

AN ICM APPROACH TO THE ASSESSMENT OF A MEDICAL ETHICS  
INTERVENTION IN THAILAND

by

PONSOPA PINIJPHON

A DISSERTATION

Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in the  
Department of Educational Psychology  
in the Graduate School of  
The University of Alabama

TUSCALOOSA, ALABAMA

2009

Copyright Ponsopa Pinijphon 2009  
ALL RIGHTS RESERVED

## **ABSTRACT**

The purpose of this study was to develop and test a new instrument, the Medical Intermediate Concept Measure of Ethical Reasoning (Hereafter the MD-ICM), designed to help educators assess the ethical reasoning of medical students in Thailand. Following earlier methods for developing an ICM measure (e.g., Bebeau & Thoma, 1999), the dissertation first presents the development of the MD-ICM . Specifically, there were four steps used to create the measure: Step 1: identify appropriate ethical dilemmas, Step2; construction of the Items used to assess choice and justification information, Step 3: obtaining expert evaluation of the ethical dilemmas and assessment items items. and Step 4: Testing for reliability and validity. Reliability and Validity of the measure was assessed in a sample of 627 Thai Medical Students representing seven different universities. These students were in the first three years of medical training.

Quantitative analysis (e.g., descriptive statistics, simple correlation, ANOVA, and regression analysis) was used to answer four research questions. The findings from this study are; 1) the MD-ICM was able to distinguish medical students who differ in medical ethics instruction and experience in the medical profession. Superficially, third year students scored higher than second year and first year students respectively. Further females scored higher on the MD-ICM across years and schools. 2) MD-ICM scores differed between students experiencing two types of professional ethics instruction. Students experiencing a specific course in medical ethics scored higher than students from schools who have adopted an integrated program. 3MD-

ICM scores relate to established measures of moral judgment development supporting the convergent validity of the measure. Further, the results of these analysis indicates that the MD-ICM offers information that is not redundant with these established measures., and 4) the relationship between MD-ICM scores and supporting measures did not vary by gender. Results are discussed in terms of the applicability of the MD-ICM approach to the assessment of professional ethics education.

## LIST OF ABBREVIATIONS AND SYMBOLS

ANOVA	Analysis of variance
CPIRD	The Collaborative Project to Increase Production of Rural Doctors
<i>df</i>	Degrees of freedom: number of values free to vary after certain restrictions have been placed on the data
DIT	Defining Issues Test of moral judgment development
<i>F</i>	Fisher's <i>F</i> ratio: A ration of two variances
ICM	Intermediate Concepts Measure
<i>M</i>	Mean: the sum of a set of measurements divided by the number of measurements in the set
<i>n</i>	Number of members in a limited portion of the total sample
<i>N</i>	Number of members in a total sample
N2SCORE	An index that assesses the prioritizing of the higher stages and discrimination and rejection of the lower stages. It is the primary index supplied by the DIT
<i>p</i>	Probability associated with the occurrence under the null hypothesis of a value as extreme as or more extreme than the observed value
<i>P</i> -score	Post-conventional score—an often used DIT score
<i>r</i>	Pearson product-moment correlation
<i>R</i> <sup>2</sup>	Measure of strength of relationship
<i>SD</i>	Standard Deviation
STAGE4P	A DIT index that is similar to Kohlberg's description of a law and order perspective (e.g., Stage 4).

<	Less than
=	Equal to
%	Percentage

## **ACKNOWLEDGEMENTS**

It is impossible to complete a Ph.D. degree at the University of Alabama without my endeavors, and supports from many.

First of all, I would like to thank to Dr. Stephen Thoma for being my advisor, mentor, and chairperson of this dissertation throughout my years at the University of Alabama. Dr. Thoma who has been incredible supportive, wonderful, understandable student's life, be constantly willing to help, and be share his knowledge. I would also like to thank all my other committee members, Drs. Allen Wilcoxon, Ali Iran-Nejad, Judy Giesen, and Nick Stinnett, for their tremendous support, inspiring input, and accessible demeanor.

I want to thank my family in Thailand, my brother, Nopadon and my sister -in- law, who help me to take care of my son, Charles Malone while I was doing data collection process. To my sister, Tusanee, for her love and support me for everything while I am studying in the U.S.

Thanks to my late father, Pun Piniyphon who instilled in me the importance of education and let me as his daughter, pursue higher education. To my mother, Sangnin, for her love and support.

## CONTENTS

ABSTRACT.....	ii
LIST OF ABBREVIATIONS AND SYMBOLS .....	iv
ACKNOWLEDGEMENTS .....	vi
LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
I. INTRODUCTION .....	1
Statement of the Problem.....	2
Purpose of the Study .....	3
Overview of the Methodology.....	4
Research Questions .....	4
Assumptions.....	5
Limitations of the Study.....	6
Definition of Terms.....	8
II. REVIEW OF THE LITERATURE.....	10
Ethical Dilemmas for Physicians: The Moral Aspect of Medical Practice .....	10
The Moral Field: Moral Psychology and Theory .....	13
Three Branches: Psychoanalysis, Learning Theory, and Social-Cogitative Development .....	14
Piaget: Foundational Work With Moral Development in Children.....	16
Kohlbert and the MJI; the Major Research in Moral Judgment .....	17

Rest: Broader Models of Moral Functioning .....	20
Expansion and Consolidation of the Field .....	22
Ethical Reasoning and Testing in Medicine .....	22
The Call for a New Test for Medical School Education.....	22
Link Between Moral Reasoning and Practice.....	23
Medical School’s Influence on Moral Development .....	25
Some Ethics Interventions Help.....	27
Problems With a “Hidden Curriculum” .....	29
Suggestions to Fix the Problems of the “Hidden Curriculum” .....	34
Other Factors That Could Inhibit Moral Growth.....	35
Problems with Existing Tests .....	37
Moral Judgment Interview .....	37
DIT and Other Moral Reasoning Tests.....	38
Other Types of Measures of Morality.....	40
Intermediate Concepts: A Framework for Building New Ethics Tests .....	42
Considerations for Testing in Thailand.....	47
Thai Officials Ask for the Test to be “Made in USA” .....	49
III. METHODOLOGY .....	52
Developing of the MD-ICM .....	52
Participants.....	55
Materials .....	57
Data Collection Procedures.....	59
Data Analysis.....	60

Research Question 1 .....	60
Research Question 2 .....	61
Research Question 3 .....	61
Research Questions 4.....	61
Summary of Methodology .....	62
IV. RESULTS .....	63
Research Question 1 .....	67
Research Question 2 .....	73
Research Question 3 .....	78
Research Question 4 .....	80
Total Ranking Index .....	81
Good Ranking Index .....	82
Bad Ranking Index .....	83
Action Choice Index .....	84
Justification Index.....	85
V. DISCUSSION .....	89
Implications for Medical Ethics Education and Assessment.....	97
Implications for ICM Development.....	99
Implications for Future Research Using the MD-ICM .....	100
Limitations of the Study.....	102
REFERENCES .....	104
APPENDIX A.....	112
APPENDIX B.....	115

APPENDIX C .....	135
APPENDIX D .....	139

## LIST OF TABLES

1. A number of Medical Students, University and Medical Center Affiliated with University.....	57
2. Demographic Information by Sample.....	64
3. Demographic Information by Sample Described by Age.....	65
4. Descriptive Statistics of ICM-Scores as Functions of Gender and Educational Level .....	66
5. Descriptive Statistics.....	79
6. Simple Correlation ( $r$ ) between the MD-ICM and the DIT2.....	80
7. Summary of Hierarchical Regression Analysis for Variables and Interactions ..... Predicting MD-ICM Score ( $N = 540$ ) .....	86

## LIST OF FIGURES

1. Total Ranking Index .....	70
2. Good Ranking Index .....	71
3. Bad Ranking Index .....	71
4. Action Choice Index .....	72
5. Justification Choice Index.....	72

## **CHAPTER I**

### **INTRODUCTION**

Physicians consistently rank as the most trusted of professionals and high scores by doctors on abstract measures of moral judgment (e.g. MJI, DIT, SRM, etc.) suggest that this trust is well placed. This is not to imply that decision making on healthcare issues is easy for physicians. The day-to-day demands placed on physicians to make decisions about medical issues are frequently complicated by the presence of ethical dilemmas. Medical ethics curriculum, now mandatory in many medical schools worldwide, should help future physicians to identify and to logically work through ethically challenging dilemmas but there is no consensus on curricular content, time allotted for study or teaching methods (Lehmann, Kasoff, Koch, & Fedeman, 2004; Loewy, 2006; Goold & Stern, 2006; Russel & O’Neill, 2006). Strangely, the medical school experience itself may not promote normal growth in moral reasoning skills (Self, Schrader, Baldwin, & Wolinsky, 1993; Patenaude, Niyonsenga, & Fafard, 2003) and the formal ethics curriculum may not address the more obvious needs of students, much less an informal or “hidden” curriculum that inhibits moral growth (Christakis & Feudtner, 1993; Charon & Fox, 1995). For all of the uncertainties in ethics interventions, only half of medical school deans surveyed recently in the U.S. and Canada reported formal assessments of student’s moral reasoning skills (e.g. Rest’s Defining Issues Test) and only one-third formally assessed students’ ability to handle ethically challenging situations, even for less complicated tasks such as giving bad news to patients (Lehmann, Kasoff, Koch, & Fedeman, 2004).

McAlpine, Kristjanson, and Poroch, (1997) and other researchers have identified the lack of instruments specific to medicine as a problem in assessing ethics programs.

### ***Statement of Problem***

Existing moral (ethical) reasoning tests such as Kohlberg's (1984) Moral Judgment Interview (MJI) and Rest's (1979) Defining Issues Test (DIT) measure the most abstract level of moral thinking, the "bedrock concepts" that, theoretically, function as default schema to guide individual's actions (Rest and Narvaez, 1994). These types of instruments are simply too course grained to provide an overall assessment of the specific curricular needs of students and ethics programs. Even less informative are instruments based on knowledge of the concrete codes of ethics of professions (e.g. the AMA code of Ethics). Such tests indicate little about thinking on ethical issues.

The new instrument, the MD-ICM, is designed at the "intermediate concepts" level (Bebeau & Thoma, 1999) around the principles reflected in the content of professional ethics curriculum in areas such as informed consent, confidentiality and so forth. Based on the work of Bebeau and Thoma (1999), this intermediate concepts measure (ICM) uses real cases specific to the medical field and relies on the thinking and the judgment of experts in the medical field for the design of items on the test and for the development of a scoring key. This design is flexible enough to address an informal curriculum in medical school that influences the moral decision making of medical students. Face validity is ensured by the design, encouraging educators to use the new instrument. This medicine specific ethical assessment tool should help educators to identify deficits in individual students' ethics skills and, more importantly, to assess the efficacy of existing ethics interventions.

Additionally, the government of Thailand has expressed interest in an ethical reasoning measure in medicine in order to look for relationships between ethics reasoning and medical

entrance and exit exam board scores for students and to examine the relationship between ethics reasoning and clinical performance evaluated near the end of medical school training.

At the present time, ICM's have not been studied in cultures outside of the United States but it should be possible to apply and validate "intermediate concepts" in cross-cultural studies. This researcher is interested in conducting an ICM in medicine in Thailand for students in the Collaborative Project to Increase Production of Rural Doctors (CPIRD) program. (The word "production," as used here, means "quantity"). The CPIRD recruits students from rural areas to become medical doctors and to return to rural areas to practice medicine. The CPIRD program began in 1995. It is administered by the Thai Department of Public Health.

Importantly, Thailand claims to follow the U.S. /U.K. medical models and some of the Thai medical text books are in English so the MD-ICM, will be written in and administered to the CPIRD students in English.

### ***Purpose of the Study***

The purpose of this study is to develop and to test a new instrument, the MD-ICM, an intermediate concepts measure (ICM) based on the work of Bebeau and Thoma (1999), designed to help educators assess the ethical reasoning of medical students. The results of this study can be beneficial to Thai medical students, faculty, physicians, and medical educators who work for the Collaborative Project to Increase Production of Rural Doctors (CPIRD) program that is administered by the Thai Department of Public Health. The MD-ICM can help Thai medical faculty to know how to construct valid and reliable tests to administer medical students. Further, this study could help to provide moral education programs in Thailand, especially at Chonburi Medical Center. This instrument allows for testing more specific within medical field, which will aid Thai medical students in becoming more professionally and morally functional as health care providers.

## ***Overview of the Methodology***

The purpose of this study is to develop and test a new instrument, the MD-ICM, designed to help educators assess the ethical reasoning of Thai medical students. The initial step is developing of the MD-ICM. There were three steps, Step 1: Development of Items, Step 2: Experts (i.e., medical ethicist and faculty) and Step 3: Testing for reliability. The instruments for this study utilized the MD-ICM and the DIT2. The MD-ICM score was represented by five main indices (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index). For the DIT2 was calculated by Post - conventional score (Pscore), N2SCORE, and Maintain Norms (STAGE4P). The samples were Thai Medical Students in seven different universities in different years (Year I, Year II and Year III). Quantitative analysis (e.g., descriptive statistics, simple correlation, ANOVA, and regression analysis) was used to answer for research question 1, 2, 3, and 4.

## ***Research Questions***

Primarily, for CPIRD students in Thailand, there are four general research questions:

1. Is the MD-ICM sensitive to ethics education?
  - a. Are pre-test and post-test MD-ICM scores different for the 2<sup>nd</sup> year ethics course?
  - b. Are MD-ICM scores different for sophomores and seniors in the 6 year integrated ethics curriculum at the other Universities?
2. Is there evidence that the MD-ICM assesses the moral domain?
  - a. Does the MD-ICM relate to established measures of moral thinking such as the DIT as represented by the N2 and U scores and type variable?
3. Do 6<sup>th</sup> year students in the new integrated ethics program score higher on the MD-ICM than 6<sup>th</sup> year students in the old style one-class curriculum (SHHU 111)?

4. Is the MD-ICM related to external measures of medical school success?
  - a. Does the MD-ICM relate to clinical practice scores beyond traditional grades?

### ***Assumptions***

1. All medical students who participated and completed the MD-ICM and the DIT2 understood the instruments and had the necessary knowledge to complete the instruments. It was further assumed that the participants' English proficiency were the same. Because of this, participants could not claim differences in proficiency as a bias in this study.

2. It is assumed that all medical students responded to the instruments by themselves with truthfulness and accuracy. They were not allowed to discuss responses with their peers while taking the test.

3. The instruments were valid and reliable for recording data.

### ***Limitations of the Study***

1. This study is designed to develop and to test a new instrument, the MD-ICM, to help educators assess the ethical reasoning of Thai medical students. There were eight university populations recruited for this study, therefore the results are limited to application in the context of Thai universities. It would be difficult to generalize the results of this study to other populations.

2. Another limitation encountered while collecting these data in Thailand was due to the unexpected political crisis which coincided with the data collection. The author had difficulty safely gaining access to Thai Medical Students because the government buildings including universities were shut down due to protests. Sample sizes were too small for some subgroup within the overall population to analyze differences between groups. Therefore, all research questions could not be completed for this study.

There are several limitations that apply to this particular study population that may have contributed to low scores on the DIT2 and the MD-ICM. The primary limitation in this study was the poor English Language skills by Thai students.

3. Because this study utilized English language for the MD-ICM and the DIT2, this could introduce unforeseen limitations due to English not being the students' primary language. One possible explanation was that it was not possible to administer these tests in Thai because the true meaning of the test might get lost in translation. This is probable even though Thai medical students are proven through recruiting statistics to be the students who earned the highest GPAs in high school. It was thus assumed that their English skills were better than that of other Thai student groups. Generally speaking, however, Thai medical students might have had a problem while taking the tests in English because Thailand is not an English speaking country. Although English is sometimes used in text books, Thai medical education is conducted primarily in the Thai language. Consequently, it might be difficult for Thai medical students to complete the MD-ICM and the DIT2 in English. Some Thai medical students complained because the test was not given in the Thai language.

4. More importantly, there were possible limitations in this study pertaining to how the experimental procedures themselves were conducted. The variable subjects' reaction times may have been reduced, therefore minimizing test scores of the MD-ICM and the DIT2 that might have been higher under more optimal experimental condition. One major distraction for the students might have been that their final examinations were scheduled for the next week. Medical students are very busy. Further distractions that might have lowered student scores include general room conditions and noise generated within a group context. Many students simply were not enthusiastic about taking the test.

5. Cultural differences might have caused a problem with the administration of MD-ICM and the DIT2. Even though the DIT2 was the standardized test and has been widely used, it may hindered the ability of (Thai medical students) to chose the best and worst action choices and justifications based upon Thai culture context. It may be possible that Thai medical students may not entirely understand certain ethical dilemmas initiated in western culture, because they may not have seen such dilemmas in Thai context. For the MD-ICM test, certain types of medical dilemmas might not have been applied to Thai medical context.

6. Although Thai medical students are assured of confidentiality, doubts could have influence responses for this study. Some medical students may have intentionally failed to complete the tests so that the results would be thrown out.

After examining the possible limitations of this study, four specific research questions emerged:

- a. Are MD-ICM scores different for the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> year integrated ethics curriculum?
- b. Do medical students in the integrated ethics program score higher on the MD-ICM than students in the old style one-class curriculum?
- c. Does the MD-ICM relate to established measures of moral thinking such as the DIT as represented by the N2 and U scores and type variable?
- d. Does the MD-ICM score relate to GPA and Gender variable?

## ***Definition of Terms***

These definitions of terms for this study is utilized for better understand this dissertation in particular some technical terms that might have not seen in general. The following definitions were used.

*Total Ranking Index:* A Total Ranking index is represented by the mean percentages of *good and bad choice and justification choices across stories* that agree with judgments derived from the responses of six experts in the filed of medical ethics education in the medical profession.

*Good Ranking Index:* A Good Ranking Index is represented by the mean percentages of *good action choice and justification choices across stories* that agree with judgments derived from the responses of six experts in the filed of medical ethics education in the medical profession.

*Bad Ranking Index:* A Bad Ranking Index is represented by the mean percentages of *bad action choice and justification choices across stories* that agree with judgments derived from the responses of six experts in the filed of medical ethics education in the medical profession.

*Action Choice Index:* An Action Choice Index is represented by the mean percentages of *action choice ranking both good and bad* that agree with judgments derived from the responses of six experts in the filed of medical ethics education in the medical profession.

*Justification Index:* An Action Choice Index is represented by the mean percentages of *justification choice ranking both good and bad* that agree with judgments derived from the responses of six experts in the filed of medical ethics education in the medical profession.

*MD-ICM:* A Medical doctor intermediated concept measures that used to measure ethical reasoning ability with medical students in Thailand. Medical dilemma cases were based primarily on the concepts such as patient autonomy, confidentiality, beneficence vs. patient

autonomy, non-maleficence (assisted suicide), and beneficence. A total score is identified by five main indices (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index, and Justification Index).

*P-score*: A score that measures Post-conventional score in Kohlberg's stage five and six. The P-score indicates respondent's preference for utilizing a postconventional moral schema.

*N2SCORE*: A N2 index is an index that assesses the prioritizing of the higher stages and discrimination and rejection of the lower stages.

*STAGE4P*: A maintaining norms score that measure in Kohlberg's stage four.

*CPIRD*: The Collaborative Project to Increase Production of Rural Doctors that is administered by the Minister of Public Health in Thailand. The CPIRD recruits students from rural areas to become medical doctors and to return to rural areas to practice medicine. The CPIRD program began in 1995.

## CHAPTER II

### REVIEW OF LITERATURE

#### *Ethical Dilemmas for Physicians: The Moral Aspect of Medical Practice*

The concern with the ethical behavior of physicians can be traced back to the Hippocratic Oath (400 B.C.) and is found in modern codes of ethics by the American Medical Association (AMA, 1985) and others. Self and Baldwin (1994) argue that doctors and other health care providers face ethical dilemmas in the daily practice of their profession. In fact, ethical considerations begin quite early in the doctor-patient relationship. This idea is aptly illustrated in an article entitled, "Every consultation has an ethical component," published in the British Medical Journal (Bradley, 1983), in which doctors discuss the ethical considerations involved, not with patient care but, with accepting or refusing to accept a new patient.

A review of literature finds several books filled with real-life based stories involving ethical dilemmas in medical practice. Beauchamp and Childress's (2001) excellent "*Principles of Biomedical Ethics*," Veatch's (1997) "Medical Ethics," provide theory and discussion together with dilemma cases. The authors examine the philosophies of Plato, Aristotle, Hobbes, Locke, Rousseau, Hume, Mills, Kant and Rawls and others and theories such as deontology and utilitarianism for their relevance to ethical medical practices. The seeds of ethical dilemmas for physicians are evident in these writings as all of the theories and the philosophies have applications in medical practice and yet fail to guide doctors effectively in their day to day decision making. For example, on the issue of lying to patients, Roger Higgs (1985) notes that Plato says that it is okay to lie if the lie benefits the patient whereas Kant says to never lie no

matter what the outcome is for the patient and Hippocrates says to beware... of telling the truth to patients. Lockwood (1985, p. 7) sums it up nicely, saying, "A common complaint against philosophers is that they are good at raising questions, but not so good at solving them." He adds, however that the philosophers generally arrive at the right answer. Noteworthy, the injunction to be truthful to patients did not appear in the AMA Code of Ethics until 1980.

In the modern era, Childress (1997) finds many moral dilemmas in medicine to be generated by the conflict between moral principles such as beneficence (doing good), nonmaleficence (not doing harm), justice and respect for autonomy. Determining which factors apply to a particular case and, more importantly, what weight to assign to each factor is commonly called "balancing." It often proves to be a difficult task. To give some examples taken from multiple sources involving "end of life" care, a physician who removes a feeding tube from a patient who suffers irreversible brain death to all but the lower brain stem could be charged with murder if a nurse decides that the deontological interest of this one patient (doing what is right for this patient without respect for other considerations) is not being served and she subsequently informs the local District Attorney of the physician's action. The physician, however, could legitimately make the utilitarian argument that the hospital needed the resources for other "salvageable" patients. Yet, a physician may be asked by a reasonable, sane patient to withdraw a feeding tube which keeps the patient alive. While this procedure could be seen as "assisted suicide" by some persons, to refuse to honor the request could be seen as an unethical violation of the patient's autonomy. On the other hand, the physician who wants to insert a feeding tube directly into the stomach of a "brain dead" patient in order to keep the patient alive could have the procedure challenged by the patient's family on the grounds that the surgical insertion of the tube would be "invasive" and a gross violation of the patient's autonomy.

Arguably, one of the most common types of dilemmas for physicians stems from “paternalism,” or the assumption that the physician should make decisions about a patient’s care because he is the person most qualified to do so (Lockwood, 1985; Beauchamp & Childress, 2001; Veatch, 1997). Many patients expect “the Doctor” to make most decisions about the patient’s care. That is, patients have the right to choose to let the physician make their health care decisions. Yet, the physician has to respect the patient’s right to make their own decisions (Beauchamp and Childress, 2001; Veatch, 1997), including the right to refuse needed treatment (as seen from the physician’s perspective). A patient who defers decisions to the physician can choose to assert their rights at any time. Quite obviously, some patients are not competent to make their own decisions and the physician makes that determination (usually). “Advance Directives” are sometimes drawn up to prepare for some future event in which a person finds himself unable to communicate his wishes about medical care. These directives don’t necessarily make decisions any easier for physicians because the directives are often vaguely worded and unusable or in conflict with the local or state laws or with accepted practice or hospital guidelines (Beauchamp & Childress, 2001). In this case, or with no clear documentation of the patient’s wishes, the physician turns to the patient’s family for direction. It is possible and often happens that the family members will disagree about the health care choices in which case the physician will be wrong in any decision that he makes (or does not make).

Can the physician rely on past decisions to guide him in current dilemmas? Not always. Self and Baldwin (1994) note that the circumstances of medicine are constantly shifting and unpredictable. This creates new dilemmas and complicates the task of dealing with ethical situations that are already familiar to physicians. For example, the dissemination of patients’ information by health care providers has been greatly restricted by codes as old as the Hippocratic Oath and as recent as the U.S. HIPAA Privacy Rule (45 CFR Parts 160 and 164;

1996) that entreat doctors to maintain the confidentiality of their clients. However, with the passage by the U.S. Congress of The Ryan White Act of 1996 and subsequent amendments, hospitals may order tests for HIV, Hepatitis B and other blood borne pathogens and selectively disseminate the test results without regard for the HIPPA laws and the basic respect for a patient's right to refuse unnecessary blood screening, when healthcare providers have been exposed to a patient's blood or other body fluids. While this example may seem to be only a simple case about the law and not to present an actual ethical dilemma, consider that a doctor could find himself legally constrained from informing a nurse that she was exposed to AIDS/Hepatitis by contact with an infected patient if the nurse was late in filing the required patient exposure form - even if other nurses had already been informed of their own exposure (HIPAA Privacy Rule, 1996). There are many similar examples.

Note that these examples of ethical dilemmas represent but the "tip of the iceberg" for physicians.

In sum, medical doctors and other health care providers face ethical dilemmas in the daily practice of their profession, and the circumstances of medicine are constantly shifting and unpredictable (Self and Baldwin, 1994). Nonetheless, ethical considerations begin quite early in the doctor-patient relationship (Self and Baldwin, 1994). Medical dilemmas in medicine, for instance, are conflicts between moral principles such as beneficence (doing good), nonmaleficence (not doing harm), justice and respect for autonomy.

### ***The Moral Field: Moral Psychology and Theory***

In the preface to "*Postconventional Moral Thinking*," James Rest and others (1999) note the importance of Lawrence Kohlberg to the field of moral psychology. Lapsley calls "Stage and Sequence" (Kohlberg & Keamer, 1969) and "From Is to Ought" (Kohlberg, 1971) "the twin

pillars... of the cognitive-developmental approach to morality and socialization” (Lapsley, 2006, p. 37).

### ***Three Branches: Psychoanalysis, Learning Theory and Social-Cognitive Development***

Philosophers, theologians and others have tried to define morality. Gert (2005) says that morality is difficult to define; nothing that you say after, “‘ Morality is ...’ seems to be quite right” (p. 3). He defines morality as, “an informal public system applying to all rational persons, governing behavior that affects others, and includes what are commonly known as the moral rules, ideals, and virtues and has the lessening of evil or harm as its goal” (p. 27). Ethical relativists tend to emphasize cultural-norms based definitions of morality that may lead to immoral behavior by this definition. Shaffer (1999) says that college students generally agree that morality involves the ability to distinguish right from wrong, to act on this distinction, and to feel good or bad according to whether or not we violate our own standards.

In defining and studying moral affects, psychoanalysts tend to emphasize strong affective emotions (e.g. guilt, shame, pride). Cognitive-developmental theorists (e.g., Kohlberg) emphasize cognitive aspects of moral reasoning. Social information-processing and social learning theorists study moral behavior to see how children learn to resist temptations to violate moral rules (e.g. lying, cheating and stealing). In short, research and theories of moral development are based in three components that represent psychoanalysis, learning theory, and cognitive-development theory; they are 1) an affective or emotional component, 2) a behavioral component and 3) a cognitive component.

