Engineering Ethical Curricula:

Assessment and Comparison of Two Approaches

Matthew J. Drake

School of Industrial and Systems Engineering

Georgia Institute of Technology

Paul M. Griffin

School of Industrial and Systems Engineering

Georgia Institute of Technology

Robert Kirkman

School of Public Policy

Georgia Institute of Technology

Julie L. Swann

School of Industrial and Systems Engineering

Georgia Institute of Technology

ABSTRACT

The paper assesses two approaches for delivery of engineering ethics: a full semester ethics course and an engineering course that includes an ethics module. The Defining Issues Test was used to compare the improvement of a student's moral reasoning ability in each class as compared to a control class. Our findings were that the module approach used did not provide any improvement in moral reasoning. In addition, although the full ethics course showed improvement when compared to the module, it was not significantly different from the control class. We also found that there was little distinction between males and females and no distinction by age, although education level did have an impact. The results suggest that to improve a student's moral reasoning and sensitivity to ethical issues, engineering ethics must be integrative, delivered at multiple points in the curriculum, and incorporate specific discipline context.

Keywords: engineering ethics, content delivery, assessment

I. INTRODUCTION

A recent survey of 254 engineering programs found that only about one in five engineering students that graduated from accredited programs in 1996-1997 were exposed to any significant ethics content.[1] In more recent years, the Accreditation Board for Engineering and Technology (ABET) requires that engineering programs introduce students to ethical issues that arise from the practice of engineering.[2] In addition, in light of the many high profile news stories about unethical business practice, many industry and academic leaders have stressed the importance of increasing ethics content in engineering courses.[3] As a result, there has been a great deal of recent literature describing how programs are integrating ethics into their curriculum. For example, Haws[4] reviewed 42 papers given from 1996-1999 at the American Society for Engineering Education (ASEE) conferences. At the 2004 ASEE Annual Conference and Exposition, over 25 papers were presented on the topic of engineering ethics. The vast majority of this literature focuses on content such as ethics cases or new course development; very little research has been done on assessing and comparing the impact of various delivery methods.

Two common approaches to teaching engineering ethics[5] are the incorporation of modules or cases into existing courses and the offering of a stand-alone course, often by an academic unit outside college of engineering. As engineering curricula already tend to be crowded, adding an additional course may prove difficult. It is not clear, however, that adding a few ethics-based modules to existing courses will sufficiently develop ethical sensitivity and judgment in the student. On the other hand, a benefit of the module approach is that it is much easier to tailor to the student's discipline; an attribute that many argue is essential.[6]

In this paper, we present a comparative assessment of these two general approaches to teaching professional ethics to undergraduate engineering students. The assessment focuses particularly on the effectiveness of each mode in improving moral or ethical judgment. (We will use the terms 'moral' and 'ethical' interchangeably to indicate reasoning or judgment concerning values and obligations in any context). Currently, the College of Engineering at the Georgia Institute of Technology (Georgia Tech) uses both approaches to teaching engineering ethics. About half of the schools, including Mechanical Engineering and Electrical and Computer Engineering, require students to take a three-credit, semester-length course related to ethics; students choose this course from a menu of options offered by the School of Public Policy and the School of International Affairs, among others. The remaining schools in the College of Engineering have opted for the introduction of ethics-related modules directly into required engineering courses, which is the approach taken by the School of Industrial and Systems Engineering (ISyE).

The comparison that we perform is between students that took a full course on ethics entitled "Ethics and the Technical Professions" (ETP) offered by the School of Public Policy, and those that took an ISyE course entitled "Modeling in Industrial Engineering" (MIE) which only contained an ethics module as part of its content.

Two previous studies of engineering ethics education have informed our research. Steneck[7] described the process used by the University of Michigan in implementing engineering ethics through their curriculum. In the study he proposed four assessment tools: i) basic knowledge – through questioning (e.g., examinations), ii) reasoning abilities – through testing of use of reasoning tools (e.g., essays and cases), iii) self evaluation – through questioning about preparation to assume professional responsibilities, and iv) professional evaluation – through feedback from engineering professionals. The assessment has not yet been completed, and as Steneck points out, one limitation of the proposed assessment framework is that does not compare different approaches to delivery, but only how their new curriculum compares to no delivery. In addition, Steneck et al[8] provide a detailed set of educational objectives related to ethics and professionalism.

Self and Ellison[9] used the Defining Issues Test (DIT) to assess if there was an increase in moral reasoning from students that took an engineering ethics course. The researchers administered the test before and after the course, and measured the difference between scores. They found that there was a significant difference between the pre- and post-test, from which they drew the conclusion that moral reasoning skills can be taught and objectively measured. One limitation of the study is that a control group was not used in the comparison. It is possible, then, that the increase in moral reasoning was due to repeated taking of the test. In addition, as was the case with Steneck, the assessment was used to compare the benefit of an engineering ethics class to the *status quo ante* of no formal instruction in ethics.

