Assessment of an Engineering Ethics Video: Incident at Morales

MICHAEL C. LOUI
Department of Electrical and Computer Engineering, and Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

ABSTRACT

A new engineering ethics video, Incident at Morales, was assessed with two different instruments: the standard Defining Issues Test and a short ad hoc survey. According to pre-tests and post-tests with the Defining Issues Test, viewers of the video increased the sophistication of their moral reasoning skills. According to the survey, viewers changed their opinions about the most important responsibilities of engineers and about meeting environmental regulations when working overseas. From these results, it appears that the video is an effective approach to teach engineering ethics.

Keywords: assessment, engineering ethics, video

I. INTRODUCTION

In universities across the United States, engineering departments are striving to meet the EC2000 criteria for engineering accreditation, which require that all engineering students demonstrate an understanding of professional and ethical responsibility. One effective method to teach engineering ethics is to use cases [1]. Cases, both text and video, are effective because they foster critical thinking, encourage student responsibility for learning, draw on both affective and cognitive skills, offer opportunities for collaboration among students, and enliven the classroom [2].

In engineering ethics, one of the most widely used fictional cases is Gilbane Gold, a video produced in 1989 by the National Institute for Engineering Ethics (NIEE) and Great Projects Film Company. NIEE and Great Projects collaborated again in 2003 to produce a new video, Incident at Morales. Whereas Gilbane Gold presents a whistle-blowing situation, Incident at Morales emphasizes everyday concerns. In the fictional case in Incident at Morales, a company plans to quickly build a new plant to manufacture a new paint remover. The company decides to construct the plant in Mexico to minimize the cost of environmental controls for the chemicals used and the byproducts produced in the manufacturing process. The process requires high temperatures and pressures, and it is designed to be automated and controlled by computer software. During the design and construction of the plant, legal, financial, safety, and ethical issues affect technical decisions about sensors, valves, piping, and environmental controls. While Incident at Morales portrays many ethical issues, it places special emphasis on three:

- Ethical considerations are an integral part of making engineering decisions.
- Although legal requirements may vary among states and nations, ethical obligations do not stop at state or national borders.
- Wherever engineers practice, they should strive to protect the health, safety, and welfare of the public.

One copy of Incident at Morales was sent free of charge to the dean of every engineering college in the United States. It is available for purchase in both VHS and DVD formats from the NIEE. A printed study guide accompanies both versions of the video, and additional materials are posted on the NIEE Web site.

Incident at Morales is 36 minutes long. It comprises three segments, with breaks for discussion after the first segment (12 minutes) and after the second segment (an additional nine minutes). The video is intended to be used in a single session of 90 to 120 minutes, or over two to three fifty-minute class sessions at a university. How much can a viewer learn about engineering ethics in this short time? It seems unlikely that a viewer’s moral reasoning skills would improve significantly, but a viewer’s opinion or attitude might change.

In this project, we assessed the effectiveness of Incident at Morales in achieving basic instructional goals in engineering ethics. A preliminary version of this paper was presented as a work-in-progress paper at the Thirty-Fifth ASEE/IEEE Frontiers in Education Conference [3].

II. METHOD

We selected two different assessment instruments: the Defining Issues Test and a five-item survey of statements about engineering practices. We chose these instruments because they are easy to administer, and because we were not convinced that the pedagogical effect of the video would be sufficiently great to justify a large assessment effort. Substantial amounts of labor and coordination are required for assessment methods based on scoring written essays [4, 5] or on individual interviews [6].

A. Five-Item Survey

We wrote a five-item survey, shown in Figure 1. We administered this survey to the following professionals and students:

- Twenty-eight students in CE 4292, mostly civil engineering majors, Texas Tech University, January 26, 2004.
- Thirty students, mostly chemical engineering majors, Texas Tech University, March 30, 2004.
Sixty-nine students in ME/IE 280, mostly mechanical engineering and industrial engineering majors, University of Illinois at Urbana-Champaign, March 30 and April 6, 2004.

Twenty-four students at a meeting of the IEEE Student Branch, mostly electrical engineering and computer engineering majors, University of Illinois at Urbana-Champaign, April 12, 2004.

We believe that these participants were broadly representative of professional engineers and engineering students in the United States. In particular, the students came from a diversity of ethnic, racial, and socioeconomic groups.