Freud theorized that children resolve an “Oedipal complex” by patterning themselves after their same sex parent and internalizing that parent’s moral standards enroute to the development of the “superego,” or conscience, sometime between the ages of 3 and 6. Freud felt that strong emotions (e.g. pride, fear and shame) were strong motivators for conduct and that

harsh parental discipline was helpful in establishing a child's identity. However, modern attachment theory tells us that a close, continuous and mutually satisfying relationship between a parent and child leads to secure attachment that is necessary for a child to thrive emotionally (Bretherton, 1992). Modern psychoanalysts believe that securely attached toddlers begin to develop a conscience long before a child reaches the age for an Oedipal morality to be experienced (Shaffer, 1999).

While there is not much research by psychoanalysts on moral reasoning or moral development on children of elementary age or beyond, the ideas of moral emotions and internalizations of moral principles expounded by Freud and the psychological construct of self are incorporated in many current and more integrated theories of morality that extend into adulthood. In the Self Model, Blasi (2004), for example, integrates moral understanding implicit in the moral cognition of Kohlberg and moral personality. Moral functioning is seen as comprised of three components: 1) moral self-identity, 2) a sense of personal responsibility that connects judgment and action and 3) self-consistency or integrity needed for us to appear moral to ourselves and to others (Walker, 2004). Moral integration leads to moral motivation. Moral motives can conflict with personal motives, leading to moral emotions (e.g. shame, guilt). Moral emotions are a sign of the healthy integration of our moral understanding with our affective system (Lapsley, 1996). Blasi (1995; as cited in Walker, 2004) also found that the need to maintain our sense of moral integrity could lead to self-deception. Walker (2004) strongly advocates for more research in the role of self in moral functioning. The concept of self figures prominently in thinking on and research in human development.

Albert Bandura has suggested that a breakdown in moral self-regulation explains why people selectively disengage from moral self-sanctions and has conducted much research in this area (Zimmerman and Schunk, 2003). Bandura (2002) says that self-regulatory mechanisms

have to be activated in order [for them] to operate and that moral knowledge and reasoning must be linked with moral conduct in a complete theory of moral agency. Social cognitive theory finds moral action to be the product “of the reciprocal interplay of personal and social influences” (2002, p. 101). Bandura, Mischel, and other social learning theorists find moral behavior to be learned through reinforcement, punishment and social modeling and find moral behavior, moral affect and moral reasoning to be specific to the situational context (doctrine of specificity) rather than being based on internalized moral principles. Perhaps this “doctrine” is overstated. Learning theorists study the ability of individuals to resist the temptation to lie, cheat or otherwise violate society’s norms when detection and punishment are unlikely (Shaffer, 1999). Given the lack of emphasis placed on internalization of moral principles, this line of study may seem paradoxical. However, learning theorists are interested in the internal motivation to obey rules. Again, much emphasis can be placed on the development of “self.” Also, the gradual type of moral learning found in behaviorism is consistent with the idea of incremental increases in cognitive ability studied by information-processing theorists

In short, research and theories of moral development are based on three components that represent psychoanalysis, learning theory, and cognitive-development theory; they are 1) an affective or emotional component, 2) a behavioral component and 3) a cognitive component.

### ***Piaget: Foundational Work With Moral Development in Children***

Social cognitive-developmental theory (e.g. Kohlberg’s theory) holds that social experience and cognitive growth are both necessary for children to construct increasingly better understandings of rules, laws and social and interpersonal obligations. Much of the foundation for Kohlberg’s work was laid by Piaget. Known for his ideas on children’s cognitive development, Piaget also studied children’s moral development. Based on observations of his own children, Piaget (1932/1965) theorized that children existed in “pre-moral” egocentric stage

with little awareness of rule up to the age of about 5 years. Between the ages of 5 and 10 years, children enter the first stage of “heteronomous or external morality” during which the child viewed rules as moral absolutes handed to them by god-like authority figures. Imminent justice is required for breaking rules. After ages 10 or 11, most children enter the second stage of “autonomous morality” in which rules are viewed as flexible and disagreements are resolved among “equal status” peers with parents less important and maybe even detrimental to moral growth. Piaget envisioned children as developing ever more complex and comprehensive moral schema over time by a gradual and incremental progression through an “invariant” sequence of moral stages, with much stage overlap and decalage. According to Piaget knowledge, including moral knowledge is structured and content specific. For all biological systems, Piaget found the “whole” (structure d’ensemble) to be more than the sum of its parts with the part-whole relationship existing as an unstable relationship between syncretism and juxtaposition. This disequilibrium produces adaptation that leads to a more ideal state of development (Lapsley, 2006).

In sum, much of the foundation for Kohlberg’s work in social cognitive-developmental theory was laid by Piaget. Known for his ideas on children’s cognitive development, Piaget also studied children’s moral development. Kohlberg’s theory states that both social experience and cognitive growth are necessary for children to construct increasingly better understanding of rules, laws and social and interpersonal obligations.

### ***Kohlberg and the MJI; the Major Research in Moral Judgment***

For Lawrence Kohlberg, the move toward a more refined and “higher” stage of moral reasoning and toward a more ideal state implied that higher moral reasoning stages were “better” than lower stages (Lapsley, 1996). These ideas resonated with Kohlberg who wished to develop a universal description of morality to counter the ideas of moral relativism. Kohlberg added

ideas from Dewey and Rawls to those of Piaget to formulate a moral stage theory consisting of 3 “hard” stages with 2 levels per stage with an emphasis on justice reasoning. Since higher stages were “better” and depended on more developed cognitive structures, there could be no decalage.

The stages are:

Level 1: Preconventional Morality depends on the egocentric aims of avoiding punishment in stage 1 and receiving satisfaction in stage 2;

Level 2: Conventional Morality depends on taking the perspectives of other people in order to gain social approval in stage 3 or to maintain social order in stage 4;

Level 3: Postconventional Morality depends on making decisions based on broad principles that serve people in a social contract in stage 5 and based on abstract universal principles of justice that are independent of society’s laws in stage 6.

In the Moral Judgment Test (MJT), Kohlberg presented moral dilemmas to teenagers from many different areas of the world using an interview process in which the reasons for decisions were more important than the decisions. Kohlberg concluded that moral reasoning was developmental and similar for all cultures and that people pass through all the stages in an invariant order but at different rates.

Widespread studies, including long-term longitudinal studies by Kohlberg (Colby and Kohlberg, 1987) and others have established the validity of Kohlberg’s basic claims but there are many criticisms of his theory. Gilligan (1982) claimed a gender bias against women that has stimulated research on issues of caring but that has not been substantiated by empirical evidence. Empirical evidence shows that Stage 5 is rare and does not appear in all cultures- indicating a Western bias (Snarey, 1985). Stage 6 is so rare that doubt is cast on the validity of the whole model. However, Lapsley notes, “some of the properties of principled reasoning [found in stage 6 and to a lesser degree in stage 5] seep down...” (2006, p. 49) into lower stages. Moral

reasoning often depends on context (Rest, Narvaez, Bebeau, & Thoma, 1999). That is, ethics is somewhat relative. Children's ability to differentiate between moral issues and social convention is underestimated by Kohlberg (and by Piaget). In fact, this distinction does not even appear in Kohlberg's theory until the higher stages. Kohlberg attempted to rid his theory of content to make it more universal but domain-specific knowledge influences thinking about moral dilemmas (Bebeau, Rest, & Narvaez, 1999). Additionally, people show evidence of decalage by reasoning at multiple stages. Kohlberg's research demonstrated a link between moral reason and moral behavior but the focus on moral reasoning over moral behavior has been criticized. Emotions, identity and motivation and other factors that affect moral reasoning, such as those discussed in Blasi's Self Model (Walker, 2004) are left out of Kohlberg's theory. Kohlberg's interview method is difficult to administer and its use of people's self-reported reasoning is questionable. Rest, Narvaez, Bebeau and Thoma (1999) provide an excellent discussion on the criticisms of Kohlberg's theory. They find Kohlberg's theory to be more incomplete than wrong.

Kohlberg also inspired many researchers. Turiel (2006; Lapsley, 1996 & 2006) sought to restrict the moral field, separating the moral domain from the domain of conventionality, with his Social-Conventional Reasoning (or transformational model) that emphasized periods of affirmation and negation in a stage-based theory similar to that of Kohlberg. He based his theories on the idea that knowledge is organized in different domains and follows different trajectories that conventional knowledge and moral knowledge have different developmental sources and that social knowledge develops differentially based on children's experiences. Disequilibrium (a Piagetian term) induced by conflicts that emerge from questioning one's own judgment and rejecting specific conventions leads to higher levels of moral reasoning. Strong emotions are involved. Importantly, Turiel suggests that moral implications are warranted only if a person's actions harm another. Strong empirical evidence to support Turiel's theory is not

found in the literature. Rest (1983; as cited in Lapsley, 2006) finds Turiel's definition of morality too narrow.

Other researcher's, (e.g., Hoffman, Gilligan, Eisenberg, Gibbs), have proposed models that are similar to Kohlberg's but with emphasis placed on "empathy," "care," "pro-social behavior," "sociomoral" or other constructs.

Turiel (1966) discovered that stage change could be induced by presenting students with moral arguments that conflicted with their level of moral reasoning. Research by Blatt and Kohlberg (1975; as cited in Lapsley, 1996) showed that moral discussions with students could increase their moral reasoning by one stage. Moral interventions in education have been inspired by this research work.

The move toward a more refined and "higher" stage of moral reasoning and toward a more ideal state implied that higher moral reasoning stages were "better" than lower stages (Lapsley, 1996). These ideas resonated with Kohlberg who wished to develop a universal description of morality to counter the ideas of moral relativism. The stages are Level 1: Preconventional Morality; Level 2: Conventional Morality; Level 3: Postconventional Morality.

In the Moral Judgment Test (MJT), Kohlberg presented moral dilemmas to teenagers from many different areas of the world using an interview process in which the reasons for decisions were more important than the decisions. Kohlberg concluded that moral reasoning was developmental and similar for all cultures and that people passed through all the stages in an invariant order but at different rates.

### ***Rest: Broader Models of Moral Functioning***

Rest has participated in a "big shift" on research in moral research toward broader models of morality that seek to better explain the link between moral judgment and moral action. He found morality to be deeply "rooted in the social condition and the human psyche" (Rest,

Bebeau, & Volker, 1986, p. 1). Rest says, “The function of morality is to provide basic guidelines for determining how conflicts in human interests are to be settled and for optimizing mutual benefit of people living together in groups” (p. 1). Rest notes several elements in human development that promote moral development. Among these are empathic feelings in the individual psyche, caring and supportive relationships, a moral self-concept and the development of a social conscience. Rest proposed the Four-component model *to explain how moral behavior occurs* (Rest et al., 1986) by identifying the processes that contribute to moral action. Four important components identified are: 1) Moral sensitivity (interpreting a situation as moral); 2) Moral judgment (determining a course of action); 3) Moral motivation (assigning priority to moral values); 4) Moral character (having the strength to execute a moral plan).

Rest’s model assumes that cognition and affect exist in all areas of moral functioning (Bebeau & Thoma, 1999). Rest and others developed the Defining Issues Test (DIT), a neo-Kohlbergian moral judgment instrument, as a paper and pencil test to replace the MJT. Rest, Narvaez, Bebeau, & Thoma (2000) label it a “device for activating [a person’s own internal] schemas.” It eliminates the subjective testing procedure of Kohlberg’s MJT and also eliminates Kohlberg’s upper stages. It poses moral dilemmas and offers a list of action choices and a list of justifications for each dilemma. These dilemmas are based on MJT. Individuals choose the best and worst action choices and justifications. DIT has been widely tested and shown to correlate well with the MJT (Rest, Narvaez, Bebeau, & Thoma, 1999). DIT2 is an updated version of DIT that, primarily, eliminates problems associated with outdated information in the original dilemmas. It correlates extremely well with DIT (1999).

In short, there has been a shift in focus on moral research toward broader models of morality. Rest tried to explain more fully the link between moral judgment and moral action.

He found morality to be deeply “rooted in the social condition and the human psyche” (Rest, Bebeau, & Volker, 1986, p. 1).

### ***Expansion and Consolidation of the Field***

Killen and Smetana (2006) argue that the psychological study of the field of moral development, by examination of affect, cognition, behavior and now, neuroscience, and of its applications in clinical and educational settings, has greatly expanded in the last two decades in terms of the number of theoretical perspectives represented and the range of topics studied with the inclusion of scholarship in social science disciplines, neurosciences and biology. Further, an understanding of morality by students and scholars from different branches of psychology, education, political science and etc., cannot be made without understanding morality’s “roots in childhood and adolescence” (p. xi). Additionally, Killen and Smetana note that there has been significant hierarchic integration of the old “grand” theories in moral psychology research.

In sum, the psychological study of the field of moral development has greatly expanded in the last two decades through the examination of affect, cognition, behavior and now, neuroscience, and of their applications in clinical and educational settings (Killen & Smetana, 2006).

### ***Ethical Reasoning and Testing in Medicine***

#### ***The Call for a New Test for Medical School Education***

It is significant that researchers have found that high scores on existing ethical reasoning tests, using *P*-scores from the DIT, correlate well with good clinical skills and low incidence of malpractice claims for practicing physicians (Baldwin et al. 1996), suggesting to researchers that good doctors are those with good moral reasoning abilities. Yet, published literature of studies in the medical field show that medical education often does not promote an increase of moral reasoning skills for medical students (Self, Schrader, Baldwin, Wolinsky, 1993; Patenaude,

Niyonsenga, & Fafard, 2003). Several factors, including bad faculty modeling and a “hidden curriculum” in medical education that instills poor attitudes and bad habits in students have been advanced to explain the poor test results (Hafferty & Franks, 1994; Hicks, Lin, Robertson, Robinson & Woodrow, 2001). Studies indicate that sufficient quantities of the right kinds of ethics instruction can enhance the moral reasoning skills of medical students (Self, Wolinsky & Baldwin, 1989; Self, Baldwin, & Olivarez, 1993; Self, Baldwin, & Olivarez, 1989). Unfortunately, current ethical reasoning tests such as Kohlberg’s Moral Judgment Interview (MJI) and its derivatives (e.g. the Defining Issues Test, Sociomoral Reflection Measure, etc.) are too abstract or fail to assess the types of ethical dilemmas specific to the medical field and to medical education (Bebeau & Thoma, 1999; Christie, Bowen, & Paarman, 2003). B. K. Redman’s (2005) entreaty that “Every field of practice has the responsibility to evaluate its outcomes...” was taken up long ago by Rest, and Narvaez (1994) and other researchers and by educators who have called for new instruments to assess morality in the professions. A new intermediate concepts measure, or ICM, should fill this gap for medical education.

### ***Link Between Moral Reasoning and Practice***

From a review of literature, it was found that researchers established a strong link between high *P*-scores on the DIT and good clinical evaluations for medical residents and nurses (Sheehan et al., 1980; Duckett & Ryan, 1994). Working with a group that included 244 pediatric residents of which 147 were from American and 97 from Foreign medical schools, researchers Sheehan et al., (1980) found a correlation of 0.57 between *P*-score and adjusted clinical performance using 18 performance characteristics based on the rating scale devised by Cook and Margolis (1974). Significantly, a high level of moral reasoning as measured by high *P*-scores was closely associated with the highest scores on the clinical performance measure and medical students with high *P*-scores were not found in the group of students with low clinical

scores (as cited in Candee, et al. 1982). Also, in similar studies with nurses, Duckett and Ryan (1994) found that nurses with higher *P*-scores at their point of entry into nursing education had much higher evaluations for their clinical performance in their later years. The correlation was very large at 0.58. These results fit quite well with the findings by Baldwin et al. (1996) for 53 orthopedic surgeons that High *P*-scores (above a mean of  $P=43.8$ ) correlated with low malpractice claims. Different findings were claimed by W. Bunch who found little correlation between DIT *P*-scores and claims per year for a group of 27 academic orthopedists. Bunch does relate, however, that this group had different characteristics than Baldwin's group with generally low malpractice claims throughout the group (Bunch, Storr, Hughes, & Baldsin (1996). In short, with little range in outcomes in Bunch's study, there were no factors to analyze.

In another study that probes the link between moral reasoning and action, Candee et al. (1982) investigated the relationship between moral reasoning as assessed by the DIT and decisions in dilemmas of neonatal care for 452 pediatricians (256 residents and 186 practicing physicians). Importantly, a questionnaire by Crane that assessed physicians' reported degree of activism in treating infants born with severe defects (designated "unsalvageable prognosis") was used rather than physicians' actual decisions on care. It was hypothesized that higher reasoning would correlate with less active treatment in cases where the family had requested a limit in the treatment or where medical treatment could offer little benefit to the patient. The hypothesis was confirmed for medical residents but showed mixed results for postresidents with activism in treatment either unrelated to or positively related to the prognosis and family attitude. This study would have been strengthened by using actual decisions on care rather than an assessment of attitudes about care. Interestingly, institution type accounted for 43% of the variance in scores, suggesting a strong socializing force in medical school education.

### ***Medical School's Influence on Moral Development***

The work of Sheehan and associates is a part of what Self & Baldwin (1994) describe as “the first wave” of empirical studies on the moral reasoning of medical students and residents (Self describes himself as a member of the “second wave”). Husted, a member of the Sheehan group reported what is perhaps the most significant finding of this research at the RIME Conference (Husted, 1978; as cited in Self & Baldwin, 1994). Using the DIT to assess the moral reasoning of 488 medical students, Husted found that medical students showed little progression in their reasoning ability with *P*-scores of 50.2 for 1<sup>st</sup> year students and 50.8 for 3<sup>rd</sup> year students. These results compare unfavorably with the expected increases in moral reasoning for other college students with similar educational levels found by Kohlberg and Colby (as cited in Rest & Narvaez, 1994) using the MJI and by many researchers using the DIT (Rest & Narvaez, 1994).

Similar results were found in longitudinal studies of 20 veterinary medical school students (Self, Schrader, Baldwin, Root, Wolinsky, and Shadduck, 1991) and on 20 medical students (Self, Schrader, Baldwin, & Wolinsky, 1993) tested at the beginning and end of their medical education using Kohlberg's (1984) original MJI. Small increases in weighted average scores (WAS) of 12.5 and 18.5 points, respectively, for the two samples were statistically insignificant, suggesting that the medical school experience inhibited normal increases in moral reasoning abilities for both groups. The researchers noted that the WAS spread narrowed significantly for the fourth year groups in each sample, indicating a strong socializing factor in the medical experience similar to that found in the study by Candee et al. (1982). No correlation was found for MCAT scores or GPA scores. Additional longitudinal studies by Self et al. (1993) for 30 medical students and Self et al. (1993) for 57 veterinary medicine students repeated these tests using the Sociomoral Reflection Measure (SRM) (Gibbs & Widaman, as cited in Nilsson, Crafoord, Hedengren, & Ekehammar, 1991). The mean increase in SRM score of 17.5 was

significant at the  $p \leq 0.5$  level for veterinary students. The mean increase of 10.67 was not significant at this level for medical students. One further study by Self et al. (1993) using Rest's (1979) DIT and a larger sample size of 68 veterinary students produced results comparable to the earlier MJI study. The strong regression to the mean evidenced by the narrowing of scores for all three veterinary medicine studies and for the medical studies suggests a powerful socializing force in medical school education. No parallel exists for college students pursuing other professional degrees suggesting that the results are no mere statistical artifact of the professional setting. Self-selecting bias in the samples may play some part in the results, especially for the studies with small sample sizes.

Moral judgment, of course, does not represent all of morality (Rest, 1986). Some researchers have studied the effects of medical school education on moral sensitivity, another component of morality. Consistent with the moral judgment assessments, the following surveys evidence a pattern of no increases in sensitivity scores during medical school education or slight increases in the early year's scores followed by a decline in scores during the later years. Bissonette, O'Shea, Horwitz, & Route (1995) asked Year-2 and Year-4 medical students to identify incidents that contained ethical concerns during clinical training. The results show a retardation of moral sensitivity during medical school. Bad modeling by faculty is suggested as a cause by the researchers and, in the discussion, it is proposed that issues relevant to student experiences be used in ethics curriculum. These themes occur quite frequently in the literature. In another study, Hebert, Meslin, and Dunn (1992) reported an increase in ethical sensitivity among medical students at the University of Toronto as assessed by an instrument containing four clinical vignettes for students between year 1 and year 2 followed by a decrease throughout the rest of the medical training.

Researchers have found that high scores on existing ethical reasoning tests, using *P*-scores from the DIT, correlate positively and significantly with good clinical skills and a low incidence of malpractice claims for practicing physicians (Baldwin et al., 1996), suggesting to researchers that good doctors are those with good moral reasoning abilities.

### ***Some Ethics Interventions Help***

The empirical assessments of the effects of different types and quantities of ethics instruction on the moral reasoning of medical students and residents has been the subject of a series of studies at Texas A&M by the “Second Wave” headed by Donnie Self. Studies by Self’s group and by Goldie, Schwartz, McConnachie, & Morrison (2002) and by Bebeau and Thoma (1994) show that certain types of ethics instruction, provided in sufficient quantities, increases medical student’s scores on ethical reasoning tests.

In one study, Self, Wolinsky, and Baldwin (1989) studied the effects on moral reasoning of teaching medical ethics by case study and by lecture methods. Both groups experienced statistically significant increases ( $p \leq 0.0001$ ) in the level of moral reasoning. The case-study method was more effective than the lecture method. In another study, Self et al. (1993) studied the effectiveness of teaching medical ethics to first-year medical students in an elective course by using weekly one-hour discussions of short films to develop moral reasoning. Using the DIT for pre- and post-tests, significant increases in moral reasoning scores were found for students who participated in one-quarter and two-quarters of study ( $p < 0.002$  and  $p < 0.008$ , respectively) compared with scores of students who did not take the course ( $p < 0.109$ ). In a later study, Self et al. (1998) use the DIT for pre- and posttests of moral reasoning skills. This time, the authors found that 20 or more hours of small group ethics case study significantly improves DIT *P*-score while less than 20 hours does not improve scores. Regression analysis revealed that age, gender,

undergraduate GPA, and MCAT scores were not related to the changes in scores for most of these studies.

Two longitudinal studies attest to the potential for effective ethics interventions in medicine. Self, Olivarez, & Baldwin (1998) tested students before and after a required freshman one-semester medical ethics course and at the end of the students' fourth year. Freshman mean pre- and posttest DIT *P*-scores of 47.7 and 53.7, respectively, and a mean post 4<sup>th</sup>-year DIT *P*-score of 56.5 seem to contradict earlier findings that medical education inhibits increases in ethical reasoning skills. The authors suggest that the current findings may be due to the efficacy of the required small group medical ethics discussion course. Interestingly, women scored higher than men on all three tests. A second study by Bebeau and Thoma (1994) of the effects of 39 hours of problem-oriented dental ethics curriculum spread over four years using pre- and posttest DIT comparisons and using open-ended questions found that ethics instruction improved students DIT score. The authors cited a high rate of regression. No sex differences were found.

Mixed results were found by Goldie et al. (2002) in a study of the impact of three years of ethics teaching in medical school in reaching consensus responses to the Ethics and Health Care Survey Instrument (EHCI) at the University of Glasgow Medical School. Researchers found that small group ethics instruction improved year-1 consensus responses on the EHCI. This result was expected. Unfortunately, lecture type teaching did not improve responses for years 2 and 3. Men's scores actually regressed in year 3. The researchers suggest that too little instruction and sample mortality affected the results of the study. They also suggest that there is a "hidden curriculum" in medical school education that inhibits normal growth in ethical reasoning but they make no proposals on how to deal with it other than to increase the amount of small-group ethics instruction. The results for this study are consistent with the early findings of Self's group that lecture is less effective than small group study and that a minimum threshold of

instruction time is necessary to improve moral reasoning scores. Perhaps there is a problem with the EHCI itself. It consists of 12 vignettes of which three are non-consensus. No information is given to indicate how a consensus scores is achieved or how the EHCI compares with other tests such as the DIT.

Certain research in regards to ethics intervention can help to improve moral reasoning skills, such as teaching medical ethics (case study method was more effective than lecture method) (Self, Wolinsky, & Baldwin, 1989), teaching medical ethics (Self, Baldwin, & Olivarez, 1993), dental ethics curriculum (Bebeau and Thoma, 1994), etc.

### ***Problems with a “Hidden Curriculum”***

The goal of ethics interventions in medical school education must be to produce physicians who are both able and willing to make sound ethical decisions, ultimately leading to better patient care. But just as James Rest (1986) notes that there is more to moral reasoning than “justice,” there is also more to ethics than is taught in educational interventions and while it seems obvious that interventions can help students to identify the ethical components in medical practice, it is not clear that learning to identify a situation as “moral” will lead physicians to make better choices in the day to day practice of medicine. Researchers Christakis and Feudtner (1993) find it puzzling that some people often choose to do something that they believe is wrong and, significantly for medical students, they argue that learning [about ethical principles in medicine, legal standards, etc.] will not necessarily lead physicians to make better choices. In fact, increased test scores brought about by ethics interventions may serve to hide underlying problems with medical education. That is, students may learn “what to say,” especially on the big questions about abortion choices, end of life issues, etc., but choose to make poor decisions in their daily practice.

What is the problem with the medical education experience and how does it inhibit moral growth for medical students? Patenaude et al. (2003) argue that students restructure their handling of ethical questions by using lower stage Kohlbergian arguments during medical school. They suggest that there is a hidden curriculum that shapes students' ethical decision making. Hafferty and Franks (1994) go much further. They argue that there is a formal and an informal teaching of medical ethics and that the critical factors that determine a physician's identity lie within the informal "hidden curriculum. "Informal" curriculum can be defined as all of the factors in the learning environment that shape the thinking of students. The "hidden" curriculum can be understood as the practices that are common in the teaching environment but unacceptable as ethical models for behavior (e.g. talking rudely in front of patients, failing to get patient consent for procedures, etc.). Ginsburg, Regehr and Lingard (2003) find that "avowed," "unavowed," and "disavowed" principles guide students actions and that "disavowed" principles (e.g. concern over grades) are dominant.

Christakis and Feudtner (1993, p. 251) say that it is the "allure – and sometimes the coercion – to be a "team player" that shapes students' thinking. Indeed, the results stated for Candee et al. (1982), show that institution type accounted for 43% of the variance in activism scores for decisions in dilemmas in neonatal care, suggesting a strong socializing force in medical school education. This study did not suggest that the effects of socialization were necessarily bad. Additionally, the longitudinal studies of the effects of medical education on moral reasoning by the Self group indicate that socializing force, too. Hafferty and Franks (1994) find that the effects of socialization, learning to think like and to act like a doctor, begin before a student enters medical school. Applicants to medical school "craft" an admission biography that justifies their ambition to enter medical school and they continue to craft biographies on into residency programs. Tara Young (1997) explains how lying and clever

deception become standard practice as students vying for entry into residency program create multiple applications tailored to several different programs with each program listed as the applicant's first choice. Dishonesty is part of the process.