We follow the approach of Self and Ellison in adopting the Defining Issues Test as our instrument. As with all such instruments, the DIT comes with built-in theoretical commitments that need to be well understood in order to properly interpret the data it yields. In this case, the DIT was developed on the basis of Kohlberg's cognitive theory of moral development. Rather than simply assimilating existing social norms, Kohlberg believed, each individual passes through a series of self-contained stages, each of which reflects a qualitatively different cognitive structure: an individual who has reached a higher stage conceives the relation of self to world differently than someone at a lower level.[10] Kohlberg initially identified six stages of development, from simple punishment-avoidance to principled reasoning about justice. Later, when he acknowledged there was little empirical basis for the existence of stage six reasoning, he reduced the number of stages to five.[11]

Kohlberg conducted his own research by way of in-depth interviews, which presented subjects with a production task: they were to make judgments and generate arguments in response to difficult – though somewhat contrived[12] – moral dilemmas. The first edition of the DIT was developed by James Rest and his collaborators as a simplified, "quick and dirty" version of the method, which takes the form of a multiple-choice test that is easy to administer and to score. Rather than a production task, the DIT presents subjects with a recognition task: they read six moral dilemmas, after each of which they are asked to both rate and rank the importance of statements that were derived from Kohlberg's own interview data. Statements were chosen from each of three levels – preconventional (stages 1 and 2), conventional (stages 3 and 4) and principled (stages 5 and, perhaps, 6); nonsensical statements were included as well in order to check respondent validity.[13]

Thirty years of testing with the DIT have established its validity and reliability by a number of measures; most importantly for our study, it has proven to be particularly sensitive to the effects of educational interventions.[13,9] Although the DIT was developed to measure general moral development, it has been found to measure moral development in educational programs of many types including medicine, business, and law among others. The testing also resulted in some refinements to the underlying theory of moral development, which in part provided the impetus for developing the second edition of the test, the DIT-2.

There have been two theoretical developments that are particularly important for interpreting the DIT and DIT-2. First, the creators of the DIT now conceive of moral development in terms of shifting distributions of schemas rather than a stepwise progression through self-contained stages.[14] The preconventional or "personal-interest" schema consists of elements that are referenced entirely to the narrow self-interest of the individual, while the elements of the conventional or "maintaining norms" schema are referenced to the demands of duty and the necessity of upholding the existing social order. The postconventional schema does not take the existing order for granted, but seeks rather to discover or generate sharable, fully reciprocal moral ideals that ought to serve as the basis of any social order.[14] At any point in development, an individual my draw from all three schemas, but will likely draw from one more than others.[13] The DIT and DIT-2 may thus be understood as revealing which schema predominates in the moral reasoning of each subject.

The second theoretical development has been an expansion of the scope of the postconventional schema. Kohlberg initially conceived of the highest levels of moral development entirely in terms of principled reasoning about justice, based on his own preference for a narrow subset of moral theory (e.g. Kant and Rawls).[10] A number of early critics took Kohlberg to task for this preference (especially Gilligan[15]), and Kolhberg himself recognized the necessity of broadening his understanding of the highest levels of moral reasoning.[11] Postconventional reasoning is more inclusive than principled reasoning, in that moral ideals may be drawn from other moral theories, e.g., virtue ethics, the ethics of care, and so on. In short, the postconventional schema is no longer meant to indicate a preference for a particular moral theory, but a general cognitive stance regarding the relationship between the individual and society, a standpoint from which it is possible to critically assess prevalent social norms.[13]

Using the DIT-2 imposes an important limitation on this study in that it focuses entirely on moral judgment, that is, how the subjects think about moral problems and arrive at decisions. Judgment is only one component of moral conduct. The creators of the DIT and DIT-2 have themselves proposed a four-component model to account for moral conduct, including moral sensitivity, moral motivation, and moral character alongside judgment.[13,16] In order to behave morally, an individual must first recognize the ethical aspects of a given situation (sensitivity, or imagination), determine what would be the right thing to do (judgment), decide to do it (motivation), and have the integrity and courage to follow through (character).[17]

II. METHODS

A. Study Population and Research Design

Three classes taught at Georgia Tech were studied with varying levels of ethical instruction.[1] The students self-enrolled in the classes; it was not possible to randomly assign them, thus the design is quasi-experimental. For each student, we tested their ethical reasoning at the beginning of the school term (prior to ethical instruction) and at the end of the term (after ethical instruction), and we matched the pre and post-tests. Tests were deleted if the individual was not at least 18 years old, if they answered randomly, or if they were not an undergraduate student). The overall number of students is shown in Table 1.

The total remaining sample who took both the pre- and post-test includes 264 students (197 males, 66 females, and one unspecified respondent) (Table 2). The majority of the students (251) are in engineering, with a few (13) in science, management, or liberal arts. Most are concentrated in mid to upper classes, with Freshmen (8), Sophomores (82), Juniors (99) and Seniors (70) comprising the entire group (Table 1). On the measure of the P score or the N2 index (described in the Measures Section), the participants who took the pre-test but not the post-

[1] Note that approval was obtained by the Georgia Tech's Institutional Review Board (IRB) for human subjects research prior to the study.

test were not significantly different from participants who completed both, either as a group or within each course.

Students who enrolled in "Ethics and the Technical Professions" (ETP) (N = 129 of final sample) received a full semester on ethical reasoning in the context of professional topics offered by the School of Public Policy and taught by a faculty member with a Ph.D. in Philosophy. 164 students were enrolled in the Spring 2004 semester, with six Public Policy graduate students serving as teaching assistants. The class met in a large lecture section on Mondays and Wednesdays, with smaller, TA-led discussion sections on Fridays. The materials for the course were largely drawn from the second edition of Harris, Pritchard, and Rabins' *Engineering Ethics: Concepts and Cases*.[17] Much of the class was focused on helping students develop their skills in analyzing cases, generating a range of possible solutions, and building an argument for their preferred solutions. Weekly writing assignments, three argumentative essays, and an essay-format final exam were likewise based on case studies.

