Interactive Learning Environments
2002, Vol. 10, No. 3, pp. 243–262

1049-4820/02/1003-243$16.00

Learning, Media, and the Case of Dax Cowart: A Comparison of Text, Film, and Interactive Multimedia

Robert Cavalier[1] and Keith Weber[2]
[1]Department of Philosophy, Center for the Advancement of Applied Ethics, Carnegie Mellon University, USA, and [2] Center for Innovation in Learning, Carnegie Mellon University, Pittsburgh, PA, USA

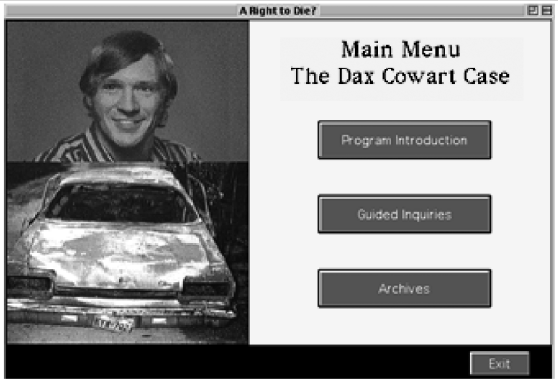
ABSTRACT

This study reports the results of a 3-year investigation into the comparative advantages of using interactive multimedia in the presentation of a case study in Ethics. The study involved undergraduate students in philosophy courses at Carnegie Mellon and the use of the case of Dax Cowart, a burn patient who wished to be allowed to die. In representing the case, we strove for functional equivalence in the content of the media chosen: a literary narrative plus expert commentary (text), a 1-hr documentary (film), and the Guided Inquiry/Archive sections of an interactive multimedia program (*A Right to Die? The Dax Cowart Case*). We also strove for functional equivalence in the use of the material: an evening assignment and a follow-up essay exam. The results of our research demonstrate a statistical difference in learning outcomes based on the medium used. Students in the interactive CD group outperformed students in the text and film groups with regard to (a) understanding the complex perspectives and positions of the principals in the case and (b) analyzing the case with respect to its morally relevant details.

INTRODUCTION

The term 'multimedia' is ubiquitous in today's computing environment, so much so that it may be as redundant to speak of 'multimedia computing' as to speak of 'Graphical User Interface Computing'. Computers naturally combine diverse forms of media such as text, graphics, sound and digital video. But the use of interactive multimedia in education is still relatively new and its value in the learning process is a matter of lively discussion (Cavalier 1996; Clark 1994a; Ehrmann 1996). This article analyses a CD-ROM program utilizing

Correspondence: Robert Cavalier, PhD, 155 Baker Hall, Carnegie Mellon University, Pittsburgh, PA 15213, USA. Tel.: +1-412-268-7643, E-mail: rc2z@andrew.cmu.edu



Main Menu selections. The case can be accessed through either a Guided Inquiry Section or a more open-ended Archives Section.

interactive multimedia, *A Right to Die? The Dax Cowart Case* (Andersen, Cavalier, & Covey, 1996), and assesses one aspect of its use in the classroom setting.

In evaluating any program, it is important to place it in an appropriate context. It would be a conceptual error to try to evaluate 'multimedia' abstracted from a particular disciplinary sphere, for it is the latter that provides the paradigm sets of problems that the particular program seeks to address. There is no Platonic Form for Multimedia that can be initially grasped independently of a particular research agenda or classroom setting. Our area of interest is in the field of Moral Philosophy and Applied Ethics. And our topical area could be classified as Medical Ethics.

A Right to Die? The Dax Cowart Case[1] investigates a burn patient's request to be allowed to die. The case relates the story of a 25-year-old man who received second and third degree burns over two thirds of his body. At the time that we encounter Dax in treatment, his injuries have left him severely scarred, his hands are badly deformed, and the sight in his one remaining eye is at risk. As a patient, he undergoes daily treatments in an antiseptic tank. These

[1]An early videodisc version of this program won the 1988 EDUCOM Best Humanities Software Award.

'tankings' are extremely painful to him and will continue over a period of months. In light of this, he has persistently asked the doctors to stop his treatment. Dax feels that the length of his projected treatment, and the quality of life that he can expect to regain, do not warrant the torment he must suffer. The doctors know that if they continue his treatment, he will live; if they stop his treatment, he will surely die. A basic question posed at this point of impasse between Dax and his care-givers is: Does Dax have a right to die? And, if so, what does this mean?