Observing and participating in unethical behavior seems to be part of the process, too, and learning to act like a doctor may compromise a student's integrity. Feudtner, Christakis, and Christakis (1994) surveyed 1,853 3<sup>rd</sup> and 4<sup>th</sup> year medical students in Pennsylvania in 1992-1993. Only 36% responded, but of these 655 students, 80% reported having done something unethical, lying to patients or both behaviors. 98% heard physicians make derogatory remarks to patients, 61% observed behavior that they thought unethical by other medical team members and half of these (54%) felt like accomplices. 61% felt guilty about their behavior as clinical clerks and 62% felt an erosion of their ethical principles. Similar results were found by Hicks et al. (2001) who surveyed 108 clinical students (with a 90% response rate). Almost half (47%) reported that they had felt pressure to act unethically in clinical situation either occasionally, frequently or very frequently and 61% claimed to have seen a clinical teacher acting unethically occasionally, frequently or very frequently. Subsequent work with four focus groups with 20 clinical clerks revealed three major types of ethical dilemmas faced by the students in their clinical training: 1) conflicts between medical education and patient care, 2) responsibilities exceeding the student's capabilities and 3) involvement in care perceived to be substandard. The students noted that these issues were seldom discussed or resolved with the clinical instructors.

The problem with bad faculty modeling and the pressure to act unethically does not end with clinical clerkship as seen in a study by Baldwin, Daugherty, & Rowley (1998). In a survey of 571 second-year residents concerning their experiences in year-one of residency, 44.5% reported observing others falsifying records, 73.8% reported directly observing the mistreatment of patients, 46.7% reported that others had taken credit for their work and 72.8% reported that

they had observed colleagues working while in an impaired physical or mental state. Perhaps the most important statistic is that 28.6% of the residents stated that they had been required to do something that was unethical, immoral or personally unacceptable. The integration of these bad habits is a likely outcome for many of these residents.

Christakis and Feudtner (1993) offer reasons for students' silence. They analyzed clinical clerk's dilemma cases submitted in 1991-1992. In addition to the general socialization aspect, they find that many dilemmas are tied to student's position within the medical team. Students are wary of challenging authority, in part for fear of jeopardizing an evaluation and, given the students inferior status on the medical team, "squealing" is not an option (St. Onge, 1997). To remain silent is in the best interest of the student but there is a price to pay for silence in the face of bad behavior. Years of exposure to unethical conduct can erode a student's personal code of ethics and habits acquired from faculty during clinical rounds and in the hallways, call rooms and cafeterias may shape the future doctor as much as the formal ethics curriculum designed by the medical schools (Hundert, Douglas-Steele, & Bickel, 1996).

Reasons suggested for physician's bad role modeling: legal concerns, making money, paternalism, lack of time and "burnout" (Hundert, Hafferty, & Christakis, 1996). Role models for physicians may set an inappropriate tone, too, as shown by a survey of academic leaders at medical schools (Eric Campbell as reported in Townhall.com; 2007). Surprisingly, 60% of department heads at medical schools in the U.S. reported ties to drug companies and 27% said that they had recently worked as paid consultants for the industry. These ties are not all negative, of course, but these academic leaders are the ones charged with the responsibility to protect students from the negative effects of receiving gifts from the pharmaceutical industry (Rogers, Mansfield, Braunack-Mayer, & Jureidini; 2004).

This literary review and discussion is not intended to imply that physicians simply act immorally. As Branch (2000) noted, instructors also feel the need to be team players and seldom give good feedback to trainees, especially negative feedback. Consider also that a physician could protect his patients by making decisions such as these:

1. Order a clinical clerk to indicate incorrectly on an elderly patient's chart that the patient's diarrhea is improving (after a prescription is written but before any test results are received) so that the patient can leave the hospital as quickly as possible in order to avoid needlessly exposing the patient to a potentially deadly nosocomial infection. (Of course, the physician should not order the clinical clerk to make the notation. If a notation is made, the physician should make it and explain the rationale to the clinical clerk).

2. Make a diagnosis of "gestational diabetes" for a pregnant patient who has refused to take the test necessary to establish the diagnosis so that the patient can be referred to a diabetes clinic that will provide diet counseling to limit the potential effects of the disease (with the bill to be paid by the patients insurance), then explain the rationale for the diagnosis only after being questioned by a confused and possibly angry patient.

In short, Patenaude et al. (2003) suggest that there is a hidden curriculum that shapes students' ethical decision making. Furthermore, Hafferty and Franks (1994) argue that there is a formal and an informal teaching of medical ethics and that the critical factors that determine a physician's identity lie within the informal "hidden" curriculum. The "informal" curriculum can be defined as all of the factors in the learning environment that shape the thinking of students. The "hidden" curriculum can be understood as the practices that are common in the teaching environment but unacceptable as ethical models for behavior (e.g. talking rudely in front of patients, failing to get patient consent for procedures, etc.).

### *Suggestions to Fix the Problems of the “Hidden Curriculum”*

Many researchers have suggested the inclusion of dilemma cases specific to the day to day decisions that medical school students make (Christakis and Feudtner, 1993; Bissonette, O’Shea, Horwitz, & Route, 1995). Others have suggested the use of practical dilemmas with less emphasis on the application of basic principle to extreme cases (St. Onge, 1997). Hafferty and Franks (1994) disagree somewhat with the idea of incorporating student specific cases into the curriculum. They suggest that less emphasis be placed on the physician-patient relationship and more be placed on medicine as an organizational and institutional entity thus establishing ethics as a core component of medical practice. Teaching the teachers is important, too, as shown in these four recommendations.

1. Instructors should become aware of the perceptions of the students.
2. Identify and address the “hidden curriculum” using a consortium of students, faculty and outside experts.
3. All faculty members should be willing and able to identify and discuss ethical issues found in the labs, at bedsides, etc. and should pass on their experiences to the students.
4. Establish the relevancy of ethics to medicine at the organizational level by recognizing hospitals, medical schools and clinics as ethical entities thus changing the framework from doctor-patient. (Of course, these “ethical entities” need to be reminded of their responsibilities as role models to students and to physicians).

Hafferty and Franks (1994) fail to emphasize that hospital administrators will have to create caring and supportive environments that foster good ethical skills for staff members.

Interestingly, one school has drafted a policy which can help protect students from being asked to behave unethically (Doyal, 2001). Developed at St. Bartholomew’s and the Royal London School of Medicine and Dentistry with the help of the students, the policy guides clinical

instructors in good ethical practices. Crafted primarily as a patient's bill of rights, the document forbids students from performing acts beyond their level of expertise and requires proper supervision by instructors. Patients must be fully informed of the student's trainee status and permission obtained for all training activities. Students are required to politely refuse to act outside of the guidelines. This seems like a simple solution to some of the problems identified with the "hidden curriculum."

In brief, Hafferty and Franks (1994) propose that teachers are important, and to help medical students develop ethical reasoning, they should implement the followings: 1) Instructors should become aware of the perceptions of the students. 2) Identify and address the "hidden curriculum" using a consortium of students, faculty and outside experts. 3) All faculty members should be willing and able to identify and discuss ethical issues found in the labs, at bedsides, etc. and should pass on their experiences to the students. 4) Establish the relevancy of ethics to medicine at the organizational level by recognizing hospitals, medical schools and clinics as ethical entities, thus changing the framework from doctor-patient.

### ***Other Factors That Could Inhibit Moral Growth***

The pressure that hospital administrators place on staff members to manage resources may negatively impact physicians and students. For example, an uncaring hospital environment created by administrators of teaching hospitals who encourage physicians to "push patients out of the door" in order to make room for new patients is cited by Dr. Jonathan Pulman (2003) as an aspect of training that inhibits the normal development of moral reasoning for medical students. Pulman argues that the practice of posting reports that rank staff members according to length-of-stay, savable days and other statistics is a form of "shaming" that encourages staff to conform and thus regress to Kohlberg's stage 3, the stage to which students tended to move in the study by Patenaude et al. (2003).

Loewy (2003) provides some pertinent thoughts on how time constraints can inhibit moral reasoning skills for students: Informational overload and fatigue (especially for clinical clerks) leaves students with little time to digest and integrate “aspects of medical care that cannot be weighed, measured or otherwise quantified” (p. 172). In other words, ethical sensitivity takes a back seat to getting some sleep for a tired student or clinical clerk. Also, talking to patients and taking time to think about patients’ care receives too little emphasis in a society that overvalues “procedures and emphasizes doing (e.g. catheterizing, operating)” (p. 174). One more often repeated quote that is repeated by Loewy (p. 174) about spending time with patients is instructive: “Good ethics starts with good facts.”

Students are not without blame, either, as the survey by Feudtner et al. (1994) shows. They participate in the unethical behavior and they are not always coerced to behave badly as a qualitative study by Ginsburg et al. (2003) illustrates. Students were shown five videotaped scenarios each of which depicted a student in a professional dilemma which required action. Students were asked to place themselves in the film student’s place and to indicate their responses and reasoning. Grounded theory analysis of students motivations revealed that the students were predominantly motivated by the consideration of implications (potential consequence of suggested action) belonging to the “disavowed” category (e.g. concern for grades, evaluations or reputation) as opposed to “avowed” professional ideals (e.g. honesty or disclosure) or “unavowed” ideals (e.g. obedience or allegiance). Ginsburg et al.(2003) argue that the preference for the “disavowed” curriculum has implications for education, feedback and evaluation and further argue that educators should teach students to balance these “avowed,” “unavowed” and “disavowed” principles and implications.

Age, undergraduate GPA, and MCAT scores were generally not related to ethical reasoning scores for most of the studies reviewed, although women often scored higher than men

on ethical reasoning measures and on ethical sensitivity measures (Bebeau and Thoma, 1994; Goldie et al., 2002; Self, Olivarez, & Baldwin, 1998).

In sum, there are certain factors that inhibit moral growth, such as 1) the pressure that hospital administrators place on staff members to manage resources which may negatively impact physicians and students and 2) time constraints which can inhibit moral reasoning skills for students (Loewy, 2003), 3), etc.

### ***Problems with Existing Tests***

Rest and Narvaez (1994) are not the only researchers to call for new tests in the professions. McAlpine, Kristjanson, & Poroeh (1997) assert the need for valid and reasonable tools to assess practitioner's cognition of ethical issues for nurses in order for educators to examine the effectiveness of ethics teaching and learning. The need for easily administered assessment tools is demonstrated by researchers Lewin & Lancken (2004) who note that the evaluation of an experimental ethics program conducted at Case Western Reserve University and at the University of Pennsylvania was a weak point in the study of "Longitudinal Small-group Learning During the First Clinical Year." The following review shows that there are numerous moral reasoning tests but they are not specific to medical school education, they fail to assess the efficacy of ethics interventions in medicine or they are too difficult to administer widely and to score.

### ***Moral Judgment Interview***

To discuss the shortcomings of existing instruments to assess moral judgment, it is good to begin with the paterfamilias of the genre, the Moral Judgment Interview (MJI) of Kohlberg. Kohlberg's interview method (a production type of testing process) is difficult to administer. The MJI consists of a 45-minute, semi-structured, oral tape-recorded interview in which subjects are asked to resolve a series of three hypothetical moral dilemmas involving justice oriented

issues (e.g., Heinz and the drug). A systematic set of open-ended probe questions is supposed to reveal the logic of a subject's moral reasoning. A transcript of the interview is scored by a specially trained person using an 800 page scoring guide. The process is time consuming and labor intensive and its use of people's self-reported reasoning is questionable with respondents often sounding like philosophers (Rest et al. 1999). According to Self and Baldwin (1994), the testing process is expensive.

The MJI was intended to show "major markers in life-span development" (Rest & Narvaez, 1994, p. 9). The test score correlates to a highly abstract stage of moral reasoning development representing a "broad characterization of cognition" (p. 9) that, theoretically, functions as a default schema that guides actions in moral situations. Almost 40 years of empirical research support Kohlberg's theory but critics have argued that measures of Kohlberg's theory are unsuited to assess ethics training because of the abstract nature of the stage properties (Rest et al. 1999). Individuals, including physicians and other professionals, could have an adequate understanding of moral schemas but lack an adequate understanding of the specific moral concepts needed for everyday functioning. Kenneth Strike (1982) put it quite well in a statement about the teaching of ethics to teachers:

The essential problem is that the emphasis is on the development of abstract principles of moral reasoning instead of instruction in the more concrete ethical principles that should inform the daily activities of the practicing teacher. ... a teacher who has a good grasp of abstract moral principles may nevertheless lack an adequate grasp of specific moral concepts, such as due process. (p. 213)

### ***DIT and Other Moral Reasoning Tests***

Rest (1979) originally developed the Defining Issues Test (DIT) as a "quick and dirty" paper and pencils multiple choice alternatives to the MJI. Rest et al. (2000) label it a "device for activating [a person's own internal] schemas." It eliminates the subjective interview testing procedure of Kohlberg's MJI. The DIT poses moral dilemmas based on the MJI and offers a list

of action choice items and a list of justification choice items for each dilemma. Items are based on participants answers gleaned from years of testing with the MJI. Participants rate and rank the test items and then choose the best and worst action choices and justifications. Years of testing leads researchers to postulate that the DIT scores relate to the two more advanced (numbers 2 and 3) of the following three basic structures in moral thinking: 1) personal interest schema, 2) maintaining norms schema and 3) postconventional schema. The most widely used index for scoring the DIT is the “*P*” score which is interpreted as the degree to which participants choose items that correspond with the upper or “postconventional” stages of the MJI. The N2 index is a more recent improvement over the *P* score. Due perhaps to being more of a “recognition or information processing” test than a “production” test (e.g. the MJI), the DIT can supply information not available in the MJI (Thoma & Rest, 1999). For example, the U score assesses the degree to which an individual shows a preference for Kohlbergian stage type arguments by looking at the pattern of answers, particularly the relationship between action choices and justification choices. Good agreement between action choices and justification choices is thought to indicate “consolidation” in thinking (at whatever stage placement the choices indicate for the individual) whereas poor agreement indicates a transitional phase in thinking about moral issues or the use of some non-Kohlbergian process to consider the issues. Good agreement also indicates a link between moral judgment and action. Unfortunately, being an information processing type of test has disadvantages, too: The DIT requires minimum reading skills at the level for a normal 12 year old, making the DIT unusable for studying the moral thinking of children - the original basis for Piaget’s (1932/1965) moral reasoning developmental work and the inspiration for Kohlberg’s work. The DIT is easy to administer and has been widely tested with over 1,000 studies in 40 countries and has been shown to correlate well with the MJI (Rest et al. 1999). DIT2 is an updated version of the DIT that, primarily,

eliminates problems associated with outdated information in the original dilemmas. It correlates extremely well with the DIT (1999). Significantly, the DIT has been shown to be sensitive to changes induced by moral education (Thoma & Rest, 1999).

Hoffman, Gilligan, Eisenberg, Gibbs and others have proposed models of ethical reasoning that are similar to Kohlberg's but with emphasis placed on "empathy," "care," "pro-social behavior," "sociomoral" or other constructs. Some of those models have interview type tests. Others have "paper and pencil" tests. Regardless of the format of the tests, the comment by Strike applies to the DIT and all of the other "Kohlbergian" measures.

#### *Other Types of Measures of Morality*

Tests such as the EHCI (Goldie et al. 2002) for medicine, the DEST and the GDEST (Bebeau, Rest, & Yamoor, 1985; Ernest, 1999) for dentistry and the JAND and the ERT (Ketefian, 1981; McAlpine et al. 1997) for nursing are profession specific but are designed to measure attitudes toward ethical dilemmas rather than moral reasoning. These types of tests can be useful, though, by informing on problems with test design. For example, essay (production) type assessments of ethical reasoning (e.g. the ERT), while good as teaching tools when accompanied by clearly established institutional criteria and ample opportunities for feedback and practice, and shown to be reliable when conducted by faculty members with backgrounds in philosophy, are impractical for general application (Bebeau, 2002). Writing skills vary, self-reported data is questionable and properly trained staff may not be available to score the test.

Additionally, a critique of the JAND by Oddi and Cassidy (1994) illustrates the inadequacy of codes used to inform actions for nurses and used to design tests, even when experts are included in the design process. The JAND uses ethical dilemma cases distilled from a pool of stories related by practicing nurses. The selected stories were assessed by nurse clinicians for realistic representation of nursing practices, with action items generated by another

set of nurse clinicians. Items are rated for compliance with the Code for Nurses (American Nurses' Association, 1976) by eight nationally recognized nurses (the experts). Scores for the test are determined by awarding a point for each selection of an action item for the dilemma cases that agree with the "Code." Oddi and Cassidy (1994) argue that the code is too abstract and contains ambiguous statements and conflicting messages, in part due to multiple revisions in the code over many years. Thus, the reliance on agreement with the "code" minimizes the value of the experts' own judgment in deciding how to deal with the ethical dilemmas and brings into question both the content validity and the scoring method for the test. A Google search of the Internet and searches on PubMed and other databases reveals no valid measure to assess the moral reasoning of medical students at the intermediate concepts level.

In sum, the following research review shows that there are numerous moral reasoning tests, but they are not specific to medical school education. The tests fail to assess the efficacy of ethics interventions in medicine, or they are too difficult to administer widely and to score. There are certain problems with existing approaches to measure medical ethics in medicine: for instance, 1) MJI (Moral Judgment Interview) is too abstract for measuring moral thinking, too difficult to administer, and too time consuming. Interviewers need to be trained, and the test cannot be used with a group because participants must be interviewed individually. 2) SRM (Sociomoral Reflection Measure) has problems similar to MJI except it can be used with a group because it is a written test instead of an interview. 3) DIT (Defining Issue Test) is a multiple-choice, easy-to-administer, general test, but it is still too abstract and not specific to ethical reasoning in medicine. 4) JAND (Judgment About Nursing decision) is a measure of nurses' perceptions of moral behaviors. 5) ERT (Ethical Reasoning tool) is an instrument to measure ethical reasoning of nurses.

### *Intermediate Concepts: A Framework for Building New Ethics Tests*

The work of James Rest and associates helps to address the problem of having moral reasoning tests that are too abstract. Rest's (1986) Four Component Model (FCM) supports a broad conception of moral functioning that integrates moral sensitivity, judgment (e.g. Kohlberg), moral motivation and the strength of character necessary to follow through on moral decisions. Fortunately, this model also supports the development of analytical tools in a wide range of problems (Rest, 1986; Walker, 2002). To better understand moral thinking in every day processes, Rest and Narvaez (1994) have proposed three levels of reasoning in the moral domain which include general schema, intermediate concepts and codes. The general schemas are the abstract or "bedrock concepts" that guide ethical decision making for individuals. The codes (e.g. AMA Code of Ethics) are the surface-level ethical prescriptions for conduct found in professional codes of ethics that serve as guides for certain actions. They are generally too specific in nature to inform every day decision making on ethical issues. In between these two ends lie the intermediate concepts (ICs). This middle ground provides the rationale for decision making on issues of "informed consent," "professional autonomy," "intellectual freedom," etc. It is these types of ethical questions about which most ethics instruction programs are organized. While numerous writers point out deficiencies with using principles to decide ethical dilemmas, no one actually suggests that there is better theoretical scaffolding on which to build an ethics curriculum (Gillon, 2003; Campbell, 2003).

The test structure of an ICM is similar to the DIT, using dilemma cases and action and justification choice items. ICMs, however, are specific to each profession and use real cases specific to the professional field and rely on the thinking and the judgment of experts in the field for the design of items on the test and for the development of a scoring key for the test. Bebeau

and Thoma's foundational work on ICMs provides a good framework on which to build an ethics exam for medical professionals.

To date, there are two empirical studies that employ ICM's. First, Bebeau and Thoma (1999) developed a prototype ICM, the Dental Ethics Reasoning Judgment Test (DERJT) to assess the outcomes of dental ethics instruction. An inspection of Bebeau and Thoma's work on the DERJT is informative. The DERJT study should prove easy to replicate (given capable and hard working researchers) because it is clearly described and shown to be valid by several measures. The research design and methods, limitations, and chain of reasoning from evidence to theory and back are well covered. The authors explain how specific research on a prototype measure of intermediate level ethical concepts for assessing the outcomes of dental ethics instruction in dental ethics is used to validate the new type of test. They focus on six areas: 1) strategies to ensure the integrity of the measure, 2) exploring consensus among ethics, 3) exploring differences among groups, 4) exploring the relationship with other measures of expertise, 5) acquisition of intermediate concepts and the role of moral judgment in transfer to novel cases, and 6) linking intermediate concepts and their components.

The test development is well covered and shows clearly that *content validity* is practically assured. Dental practitioners were asked to write up cases that focused namely on intermediate level ethical concepts of dental profession instead of knowledge of procedure/techniques, and then they were examined for validity by other practitioners. Technical terms were to be limited. Twelve dental residents and four dental faculty members were asked to study the cases and to write (1) what should be done and (2) why the proposed action choice was justified for the case development. Written responses were used to generate acceptable and unacceptable action choices and justification of each item for each case. *Item validity* is practically assured in this type of process. Then the same groups scored the action choices on a 4-point scale (i.e., highly

defensible, defensible, questionable, not at all defensible). Then, the two best and two worst action choices were ranked. The same basic procedures were used for justifications. The case was discussed and items added or modified. While experts did not necessarily arrive at a single best answer on any of the intermediate concepts, there was very good agreement on what constituted the *set* of applications of the concept found to be acceptable and unacceptable (Bebeau & Thoma, 1999). On a second-level tryout, a new group rated, ranked, discussed, modified the action choices and justifications and reached consensus on the best and worst action choices and justifications.

A sample case is presented to 14 teachers of dental ethics nationwide. (Additional comments could be made on the development process but this quantity is enough to show that the process is rigorous and methodical and, importantly, specific to the dental profession). Scoring is done by awarding points for participants' responses for best and worst action choices and justification choices that agree with the consensus best and worst choices of the experts. Points are deducted for participants' choices that disagree with the experts' choices (miscoded choices).

There is a considerable amount of data analysis. The data was collected in a way that allowed *construct validity* to be tested. Three different groups (college freshman, dental freshman, and dental senior) that are expected to differ in moral judgment development (as measured by the DIT) and in dental ethics experience were tested with the new instrument and with the DIT. Demographic data was collected. The numbers of students, school location and type and mean DIT scores for students at these institutions is given. The way in which the students were recruited for the test, the number who refused to participate and the number who failed to complete the tests is unknown. Therefore, there is no basis for commentary on sample bias. As expected, seniors scored highest (*effect size* = 1.08) and college freshmen scored lowest

(*effect size* = 0.88). As expected, novice groups also misclassified items. Measurement validity can thus be claimed to be good for this test. *Construct validity* is indicated by the findings of a correlation between DIT scores and ICM scores ( $r = 0.26$  for dental freshmen and  $r = 0.33$  for seniors). Also, as expected, GPA was unrelated to ICM scores. The transfer of learned concepts to novel cases is explored by testing expert and novices groups with cases of which some are familiar and some unfamiliar to the experts and all are unfamiliar to the novices. The authors argue that, for unfamiliar cases, the ICM scores in and of themselves offer little beyond more general moral schema tests (e.g. the DIT) and so DIT *P* scores are treated as cohort in a repeated measure ANOVA. For unfamiliar situations, some evidence that moral schema (as measured by the DIT scores) may contribute to intermediate concept measures is gleaned from this data. That is, it seems likely that a good understanding of postconventional moral schemas is necessary for transfer to occur to new problems. The interpretation of the results and the links to the theory are easy to follow. Finally, the authors suggest that a theoretical link between Component 2 and Component 4 in Rest's FCM is indicated by the close relationship between the action choices and the justifications of the seniors for both familiar and unfamiliar cases. Bebeau and Thoma do a good job of interpreting the results in terms of the theory.

There is no information about subgroups (e.g. women) in the data analysis. This is not unexpected because this study is designed with a more general aim in mind: to validate the new instrument. The authors describe research methods clearly. The test design is quantitative in nature and the data collected are analyzed by appropriate statistical methods. The *construct validity* of the instrument is demonstrated in several ways. Findings are shown only on graphs. However, the graphs are logically arranged and easy to read. Future lines of study are suggested.

The preliminary work with the DERJT supports the use of the instrument to measure medical ethics interventions in dental education (Bebeau and Thoma, 1999) and, more generally,

it supports the use of ICMs as diagnostic tools (Bebeau, 2002). Walker (2002, p. 375) sums up Bebeau and Thoma's work with intermediate concepts this way:

The preliminary evidence with the ICM indicates that it does differentiate students at different levels of education, that it is sensitive to the effects of professional-ethics training and that it provides non-redundant information to the DIT (the DIT and ICM are only modestly correlated indicating that intermediate concepts are not simply reflective of the more abstract developmental stages or schemas).

Walker adds that research with intermediate level concepts developed for concepts in everyday life could advance the field (of moral research). Thoma (2006) does exactly that in the development of the AD-icm, a measure of adolescent thinking. The DERJT was used as the prototype for the development of this ICM. A challenge for Thoma was to identify the intermediate concepts for the adolescent population. This was done by surveying the content of character education programs for core concepts (e.g. honesty, fairness, tolerance, self-discipline, etc.) similar to concepts found in the professions (e.g. informed consent, confidentiality, etc.). Thoma argues that there is a "conceptual overlap" between aspects of character found in the character literature and intermediate concepts found in the professional literature. Another challenge was to define the "experts." It was decided that graduate students in human development were not too far removed from the cohort under study and had sufficient understanding of adolescent development to act as experts. Findings from the AD-icm show that adolescents who have acted out within the school environment also have more difficulty making adequate moral decisions. The instrument is found to be a measure of conventional thinking, sensitive to high school age educational groups. Importantly, the success of the AD-icm supports the claim by Rest and Narvaez (1994) that intermediate concepts are broad-based and

can be generalized to non-professional groups (Thoma, 2006) and the portability of the measure encourages the development of ICs for other groups and settings.

With regards to the numerous moral reasoning tests used in medicine, they are not specific to medical school education, fail to assess the efficacy of ethics intervention in medicine, and are too difficult to administer widely and to score. In conclusion, Intermediate-Concept Measure (ICM) represents an expansion of component 2 of Rest's Four-component model (moral judgment). It measures concepts that are more specific than the abstract stage and more general than the codes of ethics in professions. The intermediate concepts include the concepts of day-to-day life moral decision-making, and it is less abstract than post-conventional schema (Rest, Narvaez, Bebeau, & Thoma, 1998). According to Bebeau and Thoma (1999), intermediate concepts are reflected in the content of professional ethics in areas such as informed consent, confidentiality, professional autonomy, and so forth. The intermediate concepts can be applied to persons with professional education, such as nurses, medical doctors, lawyers, and so forth.

### ***Considerations for Testing in Thailand***

1. Thailand claims to follow the U.K. / U.S. models for medicine. How well? Barbados also claims to follow the British model but a study in Barbados by Hariharan, Jonnalagadda, Walrond & Mosley (2006) found that the hospital staff knew little of the law pertinent to their work. This could be true for medical personnel and education in Thailand as well.

Also, a report by Wolfgang Hennig (2006) entitled "Bioethics in China" has some relevance to Thai medical practices. China claims to follow the U.S./European ethics model. The report says that Chinese regulations and guidelines are essentially the same as those in Europe or the USA but that the implementation in research and daily clinical practice is not comparable to Western standards. Additionally, functional control over the public health care

system is poor - especially in remote areas of the country. This report could well have been made for Thailand.