The primary difference between the ETP class and the others studied is the length of time spent on ethics (approximately 15 weeks in all). In addition, this course is not required for any major, but it may be used to fulfill the ethics requirement of several of the engineering curricula or the humanities component of the Institute's general education requirement, or both. Thus there may be some self-selection bias in who chooses to take this class, i.e., they may have a pre-existing interest in ethical dilemmas. We should point out, however, that this is the course normally taken by engineering majors with ethics requirements, and so any potential bias on this point would not be large. In addition there may be some age bias in who takes ETP, with more seniors in this class than the other classes.

A second subset of students (N = 109) took "Modeling in Industrial Engineering" (MIE), which is a general introduction to the types of models (conceptual and computational) and methodologies used in the industrial engineering field. These students received two lectures on ethics based on material from Solomon and Martin. [18] In addition, students worked through two short ethics cases in class and had an ethics case given as a group lab assignment. The cases came from Velasquez[19,20]. The final subset of students (N = 26) served as the control group, with no ethical instruction at all in an industrial engineering class called "Probability with Applications" (PA). These two courses are required for all Industrial Engineering students and are generally taken in the sophomore year. They are taught by faculty in the School of Industrial and Systems Engineering.

Kohlberg's moral development theory was not addressed in either of the experimental classes. Some researchers have concluded that teaching Kohlberg theory directly results in additional movement along the scale of moral development, although this result could be from students answering questions how they perceive they should be answered rather than from actual changes in thought. Because of this, many researchers using the DIT do not cover Kohlberg theory intentionally.

Course:	**ETP**	**MIE**	**PA (Control)**	Total
Students enrolled (N)	164	148	86	398
Pre-test only (N)	21	31	28	80
Pre response rate	91.5%	94.6%	62.8%	86.4%
All (N)	129	109	26	264
Final response rate	78.7%	73.6%	30.2%	66.3%

Table 1: Enrollment and responses of participants

Course:	ETP	MIE	PA (Control)	Total
Majors				
Industrial Engineering	4	108	23	135
Other Engineering	114	0	2	116
Non-Engineering	11	1	1	13
Educational Level				
Freshman	3	1	4	8
Sophomore	10	59	13	82
Junior	55	39	5	99
Senior	60	7	3	70
Other	1	3	1	5
Sex				
Male	113	67	17	197
Female	16	41	9	66
No response	0	1	0	1
All	129	109	26	264

Table 2: Demographics of participants who completed full experiment

B. Measures

The ethical reasoning of each student was assessed at the beginning and end of the Spring 2004 semester using the DIT-2. The P score is a numerical index of moral reasoning developed originally by Kohlberg. The P score indicates the prevalence of post-conventional thinking on the part of each subject given as a percentage from 0 to 95. It is calculated by summing the scores from test answers corresponding to Stages 5 and 6, converted to a percentage.

In developing the DIT-2, Rest and his collaborators have also developed a new index to measure moral reasoning, the N2 score, which takes into account both the prevalence of postconventional reasoning and the avoidance of preconventional reasoning or personal interest schema.[21,22] The N2 index is normalized to the same scale as the P score for comparisons. The N2 score employs both rating and ranking data from the test and typically purges more data points than the P index due to missing data rules.

The third measure of moral reasoning we report is a type indicator indicating the moral reasoning schema preferred by a test-taker and whether the person is primarily in that scheme ("consolidated") or moving to/from another schema ("transitional"). The type indicators are identified here.[23] In general, it is expected that people move up in types as they move forward in moral development.

Type 1: predominant in **Personal Interests** and consolidated

Type 2: predominant in **Personal Interests** but transitional

Type 3: predominant in **Maintaining Norms** but transitional, **Personal Interests** secondary

Type 4: predominant in **Maintaining Norms** and consolidated

Type 5: predominant in **Maintaining Norms** but transitional, **Post Conventional** secondary

Type 6: predominant in **Post Conventional** but transitional

Type 7: predominant in **Post Conventional** and consolidated

One measure that we calculate for each of the P, N2, and Type scores, is the "effect size" to indicate the effect of ethical instruction between each of the experimental classes compared to the control group. The effect size is the difference between the experimental and control groups on the change from pre test to post test, divided by the pooled standard deviation. Specifically: Effect size of class *j* compared to control group *k*= $(\mu_j^{diff} - \mu_k^{diff}) / \sigma_{jk}^{diff}$, where μ_j^{diff} is the mean over the individuals in class *j* of the post test measurement – the pre test measurement for each individual and σ_{jk}^{diff} is the pooled standard deviation. This is a common way of reporting ethical reasoning scores in experiments with control groups, and it allows us to compare to other studies.

Finally, we use other measures to test correlation between the ethical reasoning scores and political or social characteristics. For example, a variable (HumLib) has been developed[24,25], which represents the humanitarian liberal perspective. It is calculated by counting the number of times that a test-taker responds the same as professionals in political science or philosophy, and may correlate with political beliefs. Since students answer a question about their self-declared political beliefs, we can check for this correlation using the variable ConLib (values of 1 – 5, ranging from more liberal to more conservative). Correlations between reasoning scores and religious beliefs can be measured by examining a measure called Cancer10. This measure is a proxy for religious beliefs, calculated by summing the ratings and rankings corresponding to test items related to God's role in a decision, and has been found in the past to correlate with religious beliefs (see Narvaez et al.[26]).

C. Statistical Analysis

A *t*-test with equal variance was used to test for differences between groups on measures such as the P score, N2 score, and Type. A Pearson correlation coefficient was used to test for significance of correlations. Significance levels are at the 5% level unless otherwise indicated.