Participants took this survey before and after they watched the video. On the survey, each statement is motivated by a pedagogical goal of the video. Although statement 1 appears to be a dichotomous question [7], an assignment from an employer does not usually conflict with a contract from a client; in essence, statement 1 asks whether the engineer's first obligation is to fulfill an assigned task from either an employer or client. Because statement 1 seems reasonable, we expected that participants would initially tend to agree with it. We surmised that after watching the video, participants would be less willing to agree with statement 1: they would realize that although engineers should complete assigned tasks, engineers have more important responsibilities to the public.

Statement 2 appears to ask two separate questions: should an American engineer comply with local regulations in a foreign country, and should an engineer avoid more stringent American standards? In essence, statement 2 asks whether the engineer's first obligation is to fulfill an assigned task from either an employer or client. Because statement 1 seems reasonable, we expected that participants would initially tend to agree with it. We surmised that after watching the video, participants would be less willing to agree with statement 2: they would realize that although engineers should complete assigned tasks, engineers have more important responsibilities to the public.

We hypothesized that participants would initially disagree with statements 3, 4, and 5, because a priori, it is not obvious whether ethics has any relevance to the technical decisions that engineers make. We hoped that after watching the video, participants would acknowledge that ethics is relevant to technical decisions, and that ethical problems can have technical solutions. For example, suppose an engineer is asked to design a chemical process for a product that is nearly identical to a product that the engineer designed for a former employer. The engineer may feel caught between the moral obligation to serve the current employer and the moral obligation to keep trade secrets from a former employer confidential. To resolve this ethical conflict, the engineer could find a different technical solution that avoids revealing trade secrets.

When we pilot-tested the survey, we asked participants to explain the reasons for their responses. In the pilot version, statement 2 was phrased, “When working in a foreign country, an American engineer should comply with local regulations and should avoid imposing American standards for safety.” Some participants inferred that American standards were less stringent, and their responses were ambiguous.

B. Defining Issues Test

The Defining Issues Test (DIT) is a standard paper-and-pencil multiple-choice test of moral reasoning that was developed by the Center for the Study of Ethical Development at the University of Minnesota in the 1970s. The DIT has been validated through hundreds of studies with thousands of subjects [8]. The DIT presents six stories with moral dilemmas. Subjects answer questions about the most important factors that affect their judgments about the dilemmas. The score on the DIT, which ranges from 0 to 95, measures the developmental level of the subject’s reasoning according to the cognitive moral development theory of Kohlberg. More
The DIT-2 appears to be as reliable and valid as the original DIT, the DIT-2.1 Compared with the original DIT, the DIT-2 has only five stories, and these stories reflect contemporary issues.

The paired t-test, sign test, and signed rank test were performed for each comparison. Because the Shapiro–Wilk, Kolmogorov–Smirnov, Cramer-von Mises, and Anderson-Darling tests for normality rejected the normality assumption, the sign test and signed rank test were more appropriate for this data set.

Based on the sign test and the signed rank test, we drew the following conclusions for significance of differences (p < 0.05):

- For statement 1, there were significant differences between before video and after video, and before video and after discussion. The direction of the change was in the negative direction, that is, the engineers were more inclined to disagree with the statement after the video and after the discussion. The difference between after video and after discussion was significant. The direction of the change was significant based on the signed rank test, and was not significant based on the signed rank test. However, both of the p-values were close to 0.05.
- For statement 2, there was a significant difference between before video and after video. The direction of the change was in the negative direction, that is, the engineers were more inclined to disagree with the statement after the video. The difference between before video and after discussion was not significant.
- For statement 3, only the difference between before video and after discussion was significant. The direction of the change was in the negative direction, that is, the engineers were more inclined to disagree with the statement after the discussion.
- For statement 4, there were no significant differences among the responses of the three different surveys.
- For statement 5, there were no significant differences among the responses of the three different surveys.

B. Five-Item Survey: Students

The data for the four administrations with students were aggregated into a single data set of 162 students. The students had the characteristics shown in Table 2.

---

We administered the DIT-2 to 11 students at the University of Illinois at Urbana-Champaign and to 37 students at Texas Tech University. One week after they took the DIT-2, the students watched the Incident at Morales video; one week after the video, they took the DIT-2 again. Each administration of the DIT-2 ran 35 to 45 minutes. Because the DIT and DIT-2 are designed to measure improvements in moral reasoning skills over long periods of time, we did not expect significant increases in DIT-2 scores between the pre-test and the post-test.