2. British and American models are generally very similar but they differ somewhat in the allocation of resources, particularly in old age care. Doctors in Britain are less likely than their American counterparts to start antibiotics for pneumonia in older patients, especially when there is dementia (Lockwood, 1985), less likely to begin dialysis on patients over sixty five years of age or to begin chemotherapy on patients of any age with an unsalvageable prognosis (Thomasma, 1998). While the U.S. has an insurance based payment system in which physicians routinely order expensive and sometimes unnecessary tests and procedures, Britain has a two-tiered medical system in which basic services are provided free of charge and better services (usually) are provided by the private sector. Thailand has a two-tiered pay as you go system of medical service that is highly subsidized by the government so that poor people can afford most routine medical care. Better care is found in privately owned facilities and in special government promoted hospitals that cater to foreign clients. Proof of insurance or cash in hand is required for admission to these better facilities.

3. Cultural differences: The “ Noble lie” to spare feelings; physicians are gods; students at all levels are expected to listen to and to never question the teacher; extreme respect for older people and deference to people with higher rank; developmental paths for men and women are different; strong belief in “Fate.”

4. “Losing face”.... Do nothing to injure another. Never embarrass anyone. “Losing face,” as discussed in “Bioethics in China,” (Hennig, 2006) is important in discussions of Chinese failures or shortcomings in medicine and it is important in Thailand, too.

5. Thais assume that products from the U.S. are the best in the world, those from the U.K. are a close second and this includes research instruments. A medical instrument or measure constructed entirely with Thai “experts” would likely have little face value in Thailand.

6. Identifying ethical concepts in a 2<sup>nd</sup> language. A study by Ypinazar & Margolis (2004) of 62 Arabic speaking students in their 3<sup>rd</sup> week of studying medicine in the English language showed that these students could identify Western medical ethics constructs with minimal instruction. The results were independent of the student’s English language skills. Medical students in Thailand use textbooks in the English language so Thais should be able to identify basic concepts, at least, in English.

7. Thais respond slowly, if at all, to requests from lower ranking persons.

***Thai Officials Ask for the Test to be “Made in USA”***

An important consideration in the development of the test in Thailand is that the Ministry of Public Health, Thailand, represented by Dr. Anupong and Dr. Suwat, has suggested that the test be made entirely in the U.S. then administered in Thailand. While this would seem to diminish the value of the instrument for use in Thailand, I believe that a test developed in the U.S. with U.S. experts would, in fact, suit the needs and wishes of the Thai medical community better than one created with Thai experts.

Literature research on ethics curriculum, aims, models, etc. for medical programs in many western and non-western countries suggests that a universal intermediate concepts measure can be created for countries that follow western ethics models if careful attention is paid to the selection of action and justification items. The same research suggests that individualized ICMs can not reasonably be made for some countries because consensus in dealing with medical ethics dilemmas does not yet exist among indigenous experts. For example, ethics programs have existed in Thailand for about twenty years but it could be said that these programs are still in

their infancy and that their reliance on imported western ethics models leaves Thailand without a single ethics model from which “expert” opinions can be drawn. Dr. Suwat and Dr. Anupong are concerned with this issue.

Experience in medicine in Thailand by this researcher and a review of the literature concerning medical ethics in Thailand indicates that the ethical dilemmas faced by Thai doctors are essentially the same as those faced by their western counterparts and that differences in ethical perceptions by Thais are based primarily on the Asian principle of “interdependence” and to a lesser extent on the Buddhist principal of “karma”. (A good source for information on these topics is found in “*Buddhism and Medical Ethics: A Bibliographic Introduction* (Hughes & Keown, 1995).” Relative to westerners, there is more importance placed on the Thai family to make decisions on medical issues. Otherwise, issues of autonomy are similar for western and Asian cultures. “Karma” is often applied to a dilemma in a way that involves what is best for the doctor in some future life rather than what is best for the patient. For example, no Thai doctor wants to “pull the plug” on a patient even if the family, the physician and all other parties believe that this is the best course of action for the patient. The doctor could pay a price in a later life. Thus, the “disavowed” ethical decision to shift the burden for making difficult judgments to others can be somewhat justified by Buddhist religious beliefs but such a decision is inconsistent with modern western-based medical ethics programs in Thailand nonetheless.

It is possible to design action and justification choice items on a universal ethics test to identify elements that represent “formal” and “informal” curriculum for multiple cultures even when obvious non-western factors such as “karma” are present. Following the example of the work of Dr. Thoma with the Utilizer Score for the DIT (1999), it seems that the agreement between action and justification item choices could provide valuable information for medical educators in Thailand and elsewhere although assessments of “consolidation” in ethical thinking

should contain careful considerations of the relative importance of conflicts between multiple ethics models in the study population.

Thus, the population in this study, value western products and would have little confidence in a test created from 100% Thai sources. Without the acknowledged use of sources in the U.S. or the U.K., the work of this native Thai researcher would be suspect.

Thus, an ethics test created in the U.S. with U.S. experts would be more consistent with the formal ethics curriculum in Thai medical schools than a test made with Thai experts; it would have more “face” value to Thai educators; it would make the Thai government happier.

## **CHAPTER III**

### **METHODOLOGY**

This chapter describes the development of the MD-ICM, design of the study, data collection process, and statistic analysis. The purpose of this study is to develop and test a new instrument, the MD-ICM, an intermediate concepts measure (ICM) based on the work of Bebeau and Thoma (1999), which is designed to help educators assess the ethical reasoning of Thai medical students.

This chapter presents the study's methodology. The first section discusses the development of the MD-ICM. Next, the author will describe research participants and how will they recruited. This is followed by a description of the research instruments and procedures for collecting data and strategies for analyzing data relative to the study's research questions. This study was approved by the IRB (The University of Alabama Institutional Review Board) for the Protection of Human Subjects and found to be consistent with IRB policies (See appendix A).

#### ***Developing of the MD-ICM***

This section explains how the MD-ICM (Medical Doctor Intermediate Concepts) was developed. The method used to measure intermediate concepts was straightforward and simple. For medicine, students read a short story (case) that contained a dilemma relevant to the medical field. They were asked to read a list of action choices supplied with each dilemma. They rated possible action choices; then they ranked the three best and two worst actions choices. Similarly, they rated and ranked possible justifications for their choices from a list. Answers were compared to the consensus answers of ethical experts in that profession.

Step 1. This study uses medical dilemma cases based loosely on studies found in the medical literature and modified extensively to reflect cases from the researcher's own experience working in the medical field for over 15 years. The content of the cases (stories) was consistent with experiences familiar to the medical council in Thailand. The researcher developed the items for each story by making lists of possible action choices and justification choices for each story. Careful attention was given to items that represented practices that occurred frequently but may be ethically hard to defend. In order to make sure that these stories were clear, consistent, reasonable and easy to read, a small group of physicians, residents and nurses ( $n=5$  to 10 each) representing a wide range of experiences and abilities in the medical field were asked to read the stories and action items and to note any problems or changes that should be made. (Nurses were included at this stage to broaden the ethical perspectives and to more carefully and completely consider the ethical dilemmas (Hanson, 2005). Group members reviewed the list of items and rate each proposed action choice and justification choice as "defensible," "neutral" and "not defensible." They also asked to note any additional action choices and justifications for any of the stories from their own knowledge that should be used with the stories including "real world" choices that represent common practices that may be ethically questionable. Based on this information, in story 1, item 10 was removed. Several "defensible," "neutral" and "not defensible" items included for each list. Ethics curriculum was geared toward discerning the better choices from among several options so that even the "neutral" choices should be generally reasonable. The researcher was very careful to make the stories and choices easy to read because Thai students took the test in the English language.

Step 2. Expert (i.e., medical ethicist and faculty) consensus on items and scoring: Next, four or five experts, including medical instructors with experience in teaching ethics, and at least two medical ethicists were asked to participate in a panel discussion to read the revised

stories and items and to assess the action choices and justifications as “defensible,” “neutral” or “not defensible.” The experts were also asked to note any action choices and justifications from any of the stories or from their own knowledge that should be used for each story. Special attention was given to the actions that represented ethically questionable but common practices and to justifications that belong to “disavowed” principles and implications (Ginsburg et al. 2003). The researcher consulted with the experts to modify the action choices and justifications to suit these individual stories and worked to achieve a consensus on the coding of items. One hundred percent agreement was not needed for “consensus.” Rather, 100% agreement was required for two choices for best and two choices for worst actions and justifications. Calculated means determined the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> best and worst choices for the experts, but all were considered equivalent. That is, the researcher classified each item as “defensible,” “neutral,” and “not defensible.” Any combination of 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> best and worst responses by a student was considered consistent with any combination of expert’s 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> best or worst choices by the experts.

In terms of the scoring key: For each of the stories, five experts from medical schools and hospitals across the U.S. or Thailand were asked to take the test. They rated each action choice item as “defensible,” “neutral,” and “not defensible” and ranked the three best and two worst action choices. They rated each justification choice as “defensible,” “neutral,” and “not defensible” and ranked the three best and two worst action choices. The key was constructed from the experts’ responses. The lists of experts are listed below:

-Dr. Win May, Director & Professor, Standardized Patient Program,  
University of Southern California, Los Angeles, CA.

-Dr. Dixie Fisher, Assistant Professor, University of Southern California, Los  
Angeles, CA.

-Areti Tillou, M.D., Assistant Professor, Associate Program Director, David Geffen School of Medicine at UCLA, CA

-DR. Anupong Suthamaniran, Medical Ethicist, Director of Medical Education, Chonburi Medical Center, Thailand

-Dr. Sairat Noknoi, Chonburi Medical Center, Thailand

-Dr. Sunida Adeetchart, Chonburi Medical Center, Thailand

With consideration to culturally sensitive items, Thai medical instructors with experience in teaching ethics courses experts in Thailand were asked to comment on the stories and items and to rate the items and make suggestions. The researcher carefully considered the input from the Thai experts. The Thai experts were expected to make few suggestions due to cultural constraints on making negative comments. It was also thought that they would show more variability in their selection of items as “defensible,” “not defensible” etc. than U.S. /U.K. experts because the Thai medical system is based on multiple models (U.S. /U.K.) and because medical education is conducted in two languages (English and Thai), thus promoting ambiguity. The Thai experts were used primarily to investigate differences in response patterns (between U.S. and Thai experts).

Step 3. Testing for reliability: Fifty medical students were in the CPIRD project, but were not included in the sample of this study. They were used to test the reliability of the instrument. IRB approval and consent was obtained. As in step 2, the score was assessed by Cronbach’s alpha. The value was accepted if it is from 0.75 to 0.95. The test is now ready to administer.

### ***Participants***

Participants were Thai medical students drawn from seven affiliated universities in Thailand. Five universities recruited from the CPIRD were Thammasart University, Chiangmai

University, Naresuan University, Mahidol University, and Khongon University. Two universities recruited from the Non-CPIRD were BMA Medical College, Suranaree University, included from non-CPIRD program. A summary sample presents in table 1.

More interestingly in this study, each CPIRD medical student was recruited from a rural area and attended classes in basic science and preclinical studies for three years at one of the seven major public medical schools in the large metropolitan areas. In years 4-6, non-CPIRD students have clinical training at the large hospitals associated with the medical schools while CPIRD students have clinical training at the seventeen Medical Education Centers owned by the Ministry of Public Health and located in the provincial hospitals. The major public medical schools are responsible for the quality of the education and grant the MD degrees. All medical students in Thailand begin their medical studies after completion of high school.

These universities used the newer integrated ethics curricula. The curricula do vary somewhat among the universities in structure and in the year of study for the initial implementation. Content is fairly consistent for all the universities. There is only one exception, the Mahidol University and its 4 affiliated rural medical schools still use non-integrated program, which enrolled in Moral Dilemmas in Medicine (SHHU 111) at Mahidol University. SHHU 111 was part of the core curriculum that was previously required for all second year medical students in Thailand.

Table 1. A number of Medical Students, University and Medical Center Affiliated with University

Class /curriculum	Universities	N	Medical center affiliated with University
Non-integrate	Mahidol (CPIRD), Bangkok	117	1. Maharatnakhonratsrima 2. Sawanpracharak 3. Rajchaburi 4. Maharatnakhonsrithamarat
Integrate	Thammasat (CPIRD), Bangkok	134	1. Saraburi hospital 2. Surathanee hospital
	Naresuan(CPIRD), regional area	95	1. Budhajinarat hospital 2. Utharadit hospital
	Khongan(CPIRD), regional area	155	1. Khonkan hospital 2. Supasithiprasong hospital
	Chiangmai(CPIRD), regional area	26	1. Lampang hospital
	Suranaree (Non CPIRD), regional area	75	1. Maharajnakhonratsrima
BMA Medical College (Non CPIRD), Bangkok	40	1. Vajira hospital	

The samples described in Table 1 were divided into three groups (Year I, Year II and Year III). The first group consisted of 1<sup>st</sup> year students ( $N=217$ ; Male = 89, Female = 128; 18-20 years of age). The second group consisted of 2<sup>nd</sup> year students ( $N=262$ ; Male = 111, Female = 151; 19-21 years of age). The third group will consist of juniors ( $N=148$ ; Male = 63, Female = 85; 23-25 years of age).

### **Materials**

1. *DIT2-Defining Issues Test*. Rest (1979) originally developed the Defining Issues Test (DIT) as a “quick and dirty” paper and pencil multiple choice alternative to the MJI. Rest et al.

(2000) labeled it a “device for activating [a person’s own internal] schemas.” It eliminates the subjective interview testing procedure of Kohlberg’s MJI.

The DIT poses moral dilemmas based on the MJI and offers a list of action choice items and a list of justification choice items for each dilemma. Items are based on participants answers gleaned from years of testing with the MJI. Participants rate and rank the test items and then choose the best and worst action choices and justifications. Years of testing lead researchers to postulate that the DIT scores relate to the two more advanced (numbers 2 and 3) of the following three basic structures in moral thinking: 1) personal interest schema, 2) maintaining norms schema and 3) postconventional schema.

The DIT requires minimum reading skills at the level for a normal 12 year old. The DIT is easy to administer and has been widely tested with over 1,000 studies in 40 countries and has been shown to correlate well with the MJI (Rest et al., 1999). DIT2 is an updated version of the DIT that, primarily, eliminates problems associated with outdated information in the original dilemmas. It correlates extremely well with the DIT (1999). Significantly, the DIT has been shown to be sensitive to changes induced by moral education (Thoma & Rest, 1999). The DIT is a widely used test of abstract moral reasoning, which is sensitive to ethics education.

2. *MD-ICM (Medical Doctor Intermediate Concept Measure)*. This study utilized a Medical Doctor - Intermediate Concepts Measure (MD-ICM) to assess the outcomes of moral judgment development knowledge – at the ICs level (i.e., follows from the DERJT developed by Bebeau and Thoma). Medical dilemma cases for the MD-ICM were based primarily on the concepts below (Note that multiple concepts apply in all of the cases): See appendix B.

Case 1. Patient autonomy

Case 2. Confidentiality

Case 3. Beneficence vs. patient autonomy

Case 4. Non-maleficence (assisted suicide)

Case 5. Beneficence

Cases containing stories relevant to students' perceptions of their position on the "medical team" and to their perception of the medical institution as an ethical entity (the larger team) as discussed in the "hidden curriculum" section of this paper may be substituted for two of the cases shown above.

### ***Data Collection Procedures***

The Thai medical students in the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> year in the CPIRD and non-CPIRD were asked to join in this research. Willing participants were asked to sign informed consent forms (See Appendix D). Dr. Suwat, the director of the CPIRD Project had expressed interest in the development of an ethics assessment instrument for the CPIRD students and had outlined the method to be used to recruit the Thai medical students. As the director of CPIRD, Dr. Suwat sent a letter to the directors at each the 14 medical centers informing them of the study and the use of the CPIRD students. For the Non-CPIRD students (e.g., Suranaree and BMA Medical College), the letter was sent directly to university dean office in order to inform medical students for the study. The directors of the centers contacted the directors of the 7 universities to inform them that the CPIRD students doing clinical studies took part in this study. There were two options for medical students to complete the DIT2 and the MD-ICM. First, students complete a questionnaire offered on-line via SurveyMonkey. The instructors for each class advised the students of the study and posted the web address. Students had a one week window to access SurveyMonkey. Participants read the description of the study and fill in the informed consent. Participants answered some demographic questions and then they completed the MD-ICM and the DIT2. All data transferred to investigators without personal identifiers. Second, there were certain problems with completing the tests via SurveyMonkey in Thailand, and then researcher

utilized paper versions of the test were administered to each university. The researcher contacted teachers to schedule time and arrangement to administer the tests each university. Again, any personal identifiers on the tests were removed or destroyed at the time the tests were picked up.

### ***Data Analysis***

The study involves the increase of knowledge in moral judgment development scores as measured by the MD-ICM of the 1<sup>st</sup>, 2<sup>nd</sup> year and 3<sup>rd</sup> year medical students. Thai Medical Students were administered both the MD-ICM and the DIT2. The MD-ICM scores, as represents by five main indices (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index) was calculated for all statistical tests. Post-conventional score (Pscore), N2SCORE, and Maintain Norms (STAGE4P) were calculated for the DIT2. Initial analysis was descriptive statistics for all research questions. This study mainly utilized SPSS software programs to measure the following: descriptive statistics (mean, standard deviation), simple correlation, ANOVA and regression analysis. After the data was checked for accuracy, the statistical methods are described to answer the research questions as follows:

### ***Research Question 1***

1. Are MD-ICM scores different for the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> year integrated ethics curriculum?

The MD-ICM scores, as represents by five main indices (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index), a 2 X 3 analysis of variance (ANOVA) compares mean differences between gender and educational level (i.e., freshman, sophomore, and junior). An alpha level of .05 ( $p < .05$ ) was employed for all statistical tests.

### ***Research Question 2***

2. Do medical students in the integrated ethics program scores higher on the MD-ICM than students in the old styles one-class curriculum?

To analyses this research question, a 2 (Gender) X 2 (program type) analysis of variance (ANOVA) was used to assess gender and curriculum effects. Students from seven universities, medical schools in Thailand (Thammasart, Chiangmai, BMA Medical College, Suranaree, Naresuan, Mahidol, and Khongan), were recruited for this study. All but one of these universities employed a non-integrated ethics program. The other six universities used a new integrated program. Given the difference in programs, the researcher compared the MD-ICM scores between the integrated ethics program and non-integrated ethics program to assess whether the MD-ICM is sensitive to different ethics programs.

### ***Research Question 3***

3. Does the MD-ICM relate to established measures of moral thinking such as the DIT as represented by the N2 and U scores and type variable?

A correlation analysis compared five dependent variables (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index), and DIT2 scores (i.e., Utilize score, N2SCORE, and Typenew).

### ***Research Question 4***

4. Do MD-ICM scores relate to DIT2 and Gender variables individually and in combination?

To analyze this research questions a series of hierarchical regression analyses were conducted on the five dependent variables (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index), and four independent variables (i.e., Gender, Pscore, N2score, and STAGE4P). In each case the structure of analysis was

similar: in step one of the hierarchy the ICM score was regressed on gender and the moral judgment score followed by in step two, a cross-product term, which assessed a nonadditive relationship between ICM, gender and moral judgment. In other words, this relationship addressed the question of whether the relationship between ICM and moral scores vary by gender.

### ***Summary of Methodology***

The purpose of this study is to develop and test a new instrument, the MD-ICM, designed to help educators assess the ethical reasoning of Thai medical students. The initial step is developing of the MD-ICM. There were three steps: Step 1- Development of Items, Step 2- Experts (i.e., medical ethicist and faculty) and Step 3- Testing for reliability. The instruments for this study utilized the MD-ICM and the DIT2. The MD-ICM score was represented by five main indices (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index). For the DIT2 was calculated by Post - conventional score (Pscore), N2SCORE, and Maintain Norms (STAGE4P). The samples were Thai Medical Students in seven different universities in different years (Year I, Year II and Year III). Quantitative analysis (e.g., descriptive statistics, simple correlation, ANOVA, and regression analysis) was used to answer for research question 1, 2, 3, and 4.

## **CHAPTER IV**

### **RESULTS**

The purpose of this study is to develop and test a new instrument, the MD-ICM, designed to help educators assess the ethical reasoning of Thai medical students. To that end, a sample of Thai medical students are needed in order to assess the measure within the target population. Thus the sample developed for this assessment focused on Thai medical students selected from different medical schools and universities. Specifically, the sampled universities included Thammasart university ( $n=134$ ), Chiangmai university ( $n=26$ ), BMA Medical College ( $n=40$ ), Suranaree University ( $n=75$ ), Naresuan university ( $n=95$ ), Mahidol university ( $n=117$ ), and Khongon university ( $n=155$ ). The demographic information will describe in next table (table 1).

Table 2. Demographic Information by Sample

Percent of Sample:	Male	Female
<u>Year</u>		
Year I	14.44	20.41
Year II	17.43	24.18
Year III	10.05	13.50
<u>Universities</u>		
Thammasart	6.23	14.64
Chiangmai	3.12	.93
BMA Medical College	1.56	4.67
Suranaree	5.45	6.23
Naresuan	7.48	7.32
Mahidol	7.94	10.28
Khongan	9.97	14.17

Table 2 represents descriptive statistics comparing the student educational level and the universities in which the MD-ICM and the DIT2 were completed as described by gender.

Table 3. Demographic Information by Sample Described by Age

	Male	Female
	Age	Age
<u>Year</u>		
Year I	1.69 (.67)	1.75 (.70)
Year II	2.95 (.22)	2.93 (.27)
Year III	2.0 (0)	2.0 (0)
<u>Universities</u>		
Thammasart	1.18 (.39)	1.13(.34)
Chiangmai	1.7 (.47)	1.67(.52)
BMA Medical College	1.4 (.52)	1.33(.48)
Suranaree	1(0)	1.00 (0)
Naresuan	1.58 (.90)	1.30(.69)
Mahidol	1.16(.37)	1.11 (.31)
Khongan	1.13(.33)	1.11 (.31)

Age: 1=Age 19-20 year, 2=Age 21-25 year, 3= Age 26-30 year

Table 3 represents descriptive statistics comparing the student age and the universities in which the MD-ICM and the DIT2 were completed as described by gender.

Table 4. Descriptive Statistics of ICM-Scores as Functions of Gender and Educational Level

Gender	Educational Level			
	Year 1	Year 2	Year 3	Total
				<i>N</i> =627
<u>Male</u>	<i>n</i> =89	<i>n</i> =111	<i>n</i> =63	<i>N</i> =263
Total Ranking Index	.17(.19)	.24(.211)	.34(.19)	.24(.21)
Good Ranking Index	.19(.20)	.24(.21)	.34(.19)	.25(.21)
Bad Ranking Index	.14(.23)	.24(.25)	.33(.25)	.23(.26)
Action Choice Index	.19(.24)	.25(.28)	.36(.26)	.26(.27)
Justification Index	.16(.21)	.24(.21)	.32(.18)	.23(.21)
<u>Female</u>	<i>n</i> =128	<i>n</i> =151	<i>n</i> =85	<i>N</i> =364
Total Ranking Index	.28(.18)	.28(.20)	.40(.17)	.31(.19)
Good Ranking Index	.29(.19)	.28(.21)	.42(.17)	.32(.20)
Bad Ranking Index	.25(.22)	.27(.23)	.36(.20)	.29(.22)
Action Choice Index	.34(.23)	.29(.25)	.43(.21)	.34(.24)
Justification Index	.21(.20)	.27(.20)	.36(.17)	.27(.20)

Table 4 represents the descriptive statistics (*Mean and Standard deviation*) of MD-ICM scores described by indices that represent educational level and gender.

### ***Research Question 1***

1. Are MD-ICM scores different for the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> year integrated ethics curriculum?

To answer this research question, three groups of medical students in different years (I, II, III) were tested by using ANOVA analyses to assess the hypothesis that there would be a significant difference between their MD-ICM scores and the MD-ICM scores by educational level and by gender.

The MD-ICM scores represented by five main indices (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index), as the dependent variable, a 2 X 3 analysis of variance (ANOVA) compared mean differences between gender and educational level (i.e., freshman, sophomore, and junior). An alpha level of .05 ( $p < .05$ ) was employed for all statistical tests.

*Total Ranking Index:* The results indicated that, first, there was a significant a main effect for educational level,  $F(2, 621) = 24.385, p < .001, \eta^2 = .073$ . Second, the analysis identified a main effect for gender,  $F(1, 621) = 17.836, p < .001, \eta^2 = .028$ . This finding indicated that females obtained higher Total Ranking Index ICM scores than male. See Figure 1 for a portrayal of the gender difference by educational level. The partial eta squared value indicated that 2.8 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by educational level interaction,  $F(2, 621) = 1.939, p = .145, \eta^2 = .006$ .

A post hoc analysis using the LSD procedure was used to assess which educational levels were statistically different. This test revealed that mean scores of year 3 was statistically greater than year 2 and year 1, respectively.

*Good Ranking Index:* The results of the 2 (gender) X 3 (educational level) ANOVA on the Good ratings showed that there was a significant main effect for educational level,  $F(2, 621)$

= 22.632,  $p < .001$ ,  $\eta^2 = .068$ . Additionally, the analysis identified a main effect for gender,  $F(1, 621) = 19.534$ ,  $p < .001$ ,  $\eta^2 = .030$ . This finding indicated that females obtained higher Good Ranking Index ICM scores than did males. The partial eta squared value indicated that 2.8 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by educational level interaction,  $F(2, 621) = 1.469$ ,  $p = .231$ ,  $\eta^2 = .005$ . Figure 1 presents these findings in a graphical form.

A post hoc analysis using the LSD procedure was used to assess which educational levels were statistically different. Again, this test revealed that mean scores of year 3 was greater than year 2, and year 2 was equal to year 1.

*Bad Ranking Index:* A 2 (gender) X 3 (educational level) ANOVA on the bad ratings indicated a statistically significant main effect for educational level,  $F(2, 621) = 17.865$ ,  $p < .001$ ,  $\eta^2 = .054$ ; and gender,  $F(1, 621) = 8.577$ ,  $p < .05$ ,  $\eta^2 = .014$ . Again, females obtained higher Bad Ranking Index ICM scores than did males. The partial eta squared value indicated that 2.8 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by educational level interaction,  $F(2, 621) = 1.939$ ,  $p = .145$ ,  $\eta^2 = .006$  (See figure 2).

A post hoc analysis using the LSD procedure was used to compare which educational levels were statistically different. Consistent with the previous analyses, this test revealed that mean scores of year 3 was greater than year 2, and year 2 was equal year 1.

*Action Choice Index:* The results of the 2 (gender) X 3 (educational level) ANOVA on the summary of the action choice ratings showed that there was a significant main effect for educational level,  $F(2, 621) = 14.044$ ,  $p < .001$ ,  $\eta^2 = .043$  and gender,  $F(1, 621) = 19.242$ ,  $p < .001$ ,  $\eta^2 = .030$ . Again, females scored higher on this MD-ICM index than did males (See Figure 3). The partial eta squared value indicated that 2.8 % of variation could be accounted for by

gender. In addition, there was no statistically significant gender by educational level interaction,  $F(2, 621) = 3.256, p = .039, \eta^2 = .010$ . As in the previous analyses, the post hoc analysis using the LSD procedure indicated that mean scores of year 3 was greater than year 2 and year 2 was equal to year 1.