III. RESULTS

A. Did the ethical instruction result in an improvement in moral reasoning?

In order to evaluate the effect of ethical instruction, the comparison was made between individual students' scores on the DIT pre- and post-tests. A *t*-test conducted on the scores for students who completed the entire experiment versus those who only took the pre-test found no statistical difference (at the 5% level of significance) between these two groups (results not shown). Consequently, the scores of students who completed only the pre-test were omitted from the evaluation of ethical instruction without producing a strong bias. Tables 3 and 4 show the mean pre- and post-test scores for students stratified by class as well as the difference between the scores on each test.

Recall that the effect size measure estimates the effect of ethical instruction on each of the experimental groups compared to the control group. The effect sizes for the P and N2 scores in the ETP class indicate that the full-term ethics course had a small effect on these measures (since the effect sizes are less than 0.2). Contrariwise, the effect sizes for the MIE class suggest that the integrated ethics modules actually had a small negative impact on the DIT measures. The *t*-tests comparing the pre- and post-test differences for each of the experimental classes to the control group difference found that neither form of ethical instruction produced an improvement in any of the measures over that seen in the control group. (All of the *p*-values of these tests were greater than 0.25.) This suggests that any improvement seen from the pre-test to the post-test is primarily due to the students' familiarity and experience with the test and not the result of ethical instruction. Consequently, the effect sizes described above are not statistically significant; therefore, they must be assumed to be the product of random sampling variation.

Course	**N**	**Pre P Score**	**Post P Score**	***Post - Pre P***	**Pre N2 Score**	**Post N2 Score**	***Post - Pre N2***	**Pre Type**	**Post Type**	***Post - Pre Type***
ETP	129	38.37 (1.29)	42.12 (1.36)	*3.75 (1.20)*	38.50 (1.29)	43.49 (1.31)	*5.00 (1.03)*	4.93 (0.16)	5.26 (0.16)	*0.33 (0.17)*
PA (Control)	26	33.85 (2.50)	36.08 (2.82)	*2.23 (2.45)*	36.41 (2.87)	39.39 (2.89)	*2.98 (2.13)*	4.62 (0.34)	5.00 (0.29)	*0.39 (0.39)*
Effect				*0.11*			*0.18*			*-0.03*

Table 3: Average (standard error) of Pre Test and Post Scores for ETP and the Control group

Course	N	**Pre P Score**	**Post P Score**	***Post - Pre P***	**Pre N2 Score**	**Post N2 Score**	***Post - Pre N2***	**Pre Type**	**Post Type**	***Post - Pre Type***
MIE	109	34.16 (1.35)	33.35 (1.40)	*-0.81 (1.26)*	36.27 (1.34)	35.97 (1.49)	*-0.30 (1.20)*	4.66 (0.17)	4.35[2] (0.18)	*-0.29 (0.18)*
PA (Control)	26	33.85 (2.50)	36.08 (2.82)	*2.23 (2.45)*	36.41 (2.87)	39.39 (2.89)	*2.98 (2.13)*	4.62 (0.34)	5.00 (0.29)	*0.39 (0.39)*
Effect				*-0.23*			*-0.27*			*-0.35*

Table 4: Average (standard error) of Pre Test and Post Scores for MIE and the Control group

Even though the type changes were not statistically significant for any course, it is still interesting to examine the changes in type graphically for each group. Figure 1 displays histograms of the type frequencies for each class on the pre- and post-tests, and Figure 2 shows the change in frequency for each type from the pre-test to the post-test. Somewhat surprisingly, the ETP class displayed a large increase of students in type 7, which is generally thought of as an advanced, consolidated form of moral reasoning that few people exhibit. The MIE class had a large increase in type 2, which is a transitional profile that emphasizes personal interests. This suggests that students in this class switched from a broader focus on society norms and postconventional absolutes to a more inward focus. This type is transitional, though, which indicates that this inwardness is not firmly entrenched in the students' moral reasoning; it is simply the way that they lean at the time, and as such not a cause for concern. The control group exhibited very little type change from the pre-test to the post-test, which was expected because they received no formal ethics instruction.

[2] The post-test Type score for MIE has two fewer observations than the pre-test score because the students' types were inconclusive from the test.

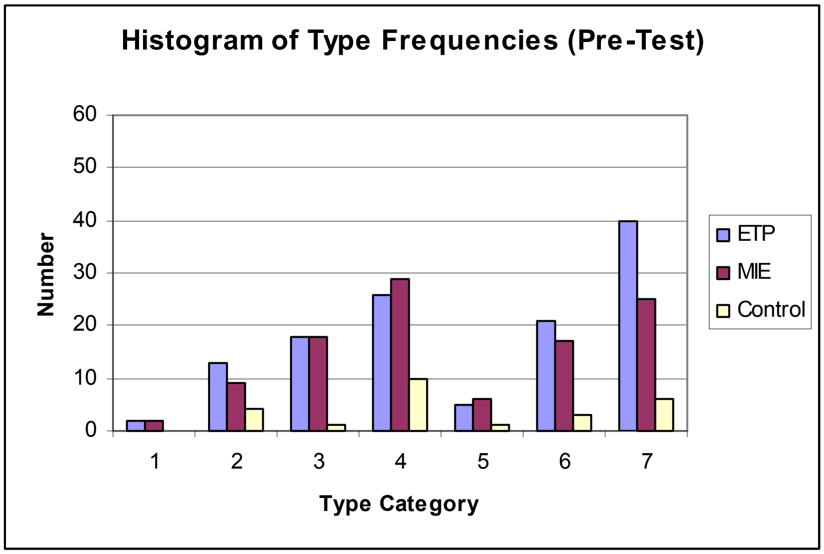

Figure 1a. Histogram of pre-test frequencies by course.