Traditional approaches to this kind of case study often rely on a four or five paragraph case study. The study is treated as an example to be discussed within the framework of prior readings. It is used to test the analytical skills of the student and to explore the relation between the general and the particular. But one pressing problem with a 'case summary' is the discrepancy between its description and its reality:

> Trying out one's theory on real situations, thick with details, is very different from the philosopher's typical hypothetical case, which, if not simply invented, is so highly abstracted from real circumstances that only enough details remain to defend selectively the particular point the philosopher wants to make thereby. His or her use of cases is much more to *illustrate* theory than to test it. But when *solving* the moral problem is the main point, the relentlessness of the details becomes readily apparent. There is no refuge; there is one quagmire after another; retreating to the theory is not a viable option. (Clouser, 1993)

The palpable complexity of real life situations seems to recede in the text book overview of such cases – hence the attempt to introduce reality into the context of case studies through the thoughtful use of interactive multimedia.

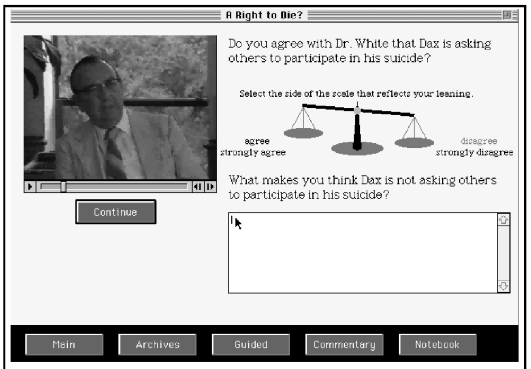
There are several theoretical backgrounds for the development and use of this interactive CD, the most general of which is described under the rubric of Project THEORIA. The acronym 'THEORIA' stands for 'testing hypotheses in ethics/esthetics: observation, reason, imagination, and affect'. The term is intended to play off of the origins of both 'theory' and 'theater' in the ancient Greek verb *theorein:* to see, to view, to behold. One stimulus for the project was Derek Bok's early claim that while computers may be useful in areas of data and calculation, they are ill-equipped to handle the often open-ended areas of the humanities (Cavalier & Covey, 1996; Covey, 1993). But if you

Menu selections for the Guided Inquiry section. This interpretation of the case focuses on the conflict between the medical professional obligations and issues relating to patient autonomy.

shift the paradigm of 'computing' away from command screen terminals to systems that display all kinds of information, including video's rich information, computers might indeed address some important issues like those raised in the case of Dax Cowart. The reason for this lies in the ability of *multimedia* to convey 'rich data' in the form of, for example, personal narratives and realistic settings. Furthermore, the *interactive* aspect of computer-mediated multimedia utilized in this program could be used to engage the student's reasoning process, eliciting an *active* and *reflective engagement* in the case.

The interactive CD provides a flexible navigation through the central, and often competing, issues of the case. In a Guided Inquiry section, users take a semi-Socratic tour of the case from the bare facts to a confrontation with doctors, patient and other principals over the issues of Medical Professionals' Obligations, Pain of Treatment, Quality of Life, and Patient's Rights and Capacities. At the end, a decision must be made as to whether Dax should have his request or not. Final sequences are determined on the basis of this recommendation.

An Archive section contains video sections on the principals and main issues as well as descriptions of other cases (e.g., the precedent-setting case of Karen Anne Quinlan). Recent State and Federal court decisions on issues like physician-assisted suicide are also included.



After listening to a doctor's perspective on one aspect of the case, the user is queried as to his or her opinion and reasons for agreeing or disagreeing with the stated position.

To test the usefulness of this approach to the case – and its effect on the learning process – a series of formative and summative evaluation studies have been implemented. Beginning in the Spring of 1997, a study of the interactive CD was formally incorporated into an 'Introduction to Ethics' class and, later, into a general 'Introduction to Philosophy' class. The assessment of the program progressed through four stages. The first stage served as a pilot study. This was followed by three further studies. Each new phase built upon the lessons learned in the previous study. We concluded the classroom tests in the Fall of 1999 and report our findings below.

PILOT STUDY (SPRING 1997)

The pilot study utilized an introductory ethics course taught by the first author to assess the effects of interactive multimedia versus film and text in developing a student's understanding of the case material. Students in the class were divided into three groups of 6 or 7 subjects, the make-up of each group being determined by the combination of Verbal SAT scores and the grade for the first exam. The design of the study required that we divide the groups into Text, Film and Interactive CD categories.

Group A read a 10-page textual narrative of the case, a book chapter on Dax's case, and an article.[2] Group B viewed a 1-hr documentary film on the case.[3] And Group C went through parts of the Guided Inquiry and Archives sections of the interactive CD.[4] We strove for informational equivalence in the content of the media chosen: a literary narrative plus expert commentary (text), a 1-hr documentary (film), and the Guided Inquiry/Archive sections of the multimedia program (interactive CD). A special evening session allowed each student to spend approximately 1 hr on task. In the pilot study, students were allowed to take notes and to bring those notes to the next day's class session. In the class session, the students took a 50-min essay exam. The questions focused on knowledge of the facts, the principal participants in the case, and the central issues of the case. This exam was very similar to the exam that we used in subsequent studies. The core questions, the answers to which were evaluated throughout the studies, involved three parts. The first part simply said: 'Present the facts of the case'. The second part asked the student to 'describe the positions of each of the principals of the case' and then presented the following list: Dr. Robert White (Dax's psychiatrist), Dr. Duane Larson (burn specialist taking care of Dax at Parkland), Dr. Baxter (initial primary care physician), Dax Cowart (patient), Ada Cowart (Dax's mother), Leslie Kerr (Dax's nurse), and Rex Houston (Dax's lawyer). The third part asked each student how he or she would describe ''the nature and relevance of the following issues in Dax's case'' and then listed the following: Physician Obligations, Pain of Treatment, Quality of Life, Patient Rights and Capacities. A fourth part of the initial exams, dealing with patient advocacy and hospital policy, proved too time consuming, difficult to quantify, and outside the scope of the assignments. It was later dropped and the study focused on the first three parts.