*Justification Index:* The results of the 2 (gender) X 3 (educational level) ANOVA on the summary of the justification ratings showed that there was a significant main effect for educational level,  $F(2, 621) = 27.128, p < .001, \eta^2 = .080$  and gender,  $F(1, 621) = 7.453, p < .05, \eta^2 = .012$ . As before, females obtained higher Total Ranking Index ICM scores than did males (see Figure 4). The partial eta squared value indicated that 2.8 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by educational level interaction,  $F(2, 621) = .218, p = .805, \eta^2 = .001$ . Similarly, the post hoc analysis using the LSD procedure revealed that mean scores of year 3 was greater than year 2 and year 1, respectively.

In sum, the results of MD-ICM scores of five dependent variables, a 2 (gender) X 3 (educational level) analysis of variance consistently found significant main effects for gender and for educational level. Eta for the five analyses ranged from .012 to .03 for all of the gender comparisons and from .043 to .08 for the educational level main effects. In addition, all of the analyses indicated that ICM score increased by educational level. Specifically, ICM scores of the 3<sup>rd</sup> year were greater than the 2<sup>nd</sup> year and the 1<sup>st</sup> year students, respectively. However, in only one instance was there a significant effect for gender\*educational level interaction.

Figure 1. Total Ranking Index

Estimated Marginal Means of good and bad choice and justification choices across stories

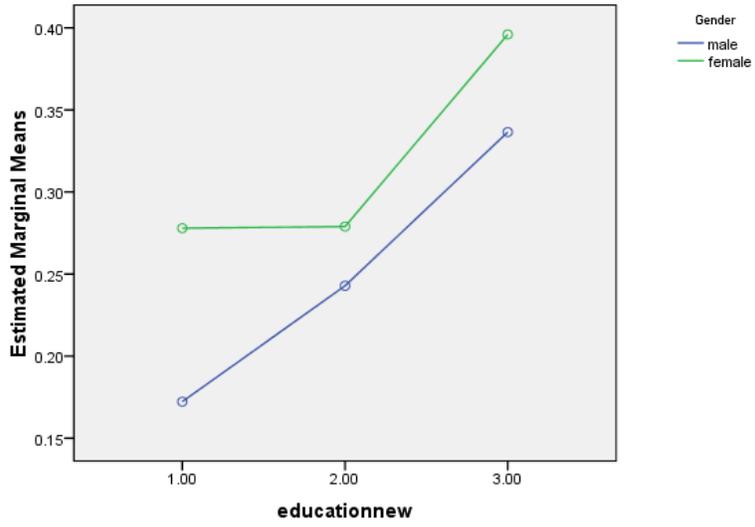


Figure 2. Good Ranking Index

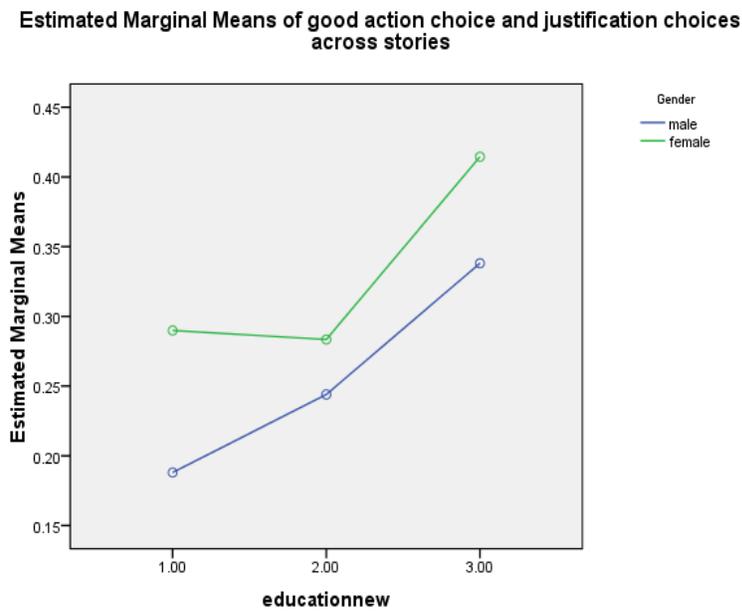


Figure 3. Bad Ranking Index

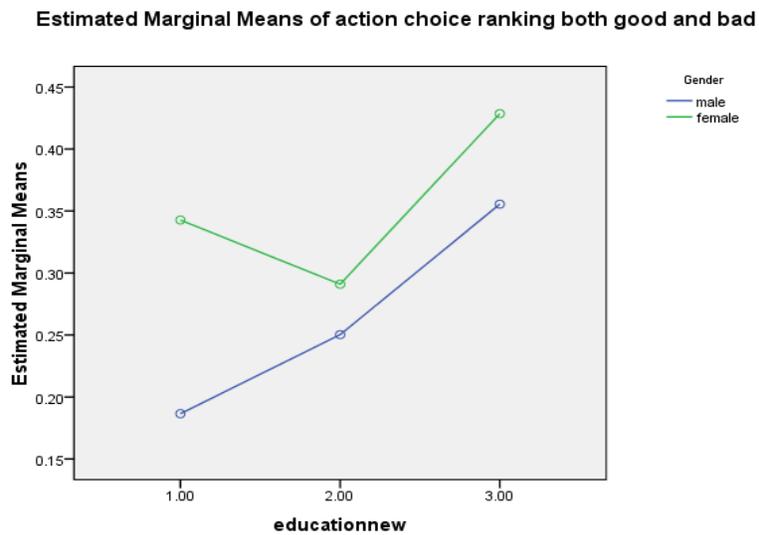


Figure 4. Action Choice Index

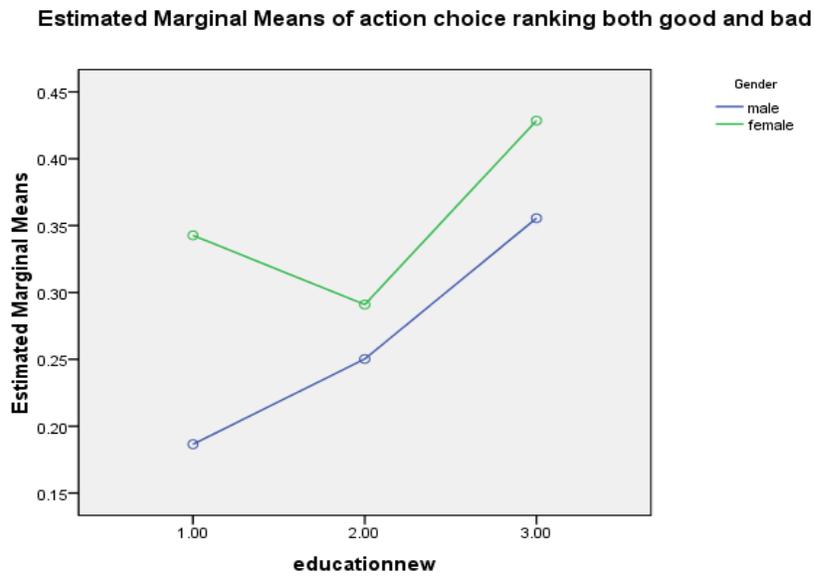
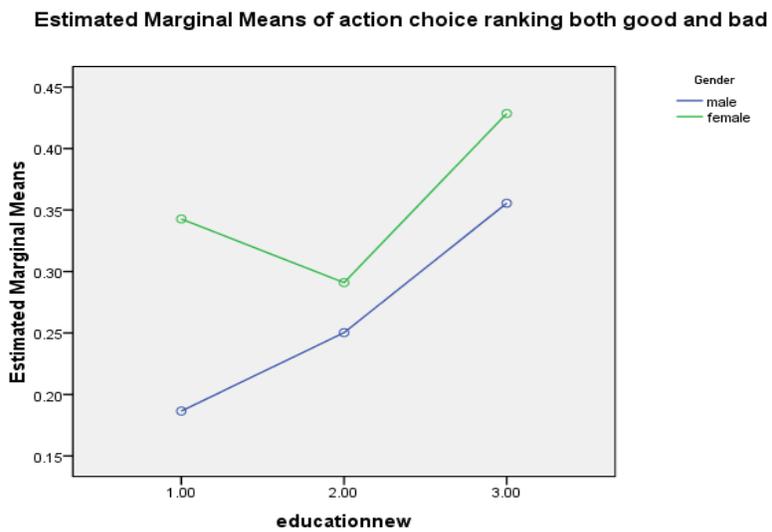


Figure 5. Justification Choice Index



## ***Research Question 2***

2. Do medical students in the integrated ethics program scores higher on the MD-ICM than students in the old style one-class curriculum?

Students from seven universities, medical schools in Thailand (Thammasart, Chiangmai, BMA Medical College, Suranaree, Naresuan, Mahidol, and Khongan), were recruited for this study. All but one of these universities employed a non-integrated ethics program. The other six universities used a new integrated program. Given the difference in programs, the researcher compared the MD-ICM scores between the integrated ethics program and non-integrated ethics program to assess whether the MD-ICM is sensitive to different ethics programs. A 2 (Gender) X 2 (program type) analysis of variance (ANOVA) was used to assess gender and curriculum effects. *Total Ranking Index:* With the first MD-ICM dependent variable, there was a significant a main effect for curriculum difference,  $F(1, 630) = 28.103, p < .001, \eta^2 = .043$  suggested that the non-intergraded program mean scores ( $Mean = .364, SD = .181$ ) was higher than integrated program ( $Mean = .262, SD = .202$ ). Not surprisingly given the findings above, the analysis of variance identified a main effect for gender,  $F(1, 630) = 6.205, p < .05, \eta^2 = .010$ . There was no statistically significant gender by curriculum interaction,  $F(1, 630) = 1.845, p = .175, \eta^2 = .003$ .

*Good Ranking Index:* With the second MD-ICM dependent variable, there was a significant main effect for curriculum difference,  $F(1, 630) = 18.903, p < .001, \eta^2 = .029$  suggested that the non-intergraded program mean scores ( $Mean = .360, SD = .198$ ) was higher than integrated program ( $Mean = .274, SD = .208$ ). Second, the Good Ranking Index of the MD-ICM scores, a 2 (gender) X 2 (curriculum) analysis of variance identified a main effect for gender,  $F(1, 630) = 7.272, p < .05, \eta^2 = .011$ . This can explain that whether a respondent's gender influenced the MD-ICM Good Ranking score. The partial eta squared value indicated

that 1.1 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by curriculum interaction,  $F(1, 630) = 1.304, p = .254, \eta^2 = .002$ .

*Bad Ranking Index:* With the third MD-ICM dependent variable, there was a significant main effect for curriculum difference,  $F(1, 630) = 33.546, p < .001, \eta^2 = .051$  suggested that the non-intergraded program mean scores ( $Mean = .711, SD = .199$ ) was higher than integrated program ( $Mean = .238, SD = .240$ ). Second, the Bad Ranking Index of the MD-ICM scores, a 2 (gender) X 2 (curriculum) analysis of variance identified a main effect for gender,  $F(1, 630) = 2.469, p = .117, \eta^2 = .004$ . This can explain that whether a respondent's gender differences the MD-ICM Bad Ranking score. The partial eta squared value indicated that 0.4 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by curriculum interaction,  $F(1, 630) = 2.058, p = .152, \eta^2 = .003$ .

*Action Choice Index:* With the fourth MD-ICM dependent variable, there was significant a main effect for curriculum difference,  $F(1, 630) = 25.043, p < .001, \eta^2 = .038$  suggested that the non-intergraded program mean scores ( $Mean = .407, SD = .257$ ) was higher than integrated program ( $Mean = .285, SD = .254$ ). Second, the Action Choice Index of the MD-ICM scores, a 2 (gender) X 2 (curriculum) analysis of variance identified a main effect for gender,  $F(1, 630) = 5.906, p < .05, \eta^2 = .009$ . The partial eta squared value indicated that 0.9 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by curriculum interaction,  $F(1, 630) = 3.111, p = .078, \eta^2 = .005$ .

*Justification Index:* With the fifth MD-ICM dependent variable, there was a significant main effect for curriculum difference,  $F(1, 630) = 16.876, p < .001, \eta^2 = .026$  suggested that the non-intergraded program mean scores ( $Mean = .321, SD = .170$ ) was higher than integrated program ( $Mean = .239, SD = .207$ ). Second, the Justification Index of the MD-ICM scores, a 2 (gender) X 2 (curriculum) analysis of variance identified a main effect for gender,  $F(1, 630) =$

3.357,  $p = .067$ ,  $\eta^2 = .005$ . The partial eta squared value indicated that 0.5 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by curriculum interaction,  $F(1, 630) = .204$ ,  $p = .652$ ,  $\eta^2 = .000$ .

In sum, results of the analysis focusing on curriculum difference on MD-ICM scores indicates a significant a main effect for curriculum difference. The direction of the effect suggests that the non-integrated program mean score was higher than the integrated programs for all five independent variables (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index). In addition, there is significant a respondent's gender difference. However, there was no statistically significant gender by curriculum interaction.

Next, we assessed whether there were gender, university, and/or interaction effects on MD-ICM scores. To explore university differences, we broke differences down the sample by the three educational levels (1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup>) and type of institution. . A post hoc analysis was also tested for this research question. These results will be described by each dependent variable.

*Total Ranking Index*: With the first MD-ICM dependent variable, there was a significant main effect for different universities,  $F(6, 620) = 20.735$ ,  $p < .001$ ,  $\eta^2 = .167$ . Second, the *Total Ranking Index* of the MD-ICM scores, a 2 (gender) X 2 (universities) analysis of variance identified a main effect for gender,  $F(1, 620) = 3.905$ ,  $p < .05$ ,  $\eta^2 = .006$ . This can explain that whether a respondent's gender differences the MD-ICM Total Ranking score. The partial eta squared value indicated that 0.6 % of variation could be accounted for gender. In addition, there was no statistically significant gender by different university interaction,  $F(6, 620) = .954$ ,  $p = .456$ ,  $\eta^2 = .009$ .

Because the main effect for university was significant, a post hoc analysis using the Student–Newman-Keuls procedure was used to compare which universities accounted for the observed difference. This test revealed that mean scores of BMA Medical College was greater than Mahidol, Chiangmai, Khongon and these were in turn greater than Suranaree and Naresuan, respectively.

*Good Ranking Index:* With the second MD-ICM dependent variable, there was a significant main effect for different universities,  $F(6, 620) = 16.850, p < .001, \eta^2 = .140$ . Second, the *Good Ranking Index* of the MD-ICM scores, a 2 (gender) X 2 (universities) analysis of variance, identified a main effect for gender,  $F(1, 620) = 3.613, p = .058, \eta^2 = .006$ . This can explain the respondent's gender differences for the MD-ICM Good Ranking score. The partial eta squared value indicated that 0.6 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by different university interaction,  $F(6, 620) = .724, p = .630, \eta^2 = .007$ .

Because the main effect for university was significant, a post hoc analysis using the Student–Newman-Keuls procedure was used to compare which universities accounted for the observed difference. Again, a post hoc analysis revealed that mean scores of BMA Medical College was greater than Mahidol, Chiangmai, and Thammasart, and also was greater than Khongon, Suranaree and Naresuan, respectively.

*Bad Ranking Index:* With the third MD-ICM dependent variable, there was significant a main effect for different universities,  $F(6, 620) = 18.823, p < .001, \eta^2 = .154$ . Second, the Bad Ranking Index of the MD-ICM scores, a 2 (gender) X 2 (universities) analysis of variance, identified a main effect for gender,  $F(1, 620) = 2.623, p = .106, \eta^2 = .004$ . This can explain the respondent's gender differences for the MD-ICM Bad Ranking score. The partial eta squared value indicated that 0.4 % of variation could be accounted for by gender. In addition, there was

no statistically significant gender by different university interaction,  $F(6, 620) = 2.036, p = .059, \eta^2 = .019$ .

For the third MD-ICM dependent variable, because the main effect for university was significant, a post hoc analysis using the Student–Newman-Keuls procedure was used to compare which universities accounted for the observed difference. Similarly, a post hoc analysis revealed that mean scores of BMA Medical College was greater than Mahidol, Chiangmai, and Thammasart, and also was greater than Khongon, Suranaree and Naresuan, respectively.

*Action Choice Index:* With the fourth MD-ICM dependent variable, there was significant a main effect for different universities,  $F(6, 620) = 18.419, p < .001, \eta^2 = .151$ . Second, the *Action Choice Index* of the MD-ICM scores, a 2 (gender) X 2 (universities) analysis of variance identified a main effect for gender,  $F(1, 620) = 3.519, p = .061, \eta^2 = .006$ . This can explain the respondent's gender differences for the MD-ICM Action Choice Index score. The partial eta squared value indicated that 0.6 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by different university interaction,  $F(6, 620) = 1.157, p = .328, \eta^2 = .011$ .

For the *Action Choice Index*, because the main effect for university was significant, a post hoc analysis using the Student–Newman-Keuls procedure was used to compare which universities accounted for the observed difference. Again, a post hoc analysis revealed that mean scores of BMA Medical College and Mahidol was grater than Thammasart and Chiangmai, and also was greater than Khongon, Suranaree and Naresuan, respectively.

*Justification Index:* With the fifth MD-ICM dependent variable, there was a significant main effect for different universities,  $F(6, 620) = 14.624, p < .001, \eta^2 = .124$ . Second, the *Justification Index* of the MD-ICM scores, a 2 (gender) X 2 (universities) analysis of variance, identified a main effect for gender,  $F(1, 620) = 2.172, p = .141, \eta^2 = .003$ . This can explain the

respondent's gender differences for the MD-ICM Justification Index score. The partial eta squared value indicated that 0.3 % of variation could be accounted for by gender. In addition, there was no statistically significant gender by different university interaction,  $F(6, 620) = .687$ ,  $p = .660$ ,  $\eta^2 = .007$ .

For the *Justification Index*, because the main effect for university was significant, a post hoc analysis using the Student–Newman-Keuls procedure was used to compare which universities accounted for the observed difference. Again, a post hoc analysis revealed that mean scores of BMA Medical College was greater than Chiangmai and Mahidol, and also was greater than Thammasart, Khongan, Naresuan and Suranaree, respectively.

In sum, there is a significant a main effect for different university on MD-ICM scores. As expected there is also significant gender difference. However, there was no statistically significant gender by university interaction. In addition, a post hoc analysis using the Student–Newman-Keuls procedure was used to compare which universities accounted for the observed difference. For overall dependent variables, a post hoc analysis revealed that mean scores of BMA Medical College was greater than Chiangmai and Mahidol, and also was greater than Thammasart, Khongan, Naresuan and Suranaree, respectively.

### ***Research Question 3***

3. Does the MD-ICM relate to established measure of moral thinking such as the DIT as represented by the N2 and U scores and type variable?

This research question compares the relationship between the DIT 2 score and the MS-ICM scores. A correlation analysis compared five dependent variables (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index), and DIT2 scores (i.e., Utilize score, N2SCORE, and Typenew). The results of the correlation analysis ( $n = 540$ ) comparing the five dependent variables on the MD-ICM scores and the DIT2

scores can be found in Table 3 below. The descriptive statistics is shown in Table 3. Among the MD-ICM scores, (Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index), the Action Choice Index is the highest mean score ( $M = .334$ ,  $SD = .257$ ). The Justification Index is the lowest mean score ( $M = .274$ ,  $SD = .196$ ). Among, the DIT2 scores, (Utilize score, N2SCORE, and Typenew), the results indicated that the means score of N2SCORE score is  $M = 22.149$  ( $SD = 9.085$ ), which is lower than expected for a college aged population.

Further correlation analysis between the MD-ICM scores and the DIT2 scores revealed significant relationships. The *Pearson product moment correlation coefficient* value ranged from .09 ( $r$ ) to .24 ( $r$ ). The important relationships with the DIT summary score, N2, were all statistically significant and in the expected positive direction. The results are shown in Table 6.

In sum, to address this research question that does the MD-ICM relate to established measure of moral thinking such as the DIT score? The results revealed that simple correlations with the DIT confirmed the expectation that the two measures should be positively related.

Table 5. Descriptive Statistics

	Mean	S.D.	N
Total Ranking Index	.304	.198	540
Good Ranking Index	.314	.206	540
Bad Ranking Index	.285	.235	540
Action Choice Index	.334	.257	540
Justification Index	.274	.196	540
N2SCORE	22.149	9.085	548
Pscore	22.000	10.061	548
Maintain Norms	28.854	10.513	548

Table 6. Simple Correlation ( $r$ ) between the MD-ICM and the DIT2

	1.	2.	3.	4.	5.	6.	7.	8.
1. Total Ranking	1							
2. Good Index	.96**	1						
3. Bad Index	.86**	.66**	1					
4. Action Choice	.91**	.87**	.77**	1				
5. Justification	.83**	.79**	.72**	.52**	1			
6. Pscore	.17**	.19**	.09*	.21**	.06	1		
7. N2score	.22**	.22**	.18**	.24**	.13**	.74**	1	
8. Maintain Norms	.21**	.21**	.17**	.17**	.21**	-.11**	.09*	1

\* =  $p < .05$ , \*\* =  $p < .01$

#### ***Research Question 4***

4. Do MD-ICM scores relate to DIT2 and Gender variables individually and in combination?

To address these questions a series of hierarchical regression analyses were conducted on the five dependent variables (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index), and four independent variables (i.e., Gender, Pscore, N2score, and STAGE4P). In each case the structure of analysis was similar: in step one of the hierarchy the ICM score was regressed on gender and the moral judgment score followed by in step two, a cross-product term, which assessed a nonadditive relationship between ICM, gender and moral judgment. In other words, this relationship addressed the question of whether the relationship between ICM and DIT scores vary by gender. These analyses presented below are blocked by the corresponding dependent variables.

### ***Total Ranking Index***

*Post-conventional score (Pscore):* In steps one, Total Ranking Index was regressed onto the two predictors (i.e., Gender and Pscore). Together, the set of predictors explained 4.4% of the variance in MD-ICM scores,  $F(2,537) = 12.29, p < .01$ . The beta weight assessed with Gender and the DIT Pscore were significant ( $t = 2.30, p < .01$  and  $t = 3.87, p < .01$ ) for gender and Pscore respectively. However, in the second step, there was not a significant interaction between Gender \* Pscore ( $t = .30, p = .76$ ) indicated that the relationship between Pscore and Total Ranking Index did not vary by gender.

*N2SCORE:* In steps one, Total Ranking Index was regressed onto the two predictors (i.e., Gender and N2score). Together, the set of predictors explained 6.5% of the variance in MD-ICM scores,  $F(2,537) = 18.79, p < .01$ . The beta weight assessed with Gender and the DIT N2SCORE were significant ( $t = 3.05, p < .01$  and  $t = 5.27, p < .01$ ) for gender and N2score respectively. However, in the second step, there was not a significant interaction between Gender \* N2SCORE ( $t = .07, p = .95$ ) indicated that the relationship between N2score and Total Ranking Index did not vary by gender.

*Maintain Norms (STAGE4P):* In steps one, Total Ranking Index was regressed onto the two predictors (i.e., Gender and STAGE4P). Together, the set of predictors explained 6.1% of the variance in MD-ICM scores,  $F(2,537) = 17.31, p < .01$ . The beta weight assessed with Gender and the DIT STAGE4P were significant ( $t_{\text{Gender}} = 4.98, p < .01$  and  $t = 3.00, p < .01$ ) for gender and STAGE4P respectively. However, in the second step, there was not a significant interaction between Gender \* STAGE4P ( $t = -.50, p = .62$ ) indicated that the relationship between STAGE4P and Total Ranking Index did not vary by gender.

### ***Good Ranking Index***

*Post-conventional score (Pscore)*: In steps one, Good Ranking Index was regressed onto the two predictors (i.e., Gender and Pscore). Together, the set of predictors explained 5.3 % of the variance in MD-ICM scores,  $F(2,537) = 15.07, p < .01$ . The beta weight assessed with Gender and the DIT Pscores were significant ( $t = 3.19, p < .01$  and  $t = 4.38, p < .01$ ) for gender and Pscore respectively. However, in the second step, there was not a significant interaction between Gender \* Pscore ( $t = .28, p = .78$ ) indicated that the relationship between Pscore and Good Ranking Index did not vary by gender.

*N2SCORE*: In steps one, Good Ranking Index was regressed onto the two predictors (i.e., Gender and N2SCORE). Together, the set of predictors explained 6.5% of the variance in MD-ICM scores,  $F(2,537) = 18.79, p < .01$ . The beta weight assessed with Gender and the DIT N2SCORE were significant ( $t = 3.25, p < .01$  and  $t = 5.14, p < .01$ ) for gender and N2SCORE respectively. However, again, in the second step, there was not a significant interaction between Gender \*N2SCORE ( $t = .31, p = .76$ ) indicated that the relationship between N2SCORE and Good Ranking Index did not vary by gender.

*Maintain Norms (STAGE4P)*: In steps one, Good Ranking Index was regressed onto the two predictors (i.e., Gender and STAGE4P). Together, the set of predictors explained 6.1% of the variance in MD-ICM scores,  $F(2,537) = 17.66, p < .01$ . The beta weight assessed with Gender and the DIT STAGE4P were significant ( $t = 4.93, p < .01$  and  $t = 3.20, p < .01$ ) for gender and STAGE4P respectively. However, in the second step, there was not a significant interaction between Gender \* STAGE4P ( $t = -.54, p = .59$ ) indicated that the relationship between STAGE4P and Good Ranking Index did not vary by gender.

### ***Bad Ranking Index***

*Post-conventional score (Pscore)*: In steps one, Bad Ranking Index was regressed onto the two predictors (i.e., Gender and Pscore). Together, the set of predictors explained 1.6 % of the variance in MD-ICM scores,  $F(2,537) = 4.31, p < .05$ . The beta weight assessed with Gender and the DIT Pscores were significant ( $t = 1.98, p < .05$  and  $t = 2.11, p < .05$ ) for gender and Pscore respectively. However, in the second step, there was not a significant interaction between Gender \*Pscore ( $t = .28, p = .78$ ) indicated that the relationship between Pscore and Bad Ranking Index did not vary by gender.

*N2SCORE*: In steps one, Bad Ranking Index was regressed onto the two predictors (i.e., Gender and N2SCORE). Together, the set of predictors explained 6.5% of the variance in MD-ICM scores,  $F(2,537) = 18.79, p < .01$ . The beta weight assessed with gender and the DIT N2SCORE were significant ( $t = 3.25, p < .01$  and  $t = 5.14, p < .01$ ) for gender and N2SCORE respectively. However, in the second step, there was not a significant interaction between Gender \*N2SCORE ( $t = .31, p = .76$ ) indicated that the relationship between N2SCORE and Bad Ranking Index did not vary by gender.