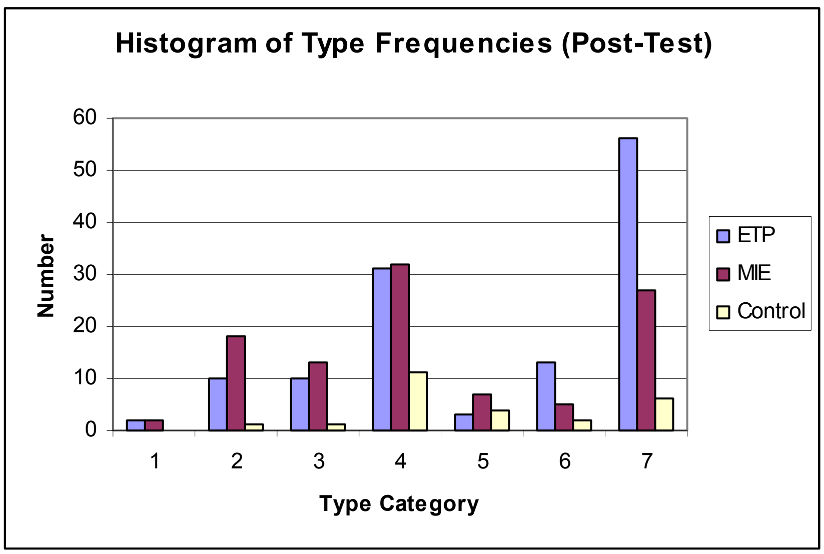

Figure 1a. Histogram of post-test frequencies by course.

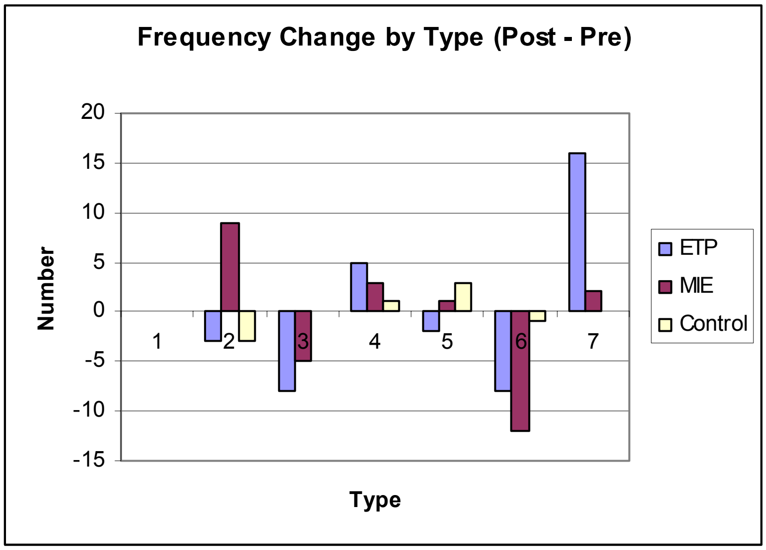

Figure 2. Change in frequency for each type from the pre-test to post-test.

B. Was the ethical reasoning development different by demographics?

In addition to measuring the impact of different types of ethical instruction, it is also interesting to stratify the student participants into different demographic groups. Comparisons were made between genders, grade levels, and majors (Table 5). The sample sizes for each of the categories differ from each other and from the overall sample size of 264 because some students failed to provide this classification data and some students were not undergraduates. Only the *p*-value for the type change was significant at the 5% level, which implies that the average male and female student's type changed differently between the pre-test and the post-test. The fact that only one measure had a significant difference between the two genders, suggests that there was little distinction between males and females in this stage of moral reasoning development and in their receptivity to ethical instruction. These results coincide with previous research that concludes that gender is not a significant factor in DIT performance.[24,27]

Because the majority of the students that participated in this study were sophomores and juniors, the few freshmen were grouped with the sophomores and the seniors with the juniors for the purposes of examining the effects of educational level. This classification also follows the traditional designation of freshmen and sophomores as "underclassmen" and juniors and seniors as "upperclassmen." The results show that the groups did not perform statistically differently on the pre-test but did exhibit a difference on the post-test and in their improvement from the pre-test to the post-test. This suggests that the juniors and seniors were either more receptive to the ethical instruction or benefited more from taking the test a second time. It is also important to note that the educational level designation is key here, because an age variable was not significantly correlated with any of the scores. This provides evidence that educational level has more of an effect on moral reasoning than age does. These results exhibit a bias, however, because most of the upperclassmen were enrolled in the ETP (full-length) course while underclassmen mainly comprised the MIE course and the control group. Too few students were enrolled in each course's minority educational level to examine the effects of this bias formally.

Since engineers are generally seen as having a more technical mindset than their counterparts in the humanities, comparisons were analyzed between the engineering students' scores and those in non-engineering majors. The only significant p-value between the two groups is the difference in the N2 score from the pre-test to the post-test, although other p-values were very close to the 5% level of significance. These results imply that engineering students are not significantly different from their counterparts in other majors in their ethical reasoning ability and receptivity to ethical instruction. This challenges the norm that political science and other liberal arts students are the "expert" group for performance on the DIT.[13] Further study about this claim is recommended, though, because the sample size for non-engineering students was quite small compared to the engineering sample. There also may be some bias in the statistics, as 85% of the non-engineering students were enrolled in the ETP full course on ethics.

