessment of moral reasoning ability does not mean that such a person is morally bad. Although the relationship between moral reasoning and moral action is complicated and not completely understood, and people with high moral reasoning assessments can indeed do morally reprehensible things in terms of the prevailing social consensus, studies show, however, that people with high level moral reasoning are more likely to act in accordance with their level of moral reasoning. Studies of the classic Milgram experiment on the psychology of inflicting harm to others and subsequent studies of the Berkeley Free Speech Movement involving sit-ins regarding the Vietnam war further demonstrate this relationship.

Now remember Kohlberg's theory is a justice based theory with the principle of justice being considered the highest form of morality. Thus, morality is defined in terms of justice, with the morally right or good being that which is most fair to all concerned based on an interpretation of justice as fairness. From this perspective, then, moral development and moral reasoning are ultimately justice development and justice reasoning. Of course, not everyone accepts or interprets morality in terms of justice. For example, Gilligan, Noddings, and other feminist theorists have argued for morality being interpreted in terms of care, compassion, and responsiveness to other persons. However, most reports of studies in moral development have been based on Kohlberg's cognitive moral development theory.

EMPIRICAL STUDIES

Now with this brief theoretical background, let us look at some of the empirical studies and their results. A good review of the instruments of assessment of moral reasoning and their pros and cons has been published in the Cambridge Quarterly of Healthcare Ethics 1996;5(2):269-77. Several studies assessing the relationship of medical education and moral reasoning have been conducted by my research group, yielding mixed results. These studies used different subject populations, different sample sizes, and different instruments of assessment. The first study, published in Medical Education in 1993 used Kohlberg's original MI, an individual audio taped interview, as the instrument of assessment. However, it is time consuming, labor intensive and costly to score since it requires specialized training to score. Still we pretested and posttested 20 medical students at the beginning of their first year of professional education and again at the end of their fourth year, just before graduation from the professional curriculum, to determine whether their moral reasoning skills had increased to the same extent as expected compared to others persons of comparable age and education. The sample of 20
students represented 41.7% of the student body for that cohort and was not randomly selected. The students were recruited as nonpaid volunteers with appropriate informed consent.

We obtained complete moral reasoning data from all students, along with demographic data including age, gender, Medical College Admission Test (MCAT) scores and GPA scores. For first-year students, the weighted average scores (WAS), or moral maturity indicator, on the MJI ranged from 315 to 482, narrowing to a range of 341 to 454 when they were fourth-year students, with all scores grouped in conventional level of moral reasoning (Stages 3 & 4). The mean increase from first to fourth years of 18.5 WAS (or moral maturity) points was not statistically significant at the P ≤ 0.05 level. Thus, the normally expected increase in moral reasoning skills did not occur over the 4 years of medical education for these students, suggesting that their medical education experience somehow inhibited their moral reasoning growth. The WAS range of 167 points for first-year students represents approximately one and a half moral development stages in Kohlberg’s theory. The WAS range of 113 points for fourth-year students represents only approximately one stage of moral development, indicating a substantial reduction in moral reasoning variance. Perhaps what is happening is the phenomenon of a regression to the mean. No significant correlations between change in moral reasoning WAS scores and age, gender, or MCAT scores were found. Changes in the global stage scores were not significantly correlated with age, gender, MCAT or GPA.

A second study published in the Cambridge Quarterly of Healthcare Ethics in 1996 sought further evidence of the relationship of medical education and moral reasoning, and used the Sociomoral Reflection Measure (SRM) of Gibbs and Widaman. Unlike the MJI, which involves an audio taped interview that is expensive to score, the SRM, being group administrable and relatively inexpensive to score, enabled the study to incorporate a larger sample size. This study pretested and posttested 30 medical students at the beginning and end of their medical education to determine what, if any, changes occurred in their moral reasoning skills. The sample of 30 students represented 62.5% of the students in that class. Again, the students were not randomly selected and were recruited as nonpaid volunteers with appropriate informed consent.