*Maintain Norms (STAGE4P)*: In steps one, Bad Ranking Index was regressed onto the two predictors (i.e., Gender and STAGE4P). Together, the set of predictors explained 3.5 % of the variance in MD-ICM scores,  $F(2,537) = 9.81, p < .01$ . The beta weight assessed with Gender and the DIT STAGE4P were significant ( $t = 3.92, p < .05$  and  $t = 1.96, p < .05$ ) for gender and STAGE4P respectively. However, in the second step, there was not a significant interaction between Gender \* STAGE4P ( $t = -.32, p = .75$ ) indicated that the relationship between STAGE4P and Bad Ranking Index did not vary by gender.

### ***Action Choice Index***

*Post-conventional score (Pscore)*: In steps one, Action Choice Index was regressed onto the two predictors (i.e., Gender and Pscore). Together, the set of predictors explained 6.4 % of the variance in MD-ICM scores,  $F(2,537) = 18.45, p < .01$ . The beta weight assessed with Gender and the DIT Pscores were significant ( $t = 3.25, p < .01$  and  $t = 5.05, p < .01$ ) for gender and Pscore respectively. However, in the second step, there was not a significant interaction between Gender \*Pscore ( $t = .020, p = .98$ ) indicated that the relationship between Pscore and Action Choice Index did not vary by gender.

*N2SCORE*: In steps one, Action Choice Index was regressed onto the two predictors (i.e., Gender and N2SCORE). Together, the set of predictors explained 7.6 % of the variance in MD-ICM scores,  $F(2,537) = 22.08, p < .01$ . The beta weight assessed with Gender and the DIT N2SCORE were significant ( $t = 3.31, p < .01$  and  $t = 5.71, p < .01$ ) for gender and N2SCORE respectively. However, in the second step, there was not a significant interaction between Gender \*N2SCORE ( $t = .49, p = .62$ ) indicated that the relationship between N2SCORE and Action Choice Index did not vary by gender. .

*Maintain Norms (STAGE4P)*: In steps one, Action Choice Index was regressed onto the two predictors (i.e., Gender and STAGE4P). Together, the set of predictors explained 4.7 % of the variance in MD-ICM scores,  $F(2,537) = 13.29, p < .01$ . The beta weight assessed with Gender and the DIT STAGE4P were significant ( $t = 3.91, p < .01$  and  $t = 3.25, p < .01$ ) for gender and STAGE4P respectively. However, in the second step, there was not a significant interaction between Gender \* STAGE4P ( $t = -.75, p = .45$ ) indicated that the relationship between STAGE4P and Action Choice Index did not vary by gender.

## Justification Index

*Post-conventional score (Pscore)*: In steps one, Justification Index was regressed onto the two predictors (i.e., Gender and Pscore). Together, the set of predictors explained .09 % of the variance in MD-ICM scores,  $F(2,537) = 2.47, p = .09$ . The beta weight assessed with Gender and the DIT Pscores were not significant ( $t = 1.81, p = .07$  and  $t = 1.24, p = .22$ ) for gender and Pscore respectively. However, in the second step, there was not a significant interaction between Gender \*Pscore ( $t = .57, p = .57$ ) indicated that the relationship between Pscore and Justification Index did not vary by gender.

*N2SCORE*: In steps one, Action Choice Index was regressed onto the two predictors (i.e., Gender and N2SCORE). Together, the set of predictors explained 2.4 % of the variance in MD-ICM scores,  $F(2,537) = 6.60, p < .01$ . The beta weight assessed with Gender and the DIT Pscores were significant ( $t = 1.80, p = .07$  and  $t = 3.12, p < .01$ ) for gender and N2SCORE respectively. However, in the second step, there was not a significant interaction between Gender \*N2SCORE ( $t = .57, p = .57$ ) indicated that the relationship between N2SCORE and Justification Index did not vary by gender.

*Maintain Norms (STAGE4P)*: In steps one, Justification Index was regressed onto the two predictors (i.e., Gender and STAGE4P). Together, the set of predictors explained 4.8 % of the variance in MD-ICM scores,  $F(2,537) = 13.60, p < .01$ . The beta weight assessed with Gender and the DIT STAGE4P were significant ( $t = 4.87, p < .01$  and  $t = 1.76, p = .08$ ) for gender and STAGE4P respectively. However, in the second step, there was not a significant interaction between Gender \* STAGE4P ( $t = .57, p = .57$ ) indicated that the relationship between STAGE4P and Justification Index did not vary by gender.

In sum, the results from this analysis found that the relationship between MD-ICM scores ( $N = 540$ ) and moral scores do not vary by gender. All independent variables (i.e., Gender,

Pscore, N2CORE and STAGE4P) were uniquely related to all MD-ICM scores. A summary of hierarchical regression analysis for variables and interactions predicting MD-ICM score ( $N = 540$ ) will be presented on Table 7. Full table for regression analyses will be in appendixes.

Table 7. Summary of Hierarchical Regression Analysis for Variables and Interactions Predicting MD-ICM Score ( $N = 540$ )

	Total ICM	Good	Bad	Action	Justification
	$\beta$	$\beta$	$\beta$	$B$	$\beta$
Gender	.13**	.13**	.09*	.14**	.08
Pscore (MJ)	.16**	.18**	.09*	.21**	.05
Gender x Pscore	.03	.03	.03	.00	.07
Model 1. $R^2$ Change	.04**	.05**	.02*	.06**	.01
Model 2. $R^2$ Change	.00	.00	.00	.00	.00
Gender	.13**	.14**	.08	.14**	.08
N2score (MJ)	.22**	.21**	.18**	.24**	.13**
Gender x N2score	.01	.04	-.04	-.03	.06
Model 1. $R^2$ Change	.07**	.07**	.04**	.08**	.02**
Model 2. $R^2$ Change	.00	.00	.00	.00	.00
Gender	.21**	.21**	.17**	.17**	.21**
STAGE4P (MJ)	.13**	.13**	.08*	.14**	.07
Gender x STAGE4P	-.07	-.07	-.04	-.10	-.00
Model 1. $R^2$ Change	.06**	.06**	.04**	.05**	.05**
Model 2. $R^2$ Change	.00	.00	.00	.00	.00

MJ = Moral Judgment,  $R^2$  = R square, \* =  $p < .05$ , \*\* =  $p < .01$

In conclusion for this chapter: the findings to answer four research questions were addressed as follows:

1. Are MD-ICM scores different for the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> year integrated ethics curriculum?

The results found that three groups of medical students in different years (I, II, III) were significant. A respondent's gender difference found that specifically female obtained higher on the MD-ICM scores than male. Furthermore, the ICM score showed the increase for the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> year. The ICM score of the 3<sup>rd</sup> year was greater than the 2<sup>nd</sup> year and the 1<sup>st</sup> year, respectively. However, there were no significant effect for gender\*educational level interaction.

2. Do medical students in the integrated ethics program scores higher on the MD-ICM than students in the old styles one-class curriculum? To answer this research question, students from seven universities, medical schools in Thailand (Thammasart, Chiangmai, BMA Medical College, Suranaree, Naresuan, Mahidol, and Conan), were recruited for this study. All but one of these universities employed a non-integrated ethics program. Six universities used new integrated program. The results found that the non-integrated program mean score was higher than integrated program for all five independent variables. There was significant a respondent's gender difference. However, there was no statistically significant gender by curriculum interaction. Next, to explore which university difference, we broke down three educational levels (1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup>) by type of institution. We found that the main effect for different universities was significant. These results found that for overall dependent variables, a post hoc analysis revealed that mean scores of BMA Medical College was greater than Chiangmai and Mahidol, and also was greater than Thammasart, Khongan, Naresuan and Suranaree, respectively.

3. Does the MD-ICM relate to established measure of moral thinking such as the DIT?

This research question compares the relationship between the DIT 2 score and the MS-ICM

scores. A correlation analysis compared five dependent variables. The results found that simple correlation was moderate to small in magnitude and in the expected direction .

4. Do MD-ICM scores relate to DIT2 and Gender variables individually and in combination? The series of hierarchical regression analysis was conducted on the five depending variable and on the four independent variables. In steps one of the hierarchies the ICM score was regressed on gender and the moral judgment score followed by in step two, a cross-product term, which assessed a relationship between ICM, gender and moral judgment. The results found that the relationship between ICM and moral scores did not vary by gender.

## CHAPTER V

### DISCUSSION

The purpose of this study was to develop and test a new instrument, the MD-ICM, designed to help educators assess the ethical reasoning of medical students.

Thai medical education needs an empirically validated tool to assess ethical reasoning in medicine. The development and validation of the MD-ICM completed in this study is an attempt to provide one such measure of ethical reasoning. This study was modeled on Bebeau and Thoma's (1999) intermediate concepts measure (ICM), which uses realistic cases that are specific to the medical field. The MD-ICM relies on the thinking and judgment of medical experts for the items used on the test and scoring key. The ICM reflects the principles in the content of professional ethics curriculum such as informed consent, confidentiality and so forth. The ICM design is flexible enough to address an informal curriculum in medical school that influences the moral decision making of medical students.

In this discussion, the author shall address four research questions in the order in which they were presented in Chapter 1.

*The first research question was, "Are MD-ICM scores different for the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> year integrated ethics curriculum"?* The data analysis showed that three groups of medical students in different years (I, II, III) scored significantly different on MD-ICM. Furthermore, the MD-ICM score showed the increase for the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> year. The MD-ICM score of the 3<sup>rd</sup> year was greater than the 2<sup>nd</sup> year and the 1<sup>st</sup> year, respectively. This study found that MD-ICM

scores for female were higher than scores for male. It was also noted that, there was no significant effect for gender\*educational level interaction.

Discussion of the findings of the first research question is addressed in three ways; 1) In terms of the number of years in the program, 2) In terms of the measurement properties of the MD-ICM, and 3) In terms of gender difference.

First, it was found that MD-ICM score increased with the number of years in the program. One possible explanation is that the 3<sup>rd</sup> year medical students have much more experience with the medical profession than do the 2<sup>nd</sup> and 1<sup>st</sup> year medical students. In support of this claim, the Thai medical curriculum includes *basic sciences* in the first year and medical ethics and certain classes specific in basic medical science in the second and in-depth clinical practice (e.g., disease, principle of clinical skills) in the third years. For example, in Chulalongkhon University, Thailand, the first year medical students have to complete foundation courses (e.g., English for Medical Profession, ,etc.), and a course each from Sciences and Mathematics group (3 credits), some Social Sciences group (3 credits), Humanities group (3 credits), Integrated Sciences group (3 credits), and 6 credits from Free electives. By contrast, medical students in the second year have to take Introduction to Basic Medical Sciences, Normal structures and function (system-based integration), Medical Ethics and Critical Thinking, Metabolism and Nutrition, Embryology and Genetics in Medicine, and The Cycle of Life. Third year medical students have to take Principles of Clinical Skills, Principles of Health and Diseases, Abnormal structures and function (system-based integration), and Medical Electives. As a result of this program structure, the researcher would expect that medical students in year III would get more experience in clinical practice and medical ethics knowledge in applying in medical practice as compared to those in years I and II. Third year medical students have experienced medical dilemmas knowledge in more depth than 1<sup>st</sup> and 2<sup>nd</sup> year medical students.

If the ICM taps ethical knowledge in the context of the medical profession then year III students should score the highest on the measure than students in year I ( who have neither medical ethics training) and year II (who have ethics training but limited experience within the profession).

The findings of this study support this contention.

More generally, Self and Baldwin (1994) argue that doctors and other health care providers face ethical dilemmas in daily practice and one should not find real advances in the quality of reasoning until students experience daily practice. The findings also support Self and Baldwin (1994) because third year students have gone through more clinical practice knowledge than first and second year students, such that they have significantly more hours of daily practice knowledge.

Second, it should be noted that in finding group differences in the expected direction, the results related to hypothesis one also support the validity of the measure. That is, an ICM measure that assesses medical ethics understanding and judgment ought to distinguish groups who differ in their training in medical ethics and experience in medical practice. Finding that mean scores of year 3 students was statistically greater than year 2 and year 1, respectively is an example of this “known group differences” validity criteria (See Thoma, 2006).

In addition to supporting the MD-ICM the findings of the current study support the overall intermediate concepts approach to the measurement of ethical thinking. The instrument developed in this study, the MD-ICM, supports the preliminary work with the Dental Ethical Reasoning And Judgment Test (DERJT). Both measures highlight the utility of the ICM approach both as a diagnostic tool and as a measure of medical ethics interventions (Bebeau and Thoma, 1999). The MD-ICM posed medical dilemmas from daily practice concepts similar to the study by Thoma (2006) who developed the AD-ICM, a measure of adolescent thinking. The AD-ICM was done by surveying the content of character education programs for core concepts

(e.g. honesty, fairness, tolerance, self-discipline, etc.), similar to concepts found in the MD-ICM in medical doctor group (e.g. informed consent, confidentiality, etc.). Finally, this study is consistent with Turner (2008) who developed and tested the ALERT (An Army Leader Intermediate Ethical Concept Measure), an ICM-based approach to the study of military ethics. Turner's findings showed evidence that the ALERT had construct and convergent validity.

Third, the finding of gender differences in this study deserves mention. Specifically, findings from this study indicate that females obtained higher ICM scores as represents by the five main MD ICM indices (i.e., Total Ranking Index, Good Ranking Index, Bad Ranking Index, Action Choice Index and Justification Index) than males. Further, these differences were consistent across class level given that there were no statistically significant gender\*educational level interaction effects.

With regards to females obtained higher ICM scores than males, it is noted that the results of this study are consistent with the gender differences found on other ICM measures (e.g., Turner, 2008; Thoma, 2006; Bebeau, 2002). Further, these results are consistent with other measures of moral judgment development notably the DIT, where when gender differences are found the overwhelming trend is that the differences favor females (e.g., Thoma, 1986, Bebeau, 2002). It may be that medical educators and preceptors should pay more attention to male students both in the classroom as well as clinical practice, given the observed gender difference.

In short, the findings from research question one are 1) The MD-ICM score increased with the number of years in the program, 2) The findings from the study present the measurement properties of the MD-ICM that supports construct validity of the measure, and 3) The findings from this study indicated significant gender difference. From these findings we can conclude that the MD-ICM is sensitive to the growth of moral reasoning development associated with the number of years one spends in the medical school.

*The second research question was, “Do medical students in the integrated ethics program score higher on the MD-ICM than students in the more traditional single-class ethics curriculum?”* The Medical students from seven universities in Thailand (Thammasart, Chiangmai, BMA Medical College, Suranaree, Naresuan, Mahidol, and Khongan), were recruited for this study. All but one of these universities employed a non-integrated ethics program. The other six universities used the new integrated program of ethics training. The current study indicates that in contrast to expectations, the non-integrated program mean score was higher than the scores associated with the newer integrated program on all five of the independent variables.

Interpreting these contrary findings in the absence of additional information is difficult. However, one possible explanation for this contrary finding is that Mahidol University is the oldest, most prestigious institution of higher learning in Thailand, internationally known and recognized for high quality research and teaching by its faculty, and outstanding achievements in teaching, research, international academic collaboration and professional service. Further, Mahidol University is located in Bangkok, the capital of Thailand a very desirable location that attracts the top students. The possibility that prestige and location are implicated in these results is further supported by the observation that other medical schools who were also located in desirable urban settings achieved significantly higher MD-ICM scores than the more rural and regional schools

In addition to prestige and location as explanatory factors in the ICM score difference by institution, it may be that the instructors in these schools are different as well. Just as students wish to attend the more prestigious schools in popular areas, doctors who serve on the faculty may be influenced in a similar manner. In Thailand, doctors who serve as preceptors (i.e., clinical instructor) are recruited to teach based on their ability. It is known that the highest

scoring medical doctors are recruited to serve in the urban areas and average medical doctors are more likely to serve in rural areas. One could argue, therefore, that the quality of instruction may also be different in the more urban and prestigious universities in contrast to the less desirable rural schools. By extension, students medical students who experience better preceptors may also develop a deeper understanding of ethical issues which then is reflected in higher scores on MD-ICM.

More generally, the finding that medical students who are most selective score higher than other students, is similar to results obtained using other ethical reasoning measures (e.g., the DIT, Rest, 1979) and supports the cognitive foundation of ethical reasoning. That is, ethical reasoning is claimed to be a cognitive process that is, in turn, related to cognitive skills. Thus, students who are selected for high cognitive skills should score higher on measures of ethical reasoning, which includes the MD-ICM.

As stated in literature review, the results in this study also support other studies showing regional differences (Bebeau, 2002). For example, Bebeau and Thoma (1994) and Bebeau (2001) summarized scores across 15 cohorts and found that the average *P*-score for entering University of Minnesota students was greatly higher than the mean for freshmen entrance scores at UCLA. Bebeau (2002) explained that the observed differences may possibly be the result of cultural mix.

More importantly, many researchers are concerned that moral reasoning is not developing normally during medical education (e.g., Patenaude, Niyonsenga, and Fafard, 2003). Although the generally trends of this study contradict this view, it may be that the Thai integrated ethics program is similar to the medical education programs criticized by Patenaude et al. (2003). For instance, the program may be too subtle to be noticed by students or perhaps it is resisted by the professors that resent giving up content time to add the ethical components. In any case the

integrated program deserves some scrutiny since it is not able to match the gains made by students in the more traditional programs.

In sum, the analysis of the data of research question two indicated that the non-integrated program mean score was higher than integrated program on all five independent variables. Clearly Thai medical researchers should pay more attention to the content of the ethics curriculum under the integrated model. To this end, researchers might benefit from a review of the intent of the ethics curricula. According to the aim of medical ethics education, Goldie, Schwartz, McConnachie, & Morrison (2003, p.109) asserted that medical student's abilities are, for example, to recognize the humanistic and ethical aspect of medical professional, to examine and affirm their own personal and professional moral commitment, to equip with a foundation of philosophical, social and legal knowledge, to utilize this knowledge in clinical reasoning. Following this view, medical educators should assess the content of medical ethics curriculum in relation to "how much curriculum time need to be devoted to ethics instruction to influence moral development" (Bebeau, 2002, p. 273).

*The third research question* asked, "Does the MD-ICM relate to established measure of moral thinking such as the DIT"? This question focused on the relationship between the DIT-2 scores and the MD-ICM scores. A correlation analysis was done to compare five dependent variables. The primary relationship between summary scores was  $r = .24$ . The small but positive relationships can be viewed as supporting the *convergent validity* of the MD-ICM. That is, it appears that the MD-ICM is related to the DIT and thus is measuring similar phenomena within the moral domain. However, the size of the correlation does not suggest that the information provided by the AD-ICM is redundant to the DIT and thus, provides independent, non-overlapping information.

Although the relationships between the MD-ICM and DIT were in the expected direction, the magnitude of the relationship was lower than anticipated. One possible reason to explain why the relationship between the MD-ICM and the DIT 2 score was low is the use of English in the assessments. Thai medical students use Thai as their primary language, while the tests (DIT1 and MD-ICM) are conducted in English. It is well-known that the DIT2 is a complex cognitive task (e.g., Rest, 1996), and the difficulties in processing the information may have taxed the students' English skills. Although English is used in text books, Thai medical education is conducted primarily in the Thai language. Therefore, it might have been difficult for Thai medical students to complete the MD-ICM and the DIT2 in English. In support of this contention, some Thai medical students informally mentioned that they would have preferred the test in the Thai language.

Cultural differences also might have played a role in poor comprehension of MD-ICM and the DIT2. Even though the DIT2 is a standardized test and has been widely used, cultural differences might have hindered the ability of Thai medical students to adequately represent their reasoning about ethical issues. It is possible that Thai medical students might not have completely understood certain ethical dilemmas initiated in western culture, because they may not have seen such dilemmas in Thai context. Nonetheless, Bebeau (2002) suggested that to use subjects whose English is not their primary language requires researchers to carefully consider making any inferences about cultural differences.

According to empirical reported by Husted (1978) the empirical article supports the idea that Thai medical students may not do well on the test because of cultural differences. Husted (1978) asserted that the mean *P*-score for U.S. trained residents ( $n=46$ , mean = 57.2) was considerably higher than the mean *P*-score for foreign-trained residents ( $n=58$ , mean =32.3). This is because the DIT is greatly dependent on reading skills (Bebeau, 2002). Studied using

nursing groups by Nokes (nd) found that the DIT scores for American-trained nurses were higher than foreign-born nurses.

Although the DIT 2 requires an English proficiency, it is interesting to note that English proficiency of Thai medical students was reported to be high. It was therefore assumed that they could read at the required 9<sup>th</sup> grade level required by the DIT and are aware of basic US English usage. Consequently, the DIT2 was not translated from English to Thai. In short, this decision may have been a mistake.

*Research question four asked, “Do MD-ICM scores relate to DIT2 and Gender variables individually and in combination?”* Series of hierarchical regression analyses was conducted on the five dependent variables and the four independent variables. In step one of the hierarchies, the ICM score was regressed on gender and the moral judgment score. Next, in step two, a cross-product term was entered, which assessed the relationship between ICM, gender, and moral judgment. The results showed that the relationship between ICM and moral scores did not vary by gender. Thus, although women scored higher than males on the MD-ICM and DIT-2 the nature of the relationship between the variables did not differ by gender. This finding suggests that it is less likely that we will have to condition our discussion about relationships between the MD-ICM and other variables by gender.

### ***Implications for Medical Ethics Education and Assessment***

As described in Chapter 2, medical ethics education currently plays an important role in medical education. The importance of medical ethics is due in part, because ethical and moral issues are a significant challenge to medical practitioners and students need exposure to ethical issues before they enter the profession (Wong & Cheung, 2003). Even though medical ethics curriculums have been often used in medical school, assessment methods are less well developed. One possible explanation is that medical ethics assessment is a difficult process in

itself. For some, ethics is “considered intangible, as being based on opinion or sentiment” (Wong & Cheung, 2003, p. 5). To these practitioners any ethics assessment is a subjective measurement and, not obviously objective like the other field. As a counter to this view, the current measure is designed to be nested within the profession and directly tied to practice. It is hoped that by focusing on the intersection between ethics and practice, concerns about the relativity of the assessment and questions of its appropriateness will be lessened. How then can researchers apply the measure to assess medical ethics education? This question is addressed below.

1. First and foremost, this study supports the use of the ICM –based measurements as a tool for the assessment of ethics intervention. The steps taken to develop the MD-ICM apply more generally and can be used by medical educators to develop their own measurements in other setting, if not modify the MD-ICM used in the current study. Following the steps outlined in this study, medical educators can further test the validity of the ICM approach (i.e., content, criterion-related, and construct validity) and assess its properties.

2. The second implication is that the MD-ICM can be a valuable instrument for assessment ethical reasoning in medicine. The method of teaching and evaluation of ethical reasoning in schools of medicine are somewhat unique because medical students are required to deal with moral and ethical issues in their life practice along with clinical practice. A literature review shows, in general, teaching and evaluation methods of ethical reasoning are Problem-based learning (PBL), Evidenced-based medicine (EBM), Performance-based assessment, and Knowledge-based assessment (Myser, Kerridge, & Mitchell, 1995; Smith, Balint, Krause, Moore-West, & Viles, 1994; Orsat et al., 2009). A study by Chaves (2000) evidenced small group discussion, case-based problem solving involving ethical issues improved scores on the DIT measure and DEST measure (the Dental Ethical Sensitivity Test).

3. The MD-ICM is a new measure for medical education. The need for ethical reasoning assessment tools in professional health education is not restricted to the U.S. ICM can be tailored not only to fit a profession (e.g., medicine, dentistry) but also to fit a particular country or culture. The government of Thailand, through the Ministry of Public Health, has expressed an interest in assessing the moral reasoning of Thai medical students and doctors. The Thai government would like to use the ICM-MD to measure the moral reasoning of medical school applicants and students currently enrolled in medical schools in Thailand.

### ***Implications for ICM Development***

The ICM approach is primarily used to assess ethical understanding of professional school students (Bebeau and Thoma, (1999). Nonetheless, this study agrees with Bebeau and Thoma (1999) who asserted that “the current evidence supports attention to both intermediated concepts and moral schemas in the development and assessment of moral thinking” (Bebeau & Thoma, 1999, p. 357). Yet, there is a need to simplify the method for teaching and assessing moral knowledge for a profession’s code of ethics (Bebeau and Thoma, 1999). This study further suggests that the MD-ICM should be used with research in the field of instruction and assessment methods of professions ethics. These suggestions are as follows:

1. To provide support for the ICM approach, there has been a need for more ICM studies conducted in other professions rather in medicine such as for instance, lawyers, CEO’s, government workers, funeral directors, etc. Currently there exist ICM measures in the military setting, in high school as well as in dentistry. The development of ICM-based assessments in other setting seems particularly worthwhile.

2. A related concern for moral psychologists working in the health professions is the need for research on the methods for promoting the development of moral thinking. Most medical educators and health profession educators acknowledge the importance of promoting ethical

reasoning (Bebeau, 2002). A literature reviewed by Bebeau (2002) suggested that “does professional education promote ethical reasoning development”? Medical educations should be focused on content for medical ethics curriculum, which is still substantially debated on learning and teaching methods (Goldie, 2000).

### ***Implication for Future Research Using the MD-ICM***

As this study is the first to employ the MD-ICM there are many more studies are needed to fully explore the measure and assess its utility in medical ethics research and education. With this in mind, the author would offer the following suggested for future research that specifically relate to field of medical education regarding using the ICM (Intermediated Concepts Measure) in Thailand.

1. Future research needs to explore the pretest-posttest MD-ICM score in medical students at 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> years, in medical school in Thailand, to determine whether summary scores increase following the medical intervention developed by the CPIRD, Ministry of Public Health, Thailand. More importantly, the researcher needs to explore “construct validity” of the MD-ICM as an intervention study. Next, research question could be “is the MD-ICM sensitive to intervention?” If differences are found between the pretest and posttest, on MD-ICM scores for medical students at year 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> year then the support for the construct validity of the test will be improved.

2. Future research needs to explore whether the MD-ICM score relates to clinical performance. Current evidence indicates that *P*-scores on the DIT are linked to good clinical evaluations for medical residents and nurses (Sheehan et al., 1980; Duckett & Ryan, 1994). This finding is important because medical educators and other health educators can utilize the DIT score or MD-ICM as the test as a predictor as a good clinical practice. Similarly, it would be important to show that the MD-ICM is useful in relating to the actions of medical professionals.

In the Thai context, we need to measure clinical performance and MD-ICM in the 4<sup>th</sup>, 5<sup>th</sup> and 6<sup>th</sup> years medical students. These students are at the point in their training where they are graded for their performance in the various specialty areas (Pediatric, OBGYN, Internal Medicine, Surgery, Community medicine, etc.). Specifically, in Thailand, 6<sup>th</sup> years medical students spend the bulk of their time working with patients and are graded on their performance. Comparing MD-ICM scores with their clinical performance we should be able to estimate the predictive validity of the MD-ICM measure.

3. Based on the findings of the current study, the MD-ICM mean score in the non-integrated program was higher than integrated program. This surprising finding suggests that it would be interesting to investigate in detail in the content of curriculum that might account for these differences (e.g. method of teaching ethical reasoning, evaluation method, etc.). One approach suggested by this finding is to compare the MD-ICM scores between integrated curriculum and non-integrated curriculum for medical students. Medical schools in Thailand have similar curriculum content except with regard to the method of delivering information on professional ethics. Thus if differences are found on measures such as the MD-ICM then it seems plausible that the change is due to something in the curriculum. The MD-ICM may be particularly suited to identifying potential differences given the close connection between the test's content and medical education. If the expected outcomes for those curricula are not significantly different, then medical curriculum can either use non-integrated or integrated programs. On the other hand, if the outcomes of those programs are significantly different, the medical educator would need to redress his or her method of teaching ethical reasoning in medicine.