Sex	N	Pre P Score	Post P Score	*Post - Pre P*	Pre N2 Score	Post N2 Score	*Post - Pre N2*	Pre Type	Post Type	*Post - Pre Type*
Male	197	36.58	39.00	*2.41*	37.74	40.85	*3.12*	4.73	4.97	*0.25*
Female	66	34.95	34.68	*-0.27*	36.36	37.40	*1.05*	4.94	4.51	*-0.42*
P-value of difference between genders		0.42	0.05	*0.16*	0.50	0.12	*0.23*	0.41	0.08	***0.02***
Grade Level										
Freshmen and sophomores	90	35.83	35.02	*-0.81*	37.41	37.32	*-0.09*	4.86	4.80	*-0.06*
Juniors and seniors	169	36.84	39.79	*2.95*	37.76	41.85	*4.09*	4.79	4.95	*0.17*
P-value of difference between grade levels		0.59	**0.02**	***0.03***	0.85	**0.02**	***0.01***	0.79	0.54	*0.38*
Major										
Engineering	251	36.38	37.85	*1.46*	37.75	39.97	*2.22*	4.82	4.84	*0.03*
Non-engineering	13	32.46	39.08	*6.62*	30.07	40.15	*10.08*	3.00	5.23	*1.08*
P-value of difference between majors		0.34	0.78	*0.18*	0.06	0.97	***0.02***	0.19	0.46	*0.06*

Table 5: Average of Pre-Test and Post Scores for Student Demographic Groups

C. How much do the ethical reasoning scores correlate with religious or political beliefs?

In order to examine the relationship between participants' demographic variables and their DIT performance, sample correlations were computed and their statistical significance analyzed. Table 6 contains only the demographic variables that were significantly correlated with the DIT scores. The HumLib sample size is smaller than the overall sample size because computation of this measure was inconclusive for certain students. The results suggest that a student's humanitarian liberal perspective is positively correlated with the P score. The more times a student's responses mimic professionals in political science and philosophy, who originally scored highly on the DIT and chose their actions consistently, the higher their P score will be on the DIT. On the other hand, religious orthodoxy is significantly negatively correlated with both scores. This is not surprising because Bebeau and Thoma[28] argue that this religious orthodoxy measure has the opposite effect as the humanitarian liberalism measure. The level of political conservatism was found to be uncorrelated with all of the DIT scores. All of these demographic variables were not significantly correlated with the improvement in each score from the pre-test to the post-test, suggesting that no one demographic was more likely to improve their scores through ethical instruction.

Demographic Variable	N	Pre P Score	Post P Score	Pre N2 Score	Post N2 Score
Humanitarian liberalism (HumLib)	239	0.1575 (0.0148)	0.1330 (0.0400)		
Religious orthodoxy (Cancer10)	264	-0.2186 (0.0003)	-0.1532 (0.0127)	-0.1679 (0.0062)	-0.1247 (0.0429)

Table 6: Significant sample correlations (p-values) between demographic variables and DIT scores.

D. How do our results compare to other studies?

Unlike the case of Self and Ellison[9], we do not show a significant impact of the full ethics course on moral reasoning. One possible reason for this difference was due to the fact that we included a control group in our study (since if the control is ignored we also see a significant impact on moral reasoning from the full ethics course). It is also possible that content played a role. For example, engineering ethics taught in the context of the student's discipline could have a larger impact than a general engineering ethics course. However, Self and Ellison did not describe the content of their course, and so this is speculation on our part.

The ETP effect sizes obtained in this study are consistent with those obtained in similar studies using the DIT.[29] In these studies, scores are compared for people who speak English primarily, see Table 7 for the scores in our study for this subset. Many of these studies reported effect sizes smaller than 0.20. The large negative effect sizes for the MIE course were surprising, although some other studies did report negative effect sizes. The *t*-test comparing the improvement from the pre-test to the post-test between the MIE course and the control group showed no significant difference between the courses, so this negative effect size is not significant.

Bebeau and Thoma[28] performed a meta-analysis on earlier studies of P scores for DIT-2. Comparison of averages (with only English speakers) to this meta-analysis studies shows that males in our study scored statistically significantly higher (39.00) than those in the meta-analysis (32.21) while scores for females in our study (34.68) were lower, though not significantly so than females in the meta-analysis (37.85).

Grade Level	N	Pre P Score	Post P Score	Pre N2 Score	Post N2 Score	Pre Type	Post Type
Freshman	6	32.67	38.00	35.76	37.04	3.67	5.00
Sophomore	77	36.71	35.58	38.57	38.57	4.97	4.90
Junior	83	38.70	41.21	39.48	42.73	4.83	4.94
Senior	63	36.81	39.40	38.05	43.01	4.87	5.05

Table 7: Average of Pre-Test and Post Scores for Primarily-English-Speaking Students

IV. CONCLUSIONS

Many engineering departments are dealing with the issue of how to incorporate engineering ethics into their already crowded curriculum without simultaneously increasing resource requirements. For this reason, it is extremely important to determine the most effective means of delivery.