### Table 1. Relationship between level in college and DIT-2 scores [11].

<table>
<thead>
<tr>
<th>Educational Level</th>
<th>Average P score</th>
<th>Average N2 score</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshman</td>
<td>32.32</td>
<td>31.05</td>
<td>2,096</td>
</tr>
<tr>
<td>Sophomore</td>
<td>32.62</td>
<td>31.24</td>
<td>1,028</td>
</tr>
<tr>
<td>Junior</td>
<td>34.45</td>
<td>32.65</td>
<td>1,333</td>
</tr>
<tr>
<td>Senior</td>
<td>37.84</td>
<td>36.85</td>
<td>2,441</td>
</tr>
</tbody>
</table>

---

The DIT-2 can be ordered from the Center for the Study of Ethical Development, University of Minnesota, 206A Burton Hall, 178 Pillsbury Drive SE, Minneapolis, MN 55455.
The paired t-test, sign test, and signed rank test were performed for each comparison of students’ survey responses before watching the video and after watching the video. Because the normality tests rejected the normality assumption, the sign test and signed rank test were more appropriate for this data set.

1) **All Students:** First, we tested for statistically significant differences \((p < 0.05)\) before and after watching the video using the whole data set (see Table 3).

- For statements 1, 2, 3, and 5, there were significant differences before and after watching the video. There was no significant difference for statement 4.
- For statements 1 and 2, the change in direction was negative, i.e., after watching the video, students were more inclined to disagree with the statements. For statements 3, 4, 5, the change in direction was positive, i.e., after watching the video, students were more inclined to agree with the statements. However, the change for statement 4 was not statistically significant.

2) **Gender Effects:** Second, we tested for statistically significant differences before and after watching the video for men and women (see Table 3).

- For statement 1, there were significant differences for both men and women. The changes in direction were negative for both men and women.
- For statement 2, there were significant differences for men but not for women. The change in direction was negative for men.
- For statement 3, there were significant differences for men but not for women. The change in direction was positive for men.
- For statement 4, there were significant differences for women but not for men. The change in direction was positive for women.
- For statement 5, there were significant differences for men. For women, the signed rank test shows a marginally significant result while the result of sign test was not significant. The changes in direction were positive for both men and women.

3) **Level of Study:** Third, we tested for statistically significant differences before and after watching the video for students at different years in college. Years 1 and 2, and years 5 and 6 were combined into groups because there were few observations in years 1 and 6. Therefore, four levels were considered: years 1 and 2, year 3, year 4, years 5 and 6 (see Table 4).

- For statement 1, there were significant differences for the first three levels (years 1 and 2, year 3, year 4). The changes in direction were negative for all levels, but the difference for years 5 and 6 was not significant. Perhaps older students had already internalized the importance of professional responsibilities beyond assigned tasks. Or perhaps the difference was not significant because of the small number of students in years 5 and 6.
- For statement 2, there were significant differences for the last two levels (year 4, year 5 and 6). The changes in direction were negative for all levels, but the differences for the first two levels (years 1 and 2, year 3) were not significant.

---

**Table 2. Characteristics of students who took the five-item survey.**

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>124</td>
</tr>
<tr>
<td>Women</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year in college</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>82</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>6+</td>
<td>4</td>
</tr>
<tr>
<td>Unknown</td>
<td>14</td>
</tr>
</tbody>
</table>

---

**Table 3. Changes from before watching the video to after watching the video. Asterisks denote statistically significant changes.**