### ***Limitations of the Study***

The current study had certain limitations including the data collection process, the use of tests in English, a number of participants, and general problem in doing research in medical schools. With these limitations, researcher would be discussed each issue as follows:

1. During the data collection phase of this study, Thailand underwent a period of political unrest. Due to this set of circumstance, the researcher could not access some medical schools. Some buildings were closed during protests and different sections of the country were inaccessible-particularly the Muslim section located in the southern part of Thailand. Thus, the uneven distribution of participants across class, institution and region was considered a limitation, particularly the number of the 6<sup>th</sup> year's medical students. As a result, there were not enough medical students available to use for data analysis for certain research questions. Further, the political events may have served as a distraction for medical students who might otherwise may have been more interested in their studies, their profession, and school-based events such as taking the MD-ICM.

2. Because this study utilized the English language for the MD-ICM and the DIT2, this could have introduced unforeseen limitations due to English not being the students' primary language. One possible explanation was that it was not possible to administer these tests in Thai because the true meaning of the test might get lost in translation. An English proficiency difficulty is possible even though Thai medical students all are required to be fluent in English and is top students in their respective classes during the equivalent of our high schools. It was thus assumed that their English skills were at least at the 9<sup>th</sup> grade level required by the DIT and MD-ICM. Generally speaking, however, Thai medical students might have had a problem while taking the tests in English. It is interesting to note that some Thai medical students complained because the test was not given in the Thai language.

3. Medical students also complained about time constraints for their profession. It is interesting to note that conducting research in medical school was considered a limitation due to limited time, heavy class load throughout semesters, a great deal of which added greater responsibility in their professions. Future studies should anticipate the reluctance of students to contribute time to a research project and develop connections with faculty and administrators to encourage participation of students.

## REFERENCES

- AMA (1985 , Dec.). Diagnostic and therapeutic technology assessment. Enhanced computed tomography in head trauma, JAMA. *Journal of the American Medical Association*, 20(23), 3370-1.
- American Nurses' Association (1976, December 15). ANA model practice act revised. *The American Nurse*, 8(16), 5, 32.
- Baldwin, D. C. Jr., Adamson, T. E., Self, D. J., Sheehan, T. J., & Oppenberg, A. A. (1996). Moral reasoning and malpractice: A pilot study of orthopedic surgeons. *The American Journal of Orthopedics*, 25(7), 481-484.
- Baldwin, D. C., Jr., Daugherty, S. R., & Rowley, B. D. (1998). Unethical and unprofessional conduct observed by residents during their first year of training. *Academic Medicine*, 73(11), 1195-1200.
- Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, 31(2), 101-119.
- Beauchamp, T. L., & Childress, J. F. (2001). *Principles of biomedical ethics* (5<sup>th</sup> ed.). Oxford, DC: Oxford University Press.
- Bebeau, M. J. (2001). Influencing the moral dimensions of professional practice: Implications for teaching and assessing for research integrity, proceedings of the ORI Conference on Research on Research Integrity. Retrieved August 25, 2008 from <http://www-pernal.umich.edu/~nsteneck/recr/index.html>
- Bebeau, M. J. (2002). The DIT and the FCM: contributions to professional education. *Journal of Moral Education*, 34(3), 271-295.
- Bebeau, M. J., & Thoma, S. J. (1994). The impact of a dental ethics curriculum on moral reasoning. *Journal of Dentist Education*, 58(9), 684-692.
- Bebeau, M. J., Rest, J., & Narvaez, D. (1999). Moving beyond the promise. A perspective for research in moral education. *Educational Research*, 28(4), 18-26.
- Bebeau, M. J., Rest, J., & Yamoore, C. M. (1985). Measuring dental students' ethical sensitivity. *Journal of Dental Education*, 49, 225-235.

- Bebeau, M. J., & Thoma, S. J. (1999). "Intermediate" concepts and connection to moral education. *Educational Psychology Review*, 11(4), 343-360.
- Bissonette, R., O'Shea, R. M., Horwitz, M., & Route, C. F. (1995). A data-generated basis for medical ethics education: Categorizing issues experienced by students during clinical training. *Academic Medicine*, 70(11), 1035-1037.
- Blasi, A. (2004). Moral functioning: Moral understanding and personality. In D.K. Lapsley & D. Narvaez (Eds.), *Moral development, self, and identity*, (pp. 335-347). New Jersey: Lawrence Erlbaum Associates.
- Bunch, W. H., Storr, C. L., Hughes, P. H., & Baldwin, D. C. Jr. (1996, February 15). Substance use by surgical residents and students entering surgery. *The Journal of Surgical Research*, 61(1), 108-12.
- Bradley, P. A. (1983). Every consultation has an ethical component. *British Medical Journal (Clinical Research ed.)*, 287(6389), 399-400.
- Branch, T. W. (2000). Supporting the moral judgment of medical students. *Journal General Internal Medicine*. 15, 503-508.
- Bretherton, I. (1992). The origin of attachment theory: John Bowlby and Mary Ainsworth. *Developmental Psychology*, 28(5), 759-775.
- Campbell, A. V., (2003). The virtues (and vices) of the four principles. *Journal of Medical Ethics*, 29, 292-296.
- Candee, D., Sheehan, T. J., Cook, C. D., Husted, S. D., & Borgen, M. (1982). Moral reasoning and decisions in dilemmas of neonatal. *Pediatric Research*, 16, 846-850.
- Charon R., & Fox R. (1995). Critiques and Remedies: medical students call for change in ethics teaching. *Journal of the American Medical Association*, 274.
- Chaves, J. F. (2000). Assessing ethics and professionalism in dental education. *Journal of Dental Association*, 79(1), 16-21.
- Childress, J. F. (1997). The normative principles of medical ethics. In R.M. Veatch, *Medical ethics* (pp. 29-56). Massachusetts, Boston: Jones and Bartlett Publishers.
- Christakis, D.A., & Feudtner, C. (1993). Ethics in a short white coat: the ethical dilemmas that medical students confront. *Academic Medicine*, 68(4), 249-254.
- Christie, R. C., Bowen, D. M., & Paarmann, C. S. (2003). Curriculum evaluation of ethical reasoning and professional responsibility. *Journal of Dental Education*, 67(1), 55-63.
- Colby, A., & Kohlberg, L. (1987). The measurement of moral judgment: vol.1. *Theoretical foundations and research validation*. Cambridge, England: Cambridge University Press.

- Doyal, L. (2001). Closing the gap between professional teaching and practice a policy can help protect students from being asked to behave unethically. *BMJ*, 322(7288), 685-686.
- Duckett, L. J., & Ryan, M. B. (1994). Education for ethical nursing practice. In J. Rest & D. Narvaez (Eds.), *Moral development in the profession: Psychology and applied ethics* (pp. 199-211). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ernest, M. (1990). *Developing and testing cases and scoring criteria for assessing geriatric dental ethical sensitivity*. Unpublished M.S. Thesis. University of Minnesota, MN.
- Feudtner, C., Christakis, D. A., & Christakis, N. A. (1994). Do clinical clerks suffer ethical erosion? Students' perceptions of their ethical environment and personal development. *Academic Medicine*, 69(8), 670-679.
- Gert, B. (2005). *Morality: Its nature and justification* (Rev. ed.). New York: Oxford University Press.
- Gilligan, C. (1982). *In a difference voice: Psychological theory and woman development*. Cambridge: Harvard University Press.
- Gillon, R. (2003). Ethics needs principles – four can encompass the rest – and respect for autonomy should be “first among equals.” *Journal of Medical Ethics*, 29, 307-312.
- Ginsburg, S., Regehr, G., & Lingard, L. (2003). Disavowed curriculum: understanding students' reasoning in professionally challenging situations. *JGIM*, 18, 1015-1022.
- Goldie, J., Schwartz, L., McConnachie, A., & Morrison, J. (2002). The impact of three year' ethics teaching, in an integrated medical curriculum, on students' proposed behavior on meeting ethical dilemmas. *Medical Education*, 36, 489-497.
- Goldie, J., Schwartz, L., McConnachie, A., & Morrison, J. (2003). Students' attitudes and potential behavior with regard to whistle blowing as they pass through a modern medical curriculum. *Medical Education*, 37, 368-375.
- Goold, S. D., & Stern, D. T. (2006). Ethics and professionalism: What does a resident need to learn? *The American Journal of Bioethics*, 6(4), 9-17.
- Hanson, S. (2005). Teaching health care ethics: why we should teach nursing and medical students together. *Nursing ethics*, 12(2), 167-176.
- Hafferty, F. W., & Franks, R. (1994). The hidden curriculum, ethics teaching, and the structure of medical education. *Academic Medicine*, 69(11), 861-871.
- Hariharan, S., Jonnalagadda, R., Walrond, E., & Moseley, H. (2006, June 9). Knowledge, attitudes and practice of healthcare ethics and law among doctors and nurses in Barbados. *BMC: Medical Ethics*, 7, E7

- Hebert, P. C., Meslin, E. M., & Dunn, E. V. (1992). Measuring the ethical sensitivity of medical students: a study at the University of Toronto. *Journal Medical Ethics*, 18(3), 142-147.
- Henning, W. (2006). Bioethics in China. *EMBO reports*, 7(9), 850-854.
- Hicks, L. K., Lin, Y., Robertson, D. W., Robinson, D. L., & Woodrow, S. I. (2001). Understanding the clinical dilemmas that shape medical students' ethical development: questionnaires survey and focus group study. *BMJ*, 322(1218), 709-710.
- Higgs, R. (1985). Medical ethics. *The Practitioner* 229(1410), 1063-4.
- Hughes, J. H. & Keown, D. (1995). Buddhism and Medical Ethics: A Bibliographic Introduction. *Journal of Buddhist Ethics*, (2).
- Hundert, E. M., Douglas-Steele, D., & Bickel, J. (1996). Context in medical education: The informal ethics curriculum. *Medical Education*, 30, 353-364.
- Hundert, E. M., Hafferty, F., & Christakis, D. (1996, June). Characteristics of the informal curriculum and trainees' ethical choices. *Academic medicine*, 71(6), 624-42.
- Husted, S. D. (1978). Assessment of moral reasoning in pediatric faculty, house officers and medical students. Proceedings of the 17<sup>th</sup> annual conference. *Research in Medical Education*, 17, 439-444.
- Ketefian, S. (1981, May-June). Moral reasoning and moral behavior among selected groups of practicing nurses. *Nursing Research*, 30(3), 171-6.
- Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Cognitive development and epistemology*, (pp.151-236). New York: Academic.
- Killen, J. & Smetana, J. (Eds.) (2006). *Handbook of moral development*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kohlberg, L. (1984). Essay on moral development: The nature and validity of moral state: The *Philosophy of Moral Development*, 2, San Francisco: Harper & Row.
- Kohlberg, L., & Kramer, R. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D.A. Goslin (Ed.), *Handbook of socialization theory* (pp.347-480). Chicago: Rand McNally.
- Lapsley, D. K. (1996). *Moral psychology*. Boulder, CO: Westview Press.
- Lapsley, D. K. (2006). Moral stage theory. In M. Killen & J.G. Smetana (Eds.), *Handbook of Moral Development*, (pp. 7-35). New Jersey: Lawrence Erlbaum Associates.

- Lehmann, L. S., Kasoff, W. S., Koch, P., & Fedeman, D. D. (2004). A survey of medical ethics education at U.S. and Canadian medical schools. *Academic Medicine*, 79(7), 682-689.
- Lewin, L. O., & Lancken, P. N. (2004). Longitudinal small-group learning during the first clinical year. *Family Medicine*, 36(January supplement), 83-88.
- Lockwood, M. (1985). *Moral dilemmas in modern medicine*. Oxford, NY: Oxford University Press.
- Loewy, E. H. (2003). Education, practice and bioethics: Growing barriers to ethical practice. *Health Care Analysis*, 11(2), 171-179.
- McAlpine, H., Kristjanson, L., & Poroch, D. (1997). Development and testing of ethical reasoning tool (ERT): An instrument to measure the ethical reasoning of nurse. *Journal of Advance Nursing*, 25(6), 1151-1161.
- Myser, C., Kerridge, I. H., Mitchell, K. R. (1995). Ethical reasoning and decision-making in the clinical setting: assessing the process. *Medical Education*, 29(1), 29-33.
- Nilsson, I., Crafoord, J., Hedengren, M. & Ekehammar, B. (1991). The sociomoral reflection measure: Applicability to Swedish children and adolescents. *Scandinavian Journal of Psychology*, 32(1), 48-56.
- Oddi, L. F., & Cassidy, V. R. (1994). The JAND as a measure of nurses' perception of moral behavior. *International Journal of Nursing Studies*, 31(1), 37-47.
- Orsat, M., Bigot, P., Roupert, M., Campillo, B., Beley, S., Chautard, D., Beaufreton, C., Richard, I., Saint-Andre, J. P. & Azzouri, A. R. (2009, March). Evidence-based medicine and French medical students: An appraisal. *Progrès en urologie*, 19(3), 215-220.
- Patenaude, J., Niyonsenga, T., & Fafard, D. (2003). Changes in the components of moral reasoning during students' medical education: a pilot study. *Medical Education*, 37(9), 822-829.
- Patenaude, J., Niyonsenga, T., & Fafard, D. (2003). Changes in students' moral development during medical school: A cohort study. *Canadian Medical Association Journal*, 168(7), 840-843.
- Piaget (1931/1965). *The Moral judgement of the child*. New York: The Free Press.
- Pulman, J. (2003). Enlightening medical students. *Journal of American Medical Colleges*, 168,11.
- Redman, B. K. (2005, July). The ethics of self-management preparation for chronic illness. *Nursing Ethics*, 12(4), 360-9.

- Rest, J., (1979). Development in Judging Moral Issues. University of Minnesota. Rest (Ed.), *Moral development advances in research and theory* (pp. 1-27). New York, Westport, Connecticut, London: Praeger Publishers.
- Rest, J. (1986). *Moral development: Advance in research and theory* (pp. 28-58). New York: Praeger.
- Rest, J., Bebeau, M., & Volker, J. (1986). An overview of the psychology of morality. In J. Rest (Ed.), *Moral development advances in research and theory* (pp. 1-27). New York, Westport, Connecticut, London: Praeger Publishers.
- Rest, J., & Narvaez, D. (1994). *Moral development in the professions: Psychology and applied ethics*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Rest, J. R., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (1998). *Postconventional moral thinking: a neo-Kohlbergian approach*. Unpublished manuscript, University of Minnesota.
- Rest, J., Narvaez, D., Bebeau, M. J., & Thoma, S. J. &. (1999). *Postconventional moral thinking: A neo-Kohlbergian approach*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Rest, J. R., Narvaez, D., Thoma, S. J. & Bebeau, M. J. (2000). A neo-Kohlbergian approach in morality research. *Journal of Moral Education*, 29(4), 381-395.
- Rogers, W. A., Mansfield, P. R., Braunack-Mayer, A. J., & Jureidini, J. N. (2004, April 19). The ethics of pharmaceutical industry relationships with medical students. *The Medical Journal of Australia*, 180(8), 411-4.
- Rothman, G. R. (1980). The relationship between moral judgment and moral behavior. In Eindmiller, M., Lambert, N., & Turiel, E., *Moral development and socialization*. (pp. 107-127). Boston, London, Sydney, Toronto: Allyn and Bacon.
- Russell, C., & O'Neill, D. (2006). Ethicist and clinicians: The case for collaboration in the teaching of medical ethics. *Irish Medical Journal*, 99(1), 25-27.
- Self, D. J., & Baldwin, D. C. (Ed.). (1994). *Moral reasoning in medicine*. Hillsdale, New Jersey, Hove, UK: Lawrence Erlbaum Associates.
- Self, D. J., Baldwin, D. C. Jr., & Olivarez, M. (1989). Teaching medical ethics to first-year students by using film discussion to develop their moral reasoning. *Academic Medicine*, 64(12), 755-759.
- Self, D. J., Baldwin, D. C., Jr., & Olivarez, M. (1993). Teaching medical ethics to first-year students by using film discussion to develop their moral reasoning. *Academic Medicine*, 68(5), 383-385.

- Self, D. J., Olivarez, M., & Baldwin, D. C., Jr. (1998). The amount of small-group case-study discussion needed to improve moral reasoning skills of medical students. *Academic Medicine*, 73, 521-523.
- Self, D. J., Schrader, D. E., Baldwin, D. C. Jr., Root., S. K., Wolinsky, F. D., Shadduck, J. A. (1991). Study of the influence of veterinary medical education on the moral development of veterinary students. *Journal of the American Veterinary Medical Association*, 198(5), 782-787.
- Self, D.J., Schrader, D.E., Baldwin, D.C., Wolinsky, F.D. (1993). The moral development of medical students: a plot study of the possible influence of medical education. *Medical Education*, 27(1), 26-34.
- Self, D. J., Skeel, J. D., & Jecker, N. S. (1993). The influence of philosophical versus theoretical education on the moral development of clinical medical ethicists. *Academic Medicine*, 68(11), 848-52.
- Self, D. J., Wolinsky, F. D., & Baldwin, D. C. Jr. (1989). The effect of teaching medical ethics on medical students' moral reasoning. *Academic Medicine*, 64(12), 755-759.
- Shaffer, D. R. (1999). *Developmental psychology: Childhood & adolescence* (5<sup>th</sup> ed.). Washington, Pacific Grove: Brooks/Cole Publishing Company.
- Sheehan, T. ., Husted, S. D. R., Candee, D., Cook, C. D., & Bargen, M. (1980). Moral judgment as a predictor of clinical performance. *Evaluation & The Health Profession*, 3(4), 393-404.
- Smith, S. R., Balint, J. A., Krause, K. C., Moore-West, M., & Viles, P. H. (1994). Performance-based assessment of moral reasoning and ethical judgment among medical students. *Academic Medicine*, 69(5), 381-386.
- Snarey, J. R. (1985). Cross-cultural universality of social-moral development: A critical review of Kohlbergian research. *Psychological Bulletin*, 97, 202-232.
- St. Onge, J. (1997). Medical education must make room for student-specific ethical dilemmas. *CMAJ: Canadian Medical Association Journal*, 15; 156(8), 1175-7.
- Strike, K. A. (1982). *Educational policy and the justice society*. Chicago: University of Illinois Press.
- Thoma, S. J.(1986). Estimating gender differences in the comprehension and references of moral issues. *Developmental Review*, 6, 165-180.
- Thoma, S. (2006). Research of the defining issues test. In M. Killen & J. Smetana (Eds.). *Handbook of moral development*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Thoma S. J., & Rest, J. (1999). The relationship between moral decision-making and patterns of consolidation and transition in moral judgment development. *Developmental Psychology*, 35, 323-334.
- Thomasma, D. C. (1998, August). Assisted death and martyrdom. *Christian Bioethics*, 4(2), 122-42.
- Turiel, E. (1966). An experimental test of the sequentiality of developmental stages in the child's moral judgment. *Journal of Personality and Social Psychology*, 3, 611-618.
- Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. Killen & J.G. Smetana (Eds.), *Handbook of moral development*, (pp. 7-35). New Jersey: Lawrence Erlbaum Associates.
- Turner, (2008). The development and testing of an army leader intermediate ethical concepts measure. Dissertation, The University of Alabama, Tuscaloosa, AL.
- Veatch, R.M. (1997). *Medical ethics* (2<sup>nd</sup> ed.). Boston: Jones and Bartlett Publishers.
- Walker, L. J. (2002). The model and the measure: An appraisal of the Minnesota approach to moral development. *Journal of Moral Education*, 31, 353-367.
- Walker, L. J. (2004). Gus in the gap: bridging the judgment-action gap in moral functioning. In D.K. Lapsley & D. Narvaez (Eds.), *Moral development, self, and identity*, (pp. 1-20). New Jersey: Lawrence Erlbaum Associates.
- Wong, J., & Cheung, E. (2003). Ethics assessment in medical students. *Medical Teacher*, 25(1), 5-8.
- Ypinazar, V. A., & Margolis, S. A. (2004). Western medical ethics taught to junior medical students can cross cultural and linguistic boundaries. *BMC Medical Ethics*, 5(4), 1-7.
- Young, T. A (1997, January 15). Teaching medical students to lie. The disturbing contradiction: Medical ideals and the resident-selection process. *CMAJ: Canadian Medical Association Journal*, 156(2), 219-22.
- Zimmerman, B. J., & Schunk, D. (2003). Albert Bandura: The scholar and his contributions to educational psychology. In B. J. Zimmerman & D.H. Schunk (Eds.), *Educational psychology: A century of contributions* (pp. 431-457). New Jersey: Lawrence Erlbaum Associates.

APPENDIX A  
INSTITUTIONAL REVIEW BOARD APPROVAL

Office for Research  
Office of the Chair,  
Institutional Review Board for the  
Protection of Human Subjects

THE UNIVERSITY OF  
**ALABAMA**  
R E S E A R C H

August 28, 2008

Ponsopa Pinijphon  
Department of Educational Psychology  
College of Education

Re: IRB # EX-08-CM-068 "An ICM Approach to the Assessment of an  
Ethics Intervention in Medical Education in Thailand"

Dear Ms. Pinijphon:

The University of Alabama Institutional Review Board has granted  
approval for your proposed research.

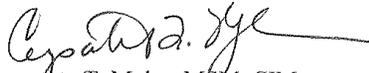
Your protocol has been given exempt approval according to 45 CFR part  
46.101(b)(2) as outlined below:

(2) Research involving the use of educational tests (cognitive, diagnostic,  
aptitude, achievement), survey procedures, interview procedures or  
observation of public behavior, unless:  
(i) information obtained is recorded in such a manner that human subjects  
can be identified, directly or through identifiers linked to the subjects; and  
(ii) any disclosure of the human subjects' responses outside the research  
could reasonably place the subjects at risk of criminal or civil liability or  
be damaging to the subjects' financial standing, employability, or  
reputation.

Should you need to submit any further correspondence regarding this  
proposal, please include the assigned IRB application number.

Good luck with your research.

Sincerely,



Carpanato T. Myles, MSM, CIM  
Director of Research Compliance & Research Compliance Officer  
Office of Research Compliance  
The University of Alabama



152 Rose Administration Building  
Box 870104  
Tuscaloosa, Alabama 35487-0104  
(205) 348-5152  
FAX (205) 348-8882

EX-08-CM-068

UNIVERSITY OF ALABAMA INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS  
REQUEST FOR APPROVAL OF RESEARCH INVOLVING HUMAN SUBJECTS

I. Identifying information

**Principal Investigator**                      Second Investigator                      Third Investigator  
Name: Ms. Ponsopa Pinijphon  
Department: Educational Psychology  
College: School of Education  
University: The University of Alabama  
Address: 4512 Culver Rd Tuscaloosa, AL  
Telephone: 205-246-6643  
FAX:  
E-mail: Pinij001@bama.ua.edu

Title of Research Project: *An FCE Approach to the Assessment of an Ethics Intervention in Medical Education in Thailand*  
Date Printed: \_\_\_\_\_  
Funding Source: \_\_\_\_\_

Type of Proposal:     New     Revision     Renewal     Completed     Exempt

Attach a renewal application

Attach a continuing review of studies form

Please enter the original IRB # at the top of the page

UA faculty or staff member signature: *[Signature]*

II. NOTIFICATION OF IRB ACTION (to be completed by IRB):

Type of Review:     Full board     Expedited

IRB Action:

Rejected    Date: \_\_\_\_\_  
 Tabled Pending Revisions                      Date: \_\_\_\_\_  
 Approved Pending Revisions                      Date: \_\_\_\_\_

Approved—this proposal complies with University and federal regulations for the protection of human subjects  
Approval is effective until the following date: *9-4-09 f.e*

Items approved:     Research protocol:    dated  
                                  Informed consent:    dated  
                                  Recruitment materials:    dated  
                                  Other:    dated

Approval signature *[Signature]*                      Date *9/4/08*

APPENDIX B  
MD-ICM INSTRUMENT

Name: \_\_\_\_\_

Student Number: \_\_\_\_\_

Year: \_\_\_\_\_

University: \_\_\_\_\_

### Intermediate Concept Measures -Medicine

#### (ICM-Med)

#### General Instructions

Thank you for agreeing to participate in this study. This questionnaire packet should take between 60 and 90 minutes to complete. Your careful consideration of each story and the questions/items that follow is greatly appreciated!

Please note that participation is strictly voluntary and you may stop at any time. Also note that your responses will be kept strictly confidential. Your name and/or student number will never be used in describing the results. Only group data will be presented. When you return the completed questionnaire this cover page will be torn off so that credit for participation can be recorded. Your name will not appear on any other page.

If you have any questions, please feel free to contact your instructor.

#### Directions:

This booklet contains five cases that present ethical challenges for practicing physicians. After each case, there is a list of action choices that could be made for the case. On the accompanying response sheet, rate each action choice. Then, pick the three best and two worst choices.

When you have finished rating and ranking your action choices, turn the page and rate and rank the reasons to justify your action choices.

Should you think of actions or justifications that seem better than the options presented, write down your idea, but please make a choice from the options presented. Keep in mind, that

there is no single right answer to each case. Rather, some choices are better than others, and some are probably not defensible.

After everyone has finished rating and ranking responses to all the cases, we will discuss each case. You will have a chance to compare your responses with responses of your peers and with the responses of a national sample of teachers of medical ethics. The emphasis will be on the reasons some responses to be more defensible than others.

### **Example**

Please read the story, then complete the questionnaire.

### **Case A**

Mr. Franklin Waldrop is a 35 year old white male who has been your patient for twelve years. He was divorced three years ago and has two young daughters. He works as an electrician with a local firm.

During his divorce, Mr. Waldrop suffered from asthma and hives. He responded well to Zantac. He in good health except for severe allergic rhinitis that flares up every spring and sometimes induces an asthma attack. Desensitization shots have partially relieved his symptoms. He would be a good candidate for a placebo effect study that you have been invited to join.

**What do you do about Mr. Waldrop and the placebo study?**

Rate the statement below using the following scale:

1. **believe that this is GOOD choice**
2. **I believe that this is a GOOD choice**
3. **I am not sure**
4. **I strongly I believe that this is a BAD choice**
5. **I strongly believe that this is a BAD choice**

**The following are examples of choices that you might think about:**

- \_\_\_1. Get Mr. Waldrop's consent and tell him that he will get a placebo.
- \_\_\_2. Refuse to join in the study.
- \_\_\_3. Get Mr. Waldrop's consent without explaining that he could get a placebo.
- \_\_\_4. Tell Mr. Waldrop that it is not appropriate to explain the research.

Your job is to rate each action choice using the scale above. For example, someone might feel that number three is a very bad choice for Case A, that number one is may not be so good, that number two is a great choice, and that number four is a pretty good choice. So, you might rate the examples this way.