The mean scores in first-year and fourth-year students were 359.10 and 369.77 respectively. The mean increase of 10.67 points in the SRM score from first-year to fourth-year was not statistically significant at the P ≤ 0.05 level. This supports the findings of the previous study which also showed no growth. Again the demographic factors of age, gender, MCAT and GPA did not correlate with moral reasoning scores.

A third study published in Academic Medicine in 1998 assessed the relationship of medical education and moral reasoning using the Defining Issues Test of Rest or (DIT) and a still larger sample size. The DIT is a paper and pencil computerized test based on Kohlberg stage theory. It provides 6 dilemmas with 12 options which are rank ordered. This third study pretested and posttested 95 medical students before and after their professional education. This sample represented students in four first-year classes. The students were tested at the beginning of their first semester and again at the end of their first semester following completion of a course in medical ethics and again at the end of their fourth year. The students were not randomly selected, but were recruited as nonpaid volunteers with informed consent.

The mean scores at the beginning of first-year was 47.7. The mean scores at the end of the first semester was 53.7, and the mean scores at the end of their fourth year was 56.5. The mean increase of 6.0 points from the beginning of first-year to the end of their first semester was statistically significant (p ≤ 0.0001) as was the +8.8 change in mean scores from the first test to the end of the fourth-year (p ≤ 0.001). The change from the end of first semester to the end of fourth-year represented only +2.8 points but was marginally significant (p ≤ 0.03). Statistical analysis revealed no significant correlation at the P ≤ 0.05 level between moral reasoning scores and age, but there was significant correlation between the moral reasoning scores and gender. The gender differences in this third study were significant with females scoring higher. I will have more to say about gender differences later.

This third study, with a larger sample and the DIT, disagreed with the findings of the first study with the MJI, and the second study with the SRM suggesting that the experience of medical education may not inhibit the increase in moral reasoning of medical students. However, the third study included an educational intervention of small group case study discussion that the first and second studies did not include. We believe that this intervention accounts for the different findings. All three studies found a narrowing of the range of moral reasoning scores as students progressed from the first year to the fourth year of their education. This regression to the mean suggests a powerful socializing influence in professional education, which results in a homogenizing effect. Higher scores coming down and lower scores coming up suggest a suppression of uniqueness, creativity, and individuality that may not necessarily be bad but should at least give one pause to think in light of the current widespread emphasis on diversity. Because all these studies were conducted at the same institution, similar studies need to be conducted in other settings. Clearly, additional research is needed to clarify fully the relationship of medical education and moral development. However, the same three types of studies have also been done with veterinary medical students at our school with essentially the same results.
AN INTERVENTION STUDY IN MEDICAL ETHICS

An intervention study was designed to assess whether a one semester course of 44 contact hours devoted exclusively to medical ethics would significantly improve the moral reasoning skills of students. It was published in Academic Medicine in 1989. The students were pretested and posttested with the SRM. The course was taught in the first semester of the first year of the medical curriculum as a required course.

The study divided students into three groups. One group of 37 students received lectures in medical ethics. Another group of 46 students received exposure to both lectures and small group case study discussions. The third group of 36 students served as a control group that received neither lectures nor small group case study discussions. Complete data were gathered on all 119 first-year students. The pretest SRM scores were 349.27 for the lecture group, 352.93 for the case study group, and 344.33 for the control group. The posttest SRM scores were 346.05 for the lecture, 356.83 for the case study group, and 323.92 for the control group. Statistical analysis revealed that both the lecture and the case study groups improved significantly over the control group and that the case study group increased their scores even more than the lecture group.

These data suggest that a one semester course of 44 contact hours of medical ethics of this format, especially the small group case study discussions, is sufficient for improving the moral reasoning skills of medical students. Clearly, additional studies need to be done considering both course length and format, such as using films, for designing a medical ethics course that will have the desired effect, yet still be practical for an already overcrowded curriculum.