<table>
<thead>
<tr>
<th></th>
<th>All students ((n = 162))</th>
<th>Men ((n = 124))</th>
<th>Women ((n = 25))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statement 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First obligation to employer or client</td>
<td>* Negative</td>
<td>* Negative</td>
<td>* Negative</td>
</tr>
<tr>
<td>Sign test (p &lt; 0.001)</td>
<td>Sign test (p &lt; 0.001)</td>
<td>Sign test (p &lt; 0.0127)</td>
<td></td>
</tr>
<tr>
<td>Signed rank (p &lt; 0.001)</td>
<td>Signed rank (p &lt; 0.001)</td>
<td>Signed rank (p &lt; 0.0084)</td>
<td></td>
</tr>
<tr>
<td><strong>Statement 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoid imposing American standards</td>
<td>* Negative</td>
<td>* Negative</td>
<td></td>
</tr>
<tr>
<td>Sign test (p &lt; 0.001)</td>
<td>Sign test (p &lt; 0.001)</td>
<td>Sign test (p &lt; 0.7539)</td>
<td></td>
</tr>
<tr>
<td>Signed rank (p &lt; 0.001)</td>
<td>Signed rank (p &lt; 0.001)</td>
<td>Signed rank (p &lt; 0.7930)</td>
<td></td>
</tr>
<tr>
<td><strong>Statement 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethics is integral in decisions</td>
<td>* Positive</td>
<td>* Positive</td>
<td></td>
</tr>
<tr>
<td>Sign test (p &lt; 0.0022)</td>
<td>Sign test (p &lt; 0.0060)</td>
<td>Sign test (p &lt; 0.2500)</td>
<td></td>
</tr>
<tr>
<td>Signed rank (p &lt; 0.0023)</td>
<td>Signed rank (p &lt; 0.0076)</td>
<td>Signed rank (p &lt; 0.2500)</td>
<td></td>
</tr>
<tr>
<td><strong>Statement 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code of ethics is helpful</td>
<td></td>
<td></td>
<td>* Positive</td>
</tr>
<tr>
<td>Sign test (p &lt; 0.0807)</td>
<td>Sign test (p &lt; 0.3915)</td>
<td>Sign test (p &lt; 0.0313)</td>
<td></td>
</tr>
<tr>
<td>Signed rank (p &lt; 0.0567)</td>
<td>Signed rank (p &lt; 0.3105)</td>
<td>Signed rank (p &lt; 0.0313)</td>
<td></td>
</tr>
<tr>
<td><strong>Statement 5</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethical problems have technical solutions</td>
<td>* Positive</td>
<td>* Positive</td>
<td>* Positive</td>
</tr>
<tr>
<td>Sign test (p &lt; 0.0049)</td>
<td>Sign test (p &lt; 0.0145)</td>
<td>Sign test (p &lt; 0.1460)</td>
<td></td>
</tr>
<tr>
<td>Signed rank (p &lt; 0.0051)</td>
<td>Signed rank (p &lt; 0.0282)</td>
<td>Signed rank (p &lt; 0.0474)</td>
<td></td>
</tr>
</tbody>
</table>
For statement 3, there were significant differences for year 4. The changes in direction were positive for all levels, but the differences for all levels except year 4 were not significant.

For statement 4, the changes in direction were positive for all levels, but the differences were not statistically significant.

For statement 5, there were significant differences for the last two levels (year 4, years 5 and 6) in the positive direction. The differences for years 1 and 2, and for year 3 were not significant.

C. Defining Issues Test
A total of 48 students took the DIT-2 test before they watched the *Incident at Morales* video (pre-test) and again after they watched the video (post-test). Four of these students were purged from the data set by the DIT-2 reliability checks. Of the remaining 44 students, 33 were men, 10 were women, and one did not provide a gender. Most students (30 of 44) were college seniors majoring in engineering. For some students the post-test P and N2 scores were lower than the pre-test scores, but for most students the post-test scores were higher. The average P scores and N2 scores before and after watching the video are shown Table 5.

The Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling tests accepted the normality assumption for the P scores and the N2 scores, and the paired t-test was used to determine the significance of the changes between the pre-test and post-test scores. Of the changes, only the increase in N2 score for men was statistically significant ($p < 0.0347$).

Finally, we tested whether women experience significantly greater increases in their P scores and N2 scores than do men. We fit an Analysis of Variance (ANOVA) model to the data. We did not find a significant difference between the changes for men and for women on either the P score ($p > 0.4383$) or the N2 score ($p = 0.1110$).

IV. DISCUSSION
Because the DIT-2 is designed to measure changes in moral reasoning skills over long periods of time, we expected that the pre-test scores would approximately equal the post-test scores. Contrary to our expectations, the DIT-2 scores changed significantly from the pre-test to the post-test: when all students were considered together, the post-test N2 score was significantly higher than the pre-test N2 score. If the N2 score measures the developmental level of the subject’s moral reasoning, then a single viewing of the video produces a significant increase in the level of moral reasoning. We cannot say whether the increase was caused by...
by the video alone or by the combination of the video and the classroom discussion, however.

Like Self and Ellison [9], we did not find major differences between men and women in the pre-test to post-test changes in DIT-2 scores, but the number of women in our sample may have been too small to find statistically significant differences. If more students had taken the DIT-2, then we might have found other significant differences; for example, the increases in P scores from pre-test to post-test could be statistically significant.