- \_\_3\_\_1. Get Mr. Waldrop's consent and tell him that he will get a placebo.
- \_\_1\_\_2. Refuse to join in the study.
- \_\_4\_\_3. Get Mr. Waldrop's consent without explaining that he could get a placebo.
- \_\_2\_\_4. Tell Mr. Waldrop that it is not appropriate to explain the research.

Following this, you will decide which is the best and the worst choices. For instance, one might say that number three is the best reason and number one is the worst choice.

Next, you will be asked to think about the reasons someone might consider in making a choice.

You will be asked to judge each reason in the following scale:

- 1. I strongly believe that this is GOOD choice**
- 2. I believe that this is a GOOD choice**
- 3. I am not sure**
- 4. I believe that this is a BAD choice**
- 5. I strongly believe that this is a BAD choice**

The following are examples of reasons that you might see:

1. It would be provide data that could improve the health of many.
2. He might benefit from participating.
3. I will get paid and, perhaps, be invited to participate in other projects.
4. I can't guarantee fully and informed consent.

Remembering Case A's story and that number three was selected as the best action choice; then someone might feel that number one is exactly what he or she was thinking for that choice. Someone might think that number two has a little bit to do with what he or she was thinking in making that choice. Someone might think that number three has nothing to do with the choice. And, someone might think that number four is not much better. So, you would rate the reasons in following way:

After rating reach item you will be asked to rank the best and worse reason. Based on the above it would make sense to rank number one as the best reason and number three or number four as the worst reason.

When you finished with this example, you may turn to the first story and begin

**Thank you!**

## Case 1

Kent Jones is a 58 year old white male. He is married and he has three children ranging from 30-38 years. He had a stroke two years ago but had recovered well enough to return to his job. He has a history of high blood pressure. He has non-insulin dependent diabetes which has not been well controlled, partly because he continued to eat sugary snacks between meals. He takes his blood pressure medicine regularly. His older sister died last year from complications following a stroke.

Five days ago, Mr. Jones suffered a major stroke. Three times in the hours after the stroke, his heart stopped but was re-started with electrical stimulation. His EEG shows only brain stem activity. He has severe, irreversible brain damage. In the last two days, he has developed adult respiratory distress syndrome. He is currently on a ventilator and feeding tubes. Mr. Jones had signed a "living will" which states his objection to being maintained on life support if there was no medical hope for recovery. However, Mrs. Jones wishes to keep her husband alive.

**What do you wish to do about Mr. Jones?**

**Case 1: Action Choices**

**For each of the action choices below, circle the number that corresponds to your view about each choice. Rate each choice using the following scale:**

- 1- I strongly believe that this is a GOOD choice**
- 2- I believe that this is a GOOD choice**
- 3- I am not sure**
- 4- I believe that this is a BAD choice**
- 5- I strongly believe that this is a BAD choice**

- \_\_\_ 1. Ask for a panel of doctors to make the decision.
- \_\_\_ 2. [Honor Mr. Jones's wishes in his" living will" and discontinue life support.](#)
- \_\_\_ 3. Remove the life support so that it is available for other patients who might need it to survive.
  
- \_\_\_ 4. [Consult experts about the problem.](#)
  
- \_\_\_ 5. Let another Doctor assume responsibility for Mr. Jones's care.
  
- \_\_\_ 6. [Honor Mrs. Jones's wishes by leaving Mr. Jones on Life support.](#)
  
- \_\_\_ 7. Tell the family that you will wait one more day then discontinue life support if Mr. Jones's condition does not improve.
  
- \_\_\_ 8 Agree to wait a few days while trying to convince Mrs. Jones and family to discontinue life support.

**Please rank the best three choices by putting the item number (1-8) in the blank**

\_\_\_ Best choice

\_\_\_ 2<sup>nd</sup> best

\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-8) in the blank**

\_\_\_ Worst choice

\_\_\_ 2<sup>nd</sup> worst

### Case 1: Justification for Action Choices

For each of the reasons below, circle the number that corresponds to your view about each choice. Rate each reason using the following scale:

1. I strongly believe that this is a GOOD reason
2. I believe that this is a GOOD reason
3. I am not sure
4. I believe that this is a BAD reason
5. I strongly believe that this is a BAD reason

- \_\_\_\_ 1. A delay in making your decision is acceptable because of the importance to the family.
- \_\_\_\_ 2. The financial costs to both his family and the hospital are very large.
- \_\_\_\_ 3. It is necessary to follow the law concerning Mr. Jones's living will.
- \_\_\_\_ 4. Sufficient information is not available to make an intelligent decision.
- \_\_\_\_ 5. The emotional impact of this choice on Mrs. Jones and her family is too great.
- \_\_\_\_ 6. Your legal and professional responsibilities to the family are as important as your responsibility to Mr. Jones.
- \_\_\_\_ 7. You should try to reconcile the family's wishes with your choice.
- \_\_\_\_ 8. You have not used all the available resources to assist them.
- \_\_\_\_ 9. You have a responsibility to other patients who might need the ventilator to survive during the delay.
- \_\_\_\_ 10. You have not used all of the relevant resources available to you.
- \_\_\_\_ 11. You have identified specific issues with which consultants might help.
- \_\_\_\_ 12. It is important to follow Mr. Jones's wishes as expressed in his "living will".  
(Respect his autonomy)
- \_\_\_\_ 13. Further consultation could make a difference in your decision.

**Please rank the best three reasons by putting the item number (1-13) in the blank**

\_\_\_\_ Best choice

\_\_\_\_ 2<sup>nd</sup> best

\_\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-13) in the blank**

\_\_\_\_ Worst choice

\_\_\_\_ 2<sup>nd</sup> worst

## Case 2

Jeffrey Darren is a 42 year old white male. He has a history of behavior that puts him at high risk for AIDS. He has bragged to his friends that he is very sexually active. He seldom uses condoms because he does not like the way that they feel. He plans to marry in two months. His fiancé insisted that he get tested for sexually transmitted diseases before the marriage so he came in for a physical exam and a battery of STD tests. His fiancé was also screened for STD's.

The couple received the appropriate counseling and a packet of information on HIV, Hepatitis, etc. Blood was drawn and sent to the lab. Both Mr. Darren and his fiancé appear to be in good general health. Results from the physical exam were within normal limits.

The lab report is negative for STD's for the fiancé.

The lab report for Mr. Darren is positive for HIV. He does not want to tell his fiancé about the test results.

**What do you wish to do about Mr. Darren's request not to tell his fiancé?**

**Case 2: Action Choices**

**For each of the action choices below, circle the number that corresponds to your view about each choice. Rate each choice using the following scale:**

- 1. I strongly believe that this is a GOOD choice**
- 2. I believe that this is a GOOD choice**
- 3. I am not sure**
- 4. I believe that this is a BAD choice**
- 5. I strongly believe that this is a BAD choice**

- \_\_\_\_ 1. Respect Mr. Darren's confidentiality by not telling.
- \_\_\_\_ 2. Try to talk him into telling his wife.
- \_\_\_\_ 3. Inform her regardless of his wishes.
- \_\_\_\_ 4. Report his HIV status to the public health authorities and let them handle it.
- \_\_\_\_ 5. Have the patient sign a form that acknowledges this receipt of information about the consequences of HIV and sexual activity.
- \_\_\_\_ 6. Try to further educate the patient.
- \_\_\_\_ 7. Consult with experts about your options.
- 

**Please rank the best three choices by putting the item number (1-7) in the blank**

\_\_\_\_ Best choice

\_\_\_\_ 2<sup>nd</sup> best

\_\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-7) in the blank**

\_\_\_\_ Worst choice

\_\_\_\_ 2<sup>nd</sup> worst

## Case 2: Justification for Action Choices

For each of the reasons below, circle the number that corresponds to your view about each choice. Rate each reason using the following scale:

1. I strongly believe that this is a GOOD reason
2. I believe that this is a GOOD reason
3. I am not sure
4. I believe that this is a BAD reason
5. I strongly believe that this is a BAD reason

- \_\_\_ 1. You have a professional obligation to the fiancé.
- \_\_\_ 2. You have to consider the clinical welfare of the fiancé.
- \_\_\_ 3. Your clinical objectives for the care of this patient are worth the continued risk of infection to others.
- \_\_\_ 4. Other people in your community have a stake in this matter. They have rights, too.
- \_\_\_ 5. You have to fulfill your fiduciary duty to your patient AND provide for his best interests.
- \_\_\_ 6. You have to show adequate respect for his right of autonomy.
- \_\_\_ 7. His right to confidentiality has to be given weight.
- \_\_\_ 8. You have to maintain his trust in you in order to provide appropriate care for him.

**Please rank the best three choices by putting the item number (1-8) in the blank**

\_\_\_ Best choice

\_\_\_ 2<sup>nd</sup> best

\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-8) in the blank**

\_\_\_ Worst choice

\_\_\_ 2<sup>nd</sup> worst

### Case 3

Sheila Curtis is a 14 year old white female who was diagnosed with non-Hodgkins lymphoma two years ago. She received both radiation and chemotherapy and seemed to have been in remission for the past year.

Sheila looks healthy now but she complains of persistent tiredness. A routine follow-up physical exam 4 months ago showed nothing unusual but a chest x-ray taken in your office last week revealed a suspicious mass. An extensive work-up diagnosed a recurrence of the cancer, with liver, spleen and bone marrow involvement. You know that her prognosis is not good. Sheila's oncologist has suggested more chemotherapy.

Sheila and her parents have come to talk with you about her care. Sheila's parents feel that the accepted standard treatment suggested by the oncologist will just let her die and will cause unnecessary suffering. The parents want you to arrange for a treatment offered in a Mexican clinic which uses shark cartilage for treating cancer and claims amazing results with no side effects. They learned about this treatment on a TV show.

**How will you handle the Curtis' request to find alternative treatment?**

Case 3: Action Choices

**For each of the action choices below, circle the number that corresponds to your view about each choice. Rate each choice using the following scale:**

- 1. I strongly believe that this is a GOOD choice**
- 2. I believe that this is a GOOD choice**
- 3. I am not sure if this choice is Good or Bad**
- 4. I believe that this is a BAD choice**
- 5. I strongly believe that this is a BAD choice**

- \_\_\_ 1. Agree to cooperate as fully as you can.
- \_\_\_ 2. Refer to a clinic that uses alternative treatments.
- \_\_\_ 3. Try to arrange cooperative treatment including both standard and alternative treatments.
- \_\_\_ 4. Refer to a psychologist.
- \_\_\_ 5. Refuse and insist on standard therapy.
- \_\_\_ 6. Try to talk with the parents and Sheila separately about her care.
- \_\_\_ 7. Thoroughly evaluate the proposed alternative therapy before you answer.
- \_\_\_ 8. Document your care and advice then “fire” this patient.
- \_\_\_ 9. Protect Sheila by challenging the alternative therapy.
- \_\_\_ 10. Agree to participate in the therapy to assure the best care for Sheila.
- \_\_\_ 11. Ask the court to appoint a guardian ad litem for Sheila.

**Please rank the best three choices by putting the item number (1-11) in the blank**

\_\_\_ Best choice

\_\_\_ 2<sup>nd</sup> best

\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-11) in the blank**

\_\_\_ Worst choice

\_\_\_ 2<sup>nd</sup> worst

### Case 3: Justification for Action Choices

For each of the reasons below, circle the number that corresponds to your view about each choice. Rate each reason using the following scale:

1. I strongly believe that this is a GOOD reason
2. I believe that this is a GOOD reason
3. I am not sure
4. I believe that this is a BAD reason
5. I strongly believe that this is a BAD reason

- \_\_\_ 1. You have to fulfill your professional obligations to your patient and his family..
- \_\_\_ 2. You have to adequately consider Sheila's clinical welfare.
- \_\_\_ 3. You can justify not following the standard care.
- \_\_\_ 4. The unapproved therapy can be justified in this case.
- \_\_\_ 5. The patient's autonomy has to be respected.
- \_\_\_ 6. Parental rights have to be respected.
- \_\_\_ 7. You must fulfill your duty of continued care.
- \_\_\_ 8. The plan may not work but the consequences of failure are the same as for the standard treatment.
- \_\_\_ 9. The standard treatment is not in best interest of the child.
- \_\_\_ 10. You should let the court should decide the best interest of the child.
- \_\_\_ 11. You cannot ignore the state's interest in the protection of human life.

**Please rank the best three choices by putting the item number (1-11) in the blank**

\_\_\_ Best choice

\_\_\_ 2<sup>nd</sup> best

\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-11) in the blank**

\_\_\_ Worst choice

\_\_\_ 2<sup>nd</sup> worst

## Case 4

Mrs. Rachel White is a 59 year old white female who retired from her job as a city clerk three years ago. At retirement, she elected to pay \$600.00 per month to keep her health insurance with the city. At age 50, she discovered a large lump in her left breast. Breast cancer with lymph node involvement was diagnosed. She had a radical mastectomy followed by radiation and chemotherapy.

Two years ago, Mrs. White began to experience pain in her right hip. She suspected arthritis. Cancer was diagnosed in the hip and in the head of the femur and also in the 3<sup>rd</sup> thoracic vertebrae. The city insurance refused to pay for bone marrow transplant necessary to accompany high dose chemotherapy and so she was given radiation and standard chemotherapy. This slowed the progression of the cancer but three months ago a bone scan showed additional lesions in the backbone at T2 and T4 and in two ribs. She has difficulty walking because of numbness in her feet. Her left arm has some numbness, too. She has severe pain in her back that never goes away and her oncologist says that it will get worse as the cancer progresses.

Mrs. White's concern is that the tumor in her back will paralyze her before the cancer can kill her. She does not want to be a burden on her family and she does not want to be severely drugged for pain in her last couple of months of life.

Mrs. White asks you for a prescription that will let her die peacefully and with some dignity.

**What do you do about Rachel White's request?**

Case 4: Action Choices

**For each of the action choices below, circle the number that corresponds to your view about each choice. Rate each choice using the following scale:**

**Action choices:**

- 1. I strongly believe that this is GOOD choice**
- 2. I believe that this is a GOOD choice**
- 3. I am not sure**
- 4. I believe that this is a BAD choice**
- 5. I strongly believe that this is a BAD choice**

- \_\_\_ 1. Agree to schedule another appointment with Mrs. White as possible to further discuss the issue.
- \_\_\_ 2. Decline to discuss this issue with Mrs. White again.
- \_\_\_ 3. Ask permission to schedule a meeting for Mrs. White with several “experts” in pain management end of life care.
- \_\_\_ 4. Respect her autonomy by agreeing to her wishes.
- \_\_\_ 5. Recommend that Mrs. White and her family seek joint counseling before any action is taken.
- \_\_\_ 6. Explain to Mrs. White how legal prescription analgesics can used for her purpose.
- \_\_\_ 7. Reassure Mrs. White that you will aggressively treat pain even to the point where the medications may shorten her life.
- \_\_\_ 8. Discuss with her possibility that her discussion could be influenced by treatable depression.
- \_\_\_ 9. Reassure Mrs. White that you will not abandon her.
- \_\_\_ 10. Recommend that she join a support group.

**Please rank the best three choices by putting the item number (1-11) in the blank**

\_\_\_ Best choice

\_\_\_ 2<sup>nd</sup> best

\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-10) in the blank**

\_\_\_ Worst choice

\_\_\_ 2<sup>nd</sup> worst

#### Case 4: Justification for Action Choices

For each of the reasons below, circle the number that corresponds to your view about each choice. Rate each reason using the following scale:

Justification for action choices:

1. I strongly believe that this is GOOD choice
2. I believe that this is a GOOD choice
3. I am not sure
4. I believe that this is a BAD choice
5. I strongly believe that this is a BAD choice

- \_\_\_ 1. Her suffering will be needlessly prolonged by waiting.
- \_\_\_ 2. It is important to respect the integrity of personal choice and self-determination.
- \_\_\_ 3. You should stay within the standards of practice to avoid claims of malpractice.
- \_\_\_ 4. You must balance respect for life with quality of life.
- \_\_\_ 5. The community's moral stance is less important than the patient's wishes.
- \_\_\_ 6. Physician-assisted suicide is nothing more than a license to kill.
- \_\_\_ 7. You don't have the expertise to deal with this issue.
- \_\_\_ 8. It is important to respect the survival and integrity of human life.
- \_\_\_ 9. Your opinions are not a substitute for her opinions of her best interest.
- \_\_\_ 10. Mrs. White may not be aware of the resources available to her.
- \_\_\_ 11. The request for assistance in dying may not reflect the patient's underlying needs.
- \_\_\_ 12. You and Mrs. Whitey may not fully understand her prospects for a life of appreciable length and quality.
- \_\_\_ 13. You should carefully examine the rationale for Mrs. White's request.
- \_\_\_ 14. Clinical research in psychology supports your decision.

**Please rank the best three choices by putting the item number (1-14) in the blank**

\_\_\_ Best choice

\_\_\_ 2<sup>nd</sup> best

\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-14) in the blank**

\_\_\_ Worst choice

\_\_\_ 2<sup>nd</sup> worst

## Case 5

Eli Smith is a 60 year old male who lives with his 84 year old mother in a poor part of town. He has been your patient for one year. He has had two other doctors during the last four years.

He started working in coal mines when he was 17. Five years ago, he was diagnosed with Chronic Obstructive Pulmonary Disease (COPD). He received a cash settlement for “Black Lung” from the mining company but he spent all of that money and now he relies on his elderly mother to take care of him. His wife died 10 years ago and he never talks with his children.

He does not take good care of his health. He smokes two packs of cigarettes per day. He gets drunk with his friends every weekend and then he has difficulty breathing. He refuses to take his medicine because he says that the doctors don’t know what they are doing and because the drug companies are “cheating people”. Without doubt, he is the most difficult and disagreeable patient that you have.

He is hospitalized one or two times every month. Yesterday, he was hospitalized again after an all night smoking and drinking binge.

You don’t want Mr. Smith as a patient.

What will you do with this patient?

**Case 5: Action Choices**

**For each of the action choices below, circle the number that corresponds to your view about each choice. Rate each choice using the following scale:**

**Action choices:**

- 1. I strongly believe that this is GOOD choice**
- 2. I believe that this is a GOOD choice**
- 3. I am not sure**
- 4. I believe that this is a BAD choice**
- 5. I strongly believe that this is a BAD choice**

- \_\_\_ 1. Sedate and restrain Mr. Smith so that he can rest and recover.
- \_\_\_ 2. Refer him to your least favorite colleague.
- \_\_\_ 3. Try harder to educate him about his health care.
- \_\_\_ 4. Refuse to continue treatment unless Mr. Smith complies with your treatment guidelines.
- \_\_\_ 5. Refer him to a county “free clinic”.
- \_\_\_ 6. Continue to provide care under the current conditions.
- \_\_\_ 7. Refer him to a pulmonary specialist.
- \_\_\_ 8. Discontinue care and let Mr. Smith find another doctor.
- \_\_\_ 9. Confront Mr. Smith’s resistance and insist on making a contract around his care.

**Please rank the best three choices by putting the item number (1-9) in the blank**

\_\_\_ Best choice

\_\_\_ 2<sup>nd</sup> best

\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-9) in the blank**

\_\_\_ Worst choice

\_\_\_ 2<sup>nd</sup> worst

**Case 5: Justification for Action Choices**

**For each of the reasons below, circle the number that corresponds to your view about each choice. Rate each reason using the following scale:**

**Justification for action choices:**

- 1. I strongly believe that this is GOOD choice**
- 2. I believe that this is a GOOD choice**
- 3. I am not sure**
- 4. I believe that this is a BAD choice**
- 5. I strongly believe that this is a BAD choice**

- \_\_\_ 1. You have to fulfill your fiduciary responsibilities to this patient.
- \_\_\_ 2. It is important to provide for his best interest including his clinical welfare.
- \_\_\_ 3. The accepted standards of care depend on the patient.
- \_\_\_ 4. You have a responsibility to care for other patient.
- \_\_\_ 5. The time and attention taken from other patients cannot be justified by the clinical outcome for this patient.
- \_\_\_ 6. The clinical objectives for the care of this patient are not being met.
- \_\_\_ 7. The cost to hospital or insurance company cannot be justified.
- \_\_\_ 8. It is unlikely that Mr. Smith will comply with treatment guidelines.
- \_\_\_ 9. It is not fair to other patients.
- \_\_\_ 10. It is important to show respect for Mr. Smith's autonomy.

**Please rank the best three choices by putting the item number (1-10) in the blank**

\_\_\_ Best choice

\_\_\_ 2<sup>nd</sup> best

\_\_\_ 3<sup>rd</sup> best

**Please rank the worst two choices by putting the item number (1-10) in the blank**

\_\_\_ Worst choice

\_\_\_ 2<sup>nd</sup> worst

APPENDIX C  
REGRESSION ANALYSIS

Table 5. Summary of Hierarchical Regression Analysis for Variables and Interactions Predicting MD-ICM score ( $N = 540$ )

<b>Pscore</b>	<b>B</b>	<b>SE B</b>	$\beta$	$R^2$
<i>Total Ranking Index</i>				
<i>Step 1</i>				
Gender	.05	.02	.13	.04
PSCORE	.00	.00	.16	
<i>Step 2</i>				
Gender*PSCORE	.00	.00	.03	
<i>Good Ranking Index</i>				
<i>Step 1</i>				
Gender	.06	.02	.13	.05
PSCORE	.00	.00	.18	
<i>Step 2</i>				
Gender*PSCORE	.00	.00	.03	
<i>Bad Ranking Index</i>				
<i>Step 1</i>				
Gender	.04	.02	.09	.02
PSCORE	.00	.00	.09	
<i>Step 2</i>				
Gender*PSCORE	.00	.00	.03	
<i>Action Choice Index</i>				
<i>Step 1</i>				
Gender	.07	.02	.14	.06
PSCORE	.01	.00	.21	
<i>Step 2</i>				
Gender*PSCORE E	.00	.00	.00	
<i>Justification Index</i>				
<i>Step 1</i>				
Gender	.03	.02	.08	.01
PSCORE	.00	.00	.05	
<i>Step 2</i>				
Gender*PSCORE	.00	.00	.07	

Table 5 Summary of Hierarchical Regression Analysis for Variables and Interactions  
Predicting MD-ICM score ( $N = 540$ )

<b>N2SCORE</b>	<b>B</b>	<b>SE B</b>	$\beta$	$R^2$
<i>Total Ranking Index</i>				
<i>Step 1</i>				
Gender	.05	.02	.13	.07
N2SCORE	.01	.01	.22	
<i>Step 2</i>				
Gender*N2SCORE	.00	.00	.01	
<i>Good Ranking Index</i>				
<i>Step 1</i>				
Gender	.06	.02	.14	.07
N2SCORE	.01	.00	.21	
<i>Step 2</i>				
Gender*N2SCORE	.00	.00	.04	
<i>Bad Ranking Index</i>				
<i>Step 1</i>				
Gender	.04	.02	.08	.04
N2SCORE	.01	.00	.18	
<i>Step 2</i>				
Gender*N2SCORE	.00	.00	-.04	
<i>Action Choice Index</i>				
<i>Step 1</i>				
Gender	.07	.02	.14	.08
N2SCORE	.01	.00	.24	
<i>Step 2</i>				
Gender*N2SCORE	.00	.00	-.03	
<i>Justification Index</i>				
<i>Step 1</i>				
Gender	.03	.02	.08	.02
N2SCORE	.00	.00	.13	
<i>Step 2</i>				
Gender*N2SCORE	.00	.00	.06	

Table 5 Summary of Hierarchical Regression Analysis for Variables and Interactions  
Predicting MD-ICM score ( $N = 540$ )

<b>STAGE4P</b>	<b>B</b>	<b>SE B</b>	$\beta$	$R^2$
----------------	----------	-------------	---------	-------

---

<i>Total Ranking Index</i>				
<i>Step 1</i>				
Gender	.00	.00	.21	.06
STAGE4P	.05	.02	.13	
<i>Step 2</i>				
Gender*STAGE4P	.00	.05	-.07	
<i>Good Ranking Index</i>				
<i>Step 1</i>				
Gender	.00	.00	.21	.06
STAGE4P	.06	.02	.13	
<i>Step 2</i>				
Gender*STAGE4P	.00	.00	-.07	
<i>Bad Ranking Index</i>				
<i>Step 1</i>				
Gender	.00	.00	.17	.04
STAGE4P	.04	.02	.08	
<i>Step 2</i>				
Gender*STAGE4P	.00	.00	-.04	
<i>Action Choice Index</i>				
<i>Step 1</i>				
Gender	.00	.00	.17	.05
STAGE4P	.07	.02	.14	
<i>Step 2</i>				
Gender*STAGE4P	-.00	.00	-.10	
<i>Justification Index</i>				
<i>Step 1</i>				
Gender	.00	.00	.21	.05
STAGE4P	.03	.02	.07	
<i>Step 2</i>				
Gender*STAGE4P	-.00	.00	-.00	

---

APPENDIX D  
INFORMED CONSENT FORM

UNIVERSITY OF ALABAMA  
INFORMED CONSENT

You are being invited to take part in a research project which will be used to help develop an instrument called “An ICM Approach to the Assessment of an Ethics Intervention in Medical Education in Thailand.” The study is being done by Ms. Ponsopa Piniijphon who is a doctoral student in Educational Psychology at the University of Alabama. Ms. Piniijphon is being supervised by Dr. Stephen J. Thoma, who is a professor of Educational Psychology at The University of Alabama.

**PURPOSE:** You are invited to participate in a research project for a new instrument which measures ethical thinking on moral dilemmas in medicine.

**PROCEDURES AND TIME COMMITMENT:** If you decide to participate in this project, you will be asked to do these things: You will fill in the consent form. Then you will be asked for some demographic information such as age, gender, GPA, ethnicity and level of education. Then you will complete the survey. The test will take approximately 45 minutes. Note: All participants should be age 19 years or older.

**CONFIDENTIALITY:** If you choose to participate, all of your data will be transferred to investigators without personal identifiers. There will be no way to identify you from any of the information, including the responses to demographic information. All of the responses will be used for data analysis and archival purposes only.

**YOUR PARTICIPATION:** Your participation in this study is voluntary. You may refuse to participate. You may withdraw from the test at any time.

**RISK AND BENEFITS:** There are no potential risks from participating in this survey. Your current skills are sufficient to participate in this survey. You are not expected to be an expert in medicine or on medical dilemmas. You will not be paid for being in this study. Your participation is expected to further research on ethical thinking in medicine.

**FOR QUESTIONS ABOUT THE RESEARCH CONTACT:** Ponsopa Piniijphon at 205-246-6643 or by e-mail at pinij001@bama.ua.edu. This consent explains the research study. Please read it carefully. If you have any questions or if there is anything that you do not understand, then ask your instructor or Dr. Thoma or email the researcher, Ponsopa. You can also reach Dr. Thoma by e-mail at sthoma@bamaed.ua.edu. You can ask questions later even if you do not ask questions now.

If you have any questions about your rights as a participant in a research project, you should contact Ms. Tanta Myles, Research Compliance Officer, at 205-348-5152 or by e-mail at cmyles@fa.ua.edu.

**PARTICIPANT STATEMENT:** I have read the explanation provided to me. I have had all my questions answered to my satisfaction, and I voluntarily agree to participate in this research project.