To further clarify the curriculum intervention needs another study published in Academic Medicine in 1998 was conducted to examine how such exposure to small-group case-study discussion is necessary to significantly increase moral reasoning skills. Using Rest’s Defining Issues Test, groups of students were tested for their moral reasoning skills both before and after the students participated in small-group case-study discussions of medical ethics. From 960 students asked to participate, complete data were collected for 729 students (75.9% response rate). Small-group case-study exposures ranged from 0 to 44 hours. Groups of students exposed to 20 hours or more demonstrated a significant increase in their moral reasoning scores. Groups with less than 20 hours of exposure demonstrated no significant increase in their scores. This study indicates that moral reasoning skills are teachable and measurable, and that small-group discussion significantly increases moral reasoning skills. Further studies are needed to increase the generalizability of these findings.

A longitudinal retention study was conducted to assess the long-term effects of teaching medical ethics courses. The hypothesis was that the increased moral reasoning skills resulting from teaching a medical ethics course in the first year would be retained over the four years of medical school. This hypothesis is consistent with cognitive moral development theory. Since there has been increased interest in moral reasoning of medical students but no longitudinal studies, this study was conducted to demonstrate whether moral reasoning skills retention occurred in subsequent years. The significance of this study was three-fold: (1) The importance of teaching medical ethics early in the curriculum, (2) The importance of offering enough exposure to significantly increase moral reasoning skills, (3) The importance of medical ethics being a required course of all medical students, not just an elective as is often the case. The Defining Issues Test was used for assessment of moral reasoning. First-year students were pretested, taught a medical ethics course during the 1st semester, and posttested at the end of the 4 subsequent years. After documenting the significantly increased, p<0.0025, moral reasoning skills, this study demonstrates retention of these moral reasoning skills at the end of 2nd year, p<0.00001; 3rd year, p<0.00001; and 4th year of medical education, p<0.00001. This retention study demonstrates that the teaching of medical ethics can be rigorously measured and tested and that it can have a sustained positive influence on the moral reasoning skills of medical students.

GENDER AND MORAL REASONING

The influence of gender has been a controversial issue in the moral development literature. Because the original sample from which Kohlberg developed the theory was an all male sample, Carol Gilligan claimed that cognitive moral development theory has a built in male bias that systematically discriminates against females. Others have carried on a lively debate in the literature regarding this point. But studies in both human medicine and veterinary medicine frequently have found the contrary to be true, that females score significantly higher than males in moral reasoning. A similar finding was reported for social workers by Dobbin in 1989. Nevertheless, several other studies of health care professionals did not find a significant relationship between gender and moral reasoning, although many of their sample sizes were small.

A study of 705 medical and veterinary medical students published in the Journal of Critical Care in 1993 convincingly demonstrated that females score statistically higher than their male classmates. The hypothesis of this study was that there is no significant difference in the moral reasoning skills of medical students with regard to gender. This hypothesis is in keeping with the theory of cognitive moral development. For this study the DIT was used to measure moral reasoning skills. Of the 705 students initially enrolled in this study, 79 were dropped when they failed the consistency and reliability check leaving a total of 626 students to be analyzed. The mean DIT score for the 312 males was 41.77
and for the 314 females, 47.18, highly statistically significant at the .0001 level. This finding appears to be in conflict with the common criticism of moral development theory. The DIT is derived from cognitive moral development theory, which, in turn, is based on the principle of justice as the foundation of moral reasoning. This study suggests that females are not less, and probably more, effective in the use of justice for resolving moral conflicts.

Many other empirical studies regarding moral reasoning and moral orientation have been conducted by my research group. The studies include assessment of other professional groups such as veterinarians, nurses, chaplains, law students, engineering students, as well as other groups. The studies include cross-sectional, longitudinal, and intervention research designs. The studies also include the relationship of moral reasoning to many issues such as clinical performance, specialty choice, personality types, self-esteem, empathy, and so forth. Many of these studies have already been published, while many others of them are in various stages still awaiting completion.

Endereço para correspondência:
Department of Humanities in Medicine
College of Medicine
The Texas A&M University
System Health Science Center
164 Reynolds Medical Building
College Station, Texas 77843-1114
Telephone: 409-845-0755
Email: dself@tamu.edu