Our findings suggest that a limited module is not sufficient for improving moral reasoning skills of engineering students. Previous studies have stated that a minimum of three weeks of ethics coverage is required to make an impact[29], and our study is consistent with these findings. However, we also find that a full course on ethics did not make a significant increase in moral reasoning (when compared to a control group). Some students have found a general course on ethics is not sufficient either--one student in the ETP course, writing an anonymous comment on the standard Georgia Tech course evaluation survey, expressed the view that outsourcing the ethics requirement to other schools in other colleges gave the impression that engineering schools place very little value on ethics education. This suggests that work needs to be done to construct some sort of comprehensive approach to ethics education for engineers.

Rest and Narvaez[30] review a number of professional programs and make some suggestions as to what components may help make a successful program. Among these are that integrating ethics across multiple courses is effective (see Duckett and Ryden[31] for example), and that the focus should be on all of the components of ethical behavior (see Bebeau[16] or Bebeau[32]), not just reasoning or judgment.

Any conclusions to be drawn from this study are subject to a number of important limitations. First, as noted earlier, the DIT-2 measures only one component of moral conduct; while it is surely important to help students to improve their moral judgment, it is just as important to foster students' ability to recognize situations that call for ethical judgment. There is a further problem that a general measure of moral judgment may not reflect the discipline-specific judgment that may be required in professional settings.[5] The effect of a course in professional ethics on this discipline-specific judgment could be greater than the effect on the more general judgment we measured. This suggests that it might be beneficial to develop a new instrument, perhaps modeled on the DIT-2, incorporating ethical dilemmas likely to be faced by engineers. Further, because discipline-specific judgment is bound together with sensitivity to ethical issues that may arise in professional practice, subsequent studies should include the development of a measure of ethical sensitivity that is also tailored to the engineering context.

Second, there were limitations of sampling in that a limited range of sizes and types of ethics classes or modular options were considered. It would be useful to compare these results with those attained from small sections (enrollment less than 40) of ETP, for example, or from small and large sections of courses in ethical theory that do not necessarily cover issues that arise from professional practice.

Finally, there is a limitation of time, in that our study spanned only one semester. Other studies have suggested that the effect of a single educational intervention on moral judgment may be short-lived.[33] For purposes of assessment, this suggests that some sort of longitudinal study may be useful, in part to determine to what extent a more sustained approach to ethics education can overcome the limitations of a single course or single module.

In the end, our study serves as an initial comparison with modest implications for curriculum design and implementation, implications that must be developed and tested further in subsequent studies at Georgia Tech and elsewhere. The study also provides evidence of the need to improve our methods for incorporating ethics education in engineering, and the research has also helped to identify some of the avenues for this improvement.

ACKNOWLEDGMENTS

This research was funded in part by a grant from the College of Engineering Undergraduate Initiative at the Georgia Institute of Technology. In addition, Dr. Swann was supported in part by NSF DMI-0348532.

We would like to acknowledge Dr. Harry Sharp for his help with the data analysis. Finally, we wish to acknowledge the help of the referees of this paper, whose comments strengthened the paper.

REFERENCES

1. Stephan, K.D., "A Survey of Ethics-Related Instruction in U.S. Engineering Programs," *Journal of Engineering Education*, vol. Oct, 1999, pp. 459-464.

2. ABET_Engineering_Accreditation_Commission. *Criteria for Accrediting Engineering Programs* 2002 [cited Sept 21, 2004. Available from http://www.abet.org/criteria_eac.html.

3. Stephan, K.D.,"Is Engineering Ethics Optional?" *IEEE Technology and Society Magazine*, Winter, 2002, 6-12.

4. Haws, D.R., "Ethics Instruction in Engineering Education: A (Mini) Meta-Analysis," *Journal of Engineering Education*, vol. April, 2001, pp. 223-229.

5. Harris, C.E., M. Davis, M.S. Pritchard, and M.J. Rabins, "Engineering Ethics: What? Why? How? and When?" *J. Engineering Education*, vol. 85, 1996, pp. 93-96.

6. Gorman, M.E.,"Turning Students into Ethical Professionals," *IEEE Technology and Society Magazine*, Winter, 2002, 21-27.

7. Steneck, N.H. 1999. Designing Teaching and Assessment Tools for an Integrated Engineering Ethics Curriculum. Paper read at 29th ASEE/IEEE Frontiers in Education Conference.

8. Steneck, N.H., B.M. Olds, and K.A. Neeley, "Recommendations for Liberal Education in Engineering," in *Proceedings of the 2002 American Society for Engineering Education Annual Conference and Exposition*, Liberal Education Division of the American Society for Engineering Education, http://asee.org/acPapers/2002-1963_Final.pdf, 2002.

9. Self, D.J., and E.M. Ellison, "Teaching Engineering Ethics: Assessment of Its Influence on Moral Reasoning Skills," *Journal of Engineering Education*, vol. 87, no. 1, 1998, pp. 29-34.

10. Kohlberg, L., "Stage and Sequence: The Cognitive-Developmental Approach to Socialization" in Kohlberg, L., ed., *The Psychology of Moral Development: The Nature and Validity of Moral Stages*, Harper and Row, San Fransisco, 1984

11. Kohlberg, L., C. Levine, and A. Hewer, "The Current Formulation of the Theory" in Kohlberg, L., ed., *The Psychology of Moral Development: The Nature and Validity of Moral Stages*, Harper and Row, San Fransisco, 1984

12. Weston, A., *A Practical Companion to Ethics*, 2nd ed, Oxford University Press, Oxford, 2002.

13. Rest, J., D. Narvaez, M.J. Bebeau, and S.J. Thoma, *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*, L. Erlbaum Associates, Mahwah, NJ, 1999.