From the results of the study with professional engineers, it appears that watching the video in itself, without the discussion, is sufficient to produce a statistically significant change of opinion on statements 1 and 2 of the five-item survey. Both the professional engineers and the students were less inclined to agree with statements 1 and 2 after watching the video. The results for statements 3, 4, and 5 were less dramatic. When we examine the content of *Incident at Morales*, we can discern possible reasons for the differences in viewers’ responses to statements 1 and 2, on the one hand, and statements 3, 4, and 5, on the other. The video clearly addresses statements 1 and 2. In the video, despite pressure from his employer to reduce costs, the principal engineer decides to spend more to protect the safety of the local residents: although local environmental regulations call for unlined evaporation ponds, the engineer decides to line the ponds to prevent poisonous byproducts from seeping into the groundwater. In contrast, the video does not directly address statements 3, 4, and 5. In one scene, the main character briefly mentions that he consulted some ethics manuals, but in general, the characters in the video do not explicitly mention ethical considerations in making decisions.

After watching the video, the professional engineers were less inclined to agree with statement 3, but the students were more inclined to agree with this statement. We did not collect sufficient data from the assessment sessions to explain this difference between the professional engineers and the students.

The results of the Defining Issues Test and the five-item survey cannot be compared directly. These two instruments were designed for different purposes: the DIT-2 measures the developmental level of the subject’s moral reasoning, and the survey gathers the subject’s opinions about engineering practices. The two instruments were administered with different groups of subjects.

This study indicates that *Incident at Morales* might be effective in improving viewers’ moral reasoning skills and in changing viewers’ opinions about some specific aspects of engineering ethics in the short term. We cannot tell whether these changes are stable over the long term, however. Although a long-term benefit would be desirable—especially to document educational outcomes for engineering accreditation under EC2000—it is difficult to assess the long-term effect of a single viewing of the video for two reasons. First, many other factors can influence an individual’s moral development. Second, students naturally experience intellectual and moral growth throughout college [13]. Further study will be needed to establish whether a single brief intervention has a lasting impact.

ACKNOWLEDGMENTS

It was an honor and a pleasure to develop the *Incident at Morales* video with Patricia Harper, E. Walter LeFevre, Steven P. Nichols, Carl M. Skooglund, Jimmy H. Smith, Frederick Suppe, Philip E. Ulmer, and Vivian Weil. My colleagues J. Craig Dutton and Jimmy Smith administered the DIT-2 and the five-item survey at various showings of the video. Undergraduates Steven Cortes and Golnaz Hashemian entered survey data into spreadsheets. Golnaz Hashemian and William Lawson suggested improvements to drafts of this paper. The statistical analyses were performed by Jianhui Zhou and Di Li, graduate students in the Department of Statistics at the University of Illinois at Urbana-Champaign.

This project was supported by the National Science Foundation under Grant SES-0138309. I take full responsibility for the views, opinions, and conclusions of this paper, which are not necessarily those of the National Institute for Engineering Ethics, the National Science Foundation, or the University of Illinois.
REFERENCES


AUTHOR’S BIOGRAPHY

Michael C. Loui earned the B.S. degree at Yale University in 1975, and the S.M. and Ph.D. degrees at the Massachusetts Institute of Technology in 1977 and 1980.

Since 1981, he has taught at the University of Illinois at Urbana-Champaign, where he is a professor of electrical and computer engineering and a research professor in the Coordinated Science Laboratory. His interests include computational complexity theory, professional ethics, and the scholarship of teaching and learning.

From 1990 to 1991, Professor Loui directed the theory of computing program at the National Science Foundation in Washington, D.C. From 1996 to 2000, he was an associate dean of the Graduate College at Illinois, with administrative responsibility for all graduate academic programs on campus. He currently serves on the editorial boards of *Information and Computation, Teaching Ethics, College Teaching*, and *Accountability in Research*. He is a member of the Advisory Board for the Online Ethics Center for Engineering and Science, the Executive Board of the National Institute for Engineering Ethics, and the Board of Governors of the IEEE Society on Social Implications of Technology.

In 1985, Professor Loui won the Dow Outstanding Young Faculty Award of the American Society for Engineering Education. At Illinois, in 1995, he received the campus’s Luckman Undergraduate Distinguished Teaching Award, and in 2001, he was named a University Distinguished Teacher/Scholar. In 2003, he was selected as a Carnegie Scholar by the Carnegie Foundation for the Advancement of Teaching.

Address: Coordinated Science Laboratory, 1308 W. Main St., Urbana, IL 61801; telephone: (+1) 217.333.2595; e-mail: loui@uiuc.edu.