14. Narvaez, D., and T. Bock, "Moral schemas and tacit judgement or how the Defining Issues Test is supported by cognitive science," *Journal of Moral Education*, vol. 31, no. 3, 2002, pp. 297-314.

15. Gilligan, C., *In a Different Voice: Psychological Theory and Women's Development*, Harvard University Press, Cambridge, Mass., 1982.

16. Bebeau, M.J., "The Defining Issues Test and the Four Component Model: Contributions to professional education," *Journal of Moral Education*, vol. 31, no. 3, 2002, pp. 271-295.

17. Harris, J., C. E., M.S. Pritchard, and M.J. Rabins, *Engineering Ethics: Concepts and Cases*, 2nd ed, Wadsworth, Belmont, CA, 2000.

18. Solomon, R.C., and M. Martin, *Above the Bottom Line: An Introduction to Business Ethics*, 3rd ed, Wadsworth/Thomson Learning, Belmont, CA, 2004.

19. Velasquez, M.G., *Business Ethics: Concepts and Cases*, 5th ed, Prentice Hall, Upper Saddle River, NJ, 2002.

20. Velasquez, M.G., C. Andre, T. Shanks, and M.J. Meyer,"Thinking Ethically, A Framework for Moral Decision Making," *Issues in Ethics*, Winter, 1996, 2-5.

21. Rest, J., S.J. Thoma, D. Narvaez, and M.J. Bebeau, "Alchemy and Beyond: Indexing the Defining Issues Test," *Journal of Educational Psychology*, vol. 89, no. 3, 1997, pp. 498-507.

22. Rest, J., D. Narvaez, S.J. Thoma, and M.J. Bebeau, "DIT2: Devising and Testing a Revised Instrument of Moral Judgment," *Journal of Educational Psychology*, vol. 91, no. 4, 1999, pp. 644-659.

23. Thoma, S.J., and J. Rest, "The relationship between decision-making and patterns of consolidation and transition in moral judgment development.," *Developmental Psychology*, in press.

24. Rest, J., *Development in judging moral issues*, University of Minnesota Press, Minneapolis, 1979.

25. Thoma, S.J., "An overview of the Minnesota approach to research in moral development," *Journal of Moral Education*, vol. 31, no. 3, 2002, pp. 225-245.

26. Narvaez, D., I. Geta, S.J. Thoma, and J. Rest, "Individual moral judgment and cultural ideology," *Developmental Psychology*, in press.

27. Thoma, S.J., "Do moral education programs facilitate moral judgment? A meta-anaylsis of studies using the Defining Issues Test," *Moral Education Forum*, vol. 9, no. 4, 1984, pp. 20-25.

28. Bebeau, M.J., and S.J. Thoma, "Guide for DIT-2," Center for the Study of Ethical Development, University of Minnesota, Minneapolis, 2003.

29. Schaefli, A., J. Rest, and S. Thoma, "Does moral education improve moral judgment? A meta-analysis of intervention studies using the Defining Issues Test," *Review of Educational Research*, vol. 55, no. 3, 1985, pp. 319-352.

30. Rest, J., and D. Narvaez, "Summary: What's possible?" in Rest, J.R. and D. Narvaez, eds., *Moral development in the professions: Psychology and applied ethics*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1994

31. Duckett, L.J., and M.B. Ryden, "Education for ethical nursing practice" in Rest, J.R. and D. Narvaez, eds., *Moral development in the professions: Psychology and applied ethics*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1994

32. Bebeau, M.J., "Influencing the moral dimension of dental practice" in Rest, J.R. and D. Narvaez, eds., *Moral development in the professions: Psychology and applied ethics*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1994

33. Weber, J., "Measuring the Impact of Teaching Ethics to Future Managers - a Review, Assessment, and Recommendations," *Journal of Business Ethics*, vol. 9, no. 3, 1990, pp. 183-190.

AUTHORS' BIOGRAPHIES

Matthew J. Drake is a graduate student in the School of Industrial and Systems Engineering at Georgia Tech. He received a B.S. in Business Administration from Duquesne University in 2002. His research interests are supply chain coordination, strategic procurement, and supply chain management education.

Address: School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0205; email: mdrake@isye.gatech.edu.

Paul M. Griffin is Associate Chair for Undergraduate Studies and Associate Professor in the School of Industrial and Systems Engineering at Georgia Tech. His Ph.D. is in Industrial Engineering from Texas A&M University. His research and teaching interests are in manufacturing, logistics and economic decision analysis.

Address: School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0205; telephone: 404-894-2431; fax: 404-894-2301; email: pgriffin@isye.gatech.edu.

Robert Kirkman is an Assistant Professor in the School of Public Policy at Georgia Tech, and serves as Director of the program in Philosophy, Science and Technology. His teaching responsibilities include environmental ethics, political philosophy, and moral theory in addition to engineering ethics. His primary research interests are in environmental philosophy and the ethics of metropolitan growth.
Address: School of Public Policy, Georgia Institute of Technology, Atlanta, GA, 30332-0205; telephone: 404-385-4258; email: robert.kirkman@spp.gatech.edu.

Julie Swann is an Assistant Professor in the School of Industrial and Systems Engineering at Georgia Tech. She received her Ph.D. in Industrial Engineering and Management Science from Northwestern University. Her research interests are in integrating methods from optimization and economics to solve problems in supply chain management and healthcare policy. Her teaching interests include these areas as well as improving ethics in engineering curricula. (Correspondence regarding the paper should be addressed to this author.)
Address: School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0205; telephone: 404-385-3054; fax: 404-894-2301; email: jswann@isye.gatech.edu.