A pair of graders familiar with the case study assigned each student essay an integer score between 1 and 4. The graders scored the essays holistically; that is, they were given no instructions on how to grade the essays. Perhaps

[2] The texts were carefully selected to accommodate the dramatic aspects as well as the factual aspects of the case. The narrative comes from the Teachers's Guide for *A Right to Die? The Dax Cowart Case* (Cavalier, & Covey 1996) and the chapter ('A Memoir: Dax's Case Twelve Years Later', by Robert White) comes from Lonnie Kliever, Ed., *Dax's Case: Essays in Medical Ethics and Human Meaning* (Kliever, 1989).

[3] The film, *Dax's Case* (Concern for Dying, 1985), uses both original footage from Dax's treatment as well as interviews with principals conducted 10 years later.

[4] Much of the video footage used in the CD came from the film on Dax's case. *A Right to Die? The Dax Cowart Case* (Andersen, Cavalier, & Covey, 1996).

because no scoring criteria were given to the graders, the interrater reliability was quite poor. The graders agreed on the scoring of less than 40% of the essays. This lack of agreement motivated the more precise scoring criteria used in the subsequent studies. Each student's score on the Dax Cowart exam was obtained by totaling the mean score of the two graders for each essay.

The results of the pilot study were encouraging. Students learning about the case study using the interactive CD had a higher mean score than both students learning about the case via text and students learning about the case via film (33.47 for the interactive CD; 33.18 for the film; and 31.46 for the text). Students in the interactive CD group performed at least as well as the other students despite the fact that students taking notes on the linear presentation of the text and film were at an advantage over students who took the Guided Inquiry of the interactive CD. For example, the text input/output from the Guided Inquires was nested around the particular questions posed by the program. These questions stressed critical reasoning, but did not take the form of an outline that could be referred to with standard note-taking procedures. The lack of functional equivalence here led us to drop the use of notes in the test taking for the final two studies. And this, in fact, more appropriately models the typical class assignment (viz., 'go over this material and be prepared to discuss it tomorrow').

The pilot study had several drawbacks, however. The two graders who scored the essays did not have a high level of agreement. Also, the variance of scores within each group was quite high and thus the magnitude of the results we obtained did not approach statistical significance.

In addition to the quantitative pilot study, we ran an informal qualitative study in the Spring 1997 Introduction to Philosophy class ('What Philosophy Is'). Of the 160 students enrolled, 90 used the interactive CD. A set of questions was distributed, and slightly over 50% of those who used the interactive CD responded. On a scale of 1 (strongly disagree) to 3 (neutral) to 5 (strongly agree), the student responses were as follows:

> Question 1: "I enjoyed working with the CD-ROM." [1 (**1**); 2 (**8**); 3 (**15**); 4 (**15**); 5 (**8**)] Mean: 3.45 (out of 5) Students later commented that while the interactive CD was very powerful and compelling, 'enjoyment' did not accurately describe their experience.
>
> Question 2: "I gained a unique perspective on the case because it was presented to me through a multimedia environment." [1 (**4**); 2 (**5**); 3 (**10**); 4 (**13**); 5 (**17**)] Mean: 3.69.

Question 3: ''The CD-ROM presented the case in an unbiased way.'' [1 (**3**); 2 (**7**); 3 (**9**); 4 (**18**); 5 (**9**)] Mean: 3.50. Since the program asks questions and branches in such a way as to put pressure on the student's responses, we suspect that this accounts for some students feeling that the program pushed them in a certain direction.

Question 4: ''Using the CD-ROM helped me learn the material.'' [1 (**0**); 2 (**6**); 3 (**6**); 4 (**26**); 5 (**8**)] Mean: 3.78. This is a key question and its results, even though qualified by the assumption that these students are somewhat self-selecting, should be emphasized.

Question 5: ''Where appropriate, more of this kind of multimedia should be used in courses.'' [1 (**1**); 2 (**2**); 3 (**10**); 4 (**21**); 5 (**11**)] Mean: 3.87.

In follow-up interviews with the class instructors, all four (one faculty and three TAs) felt that the interactive CD was helpful and that each would use it again.

Although there were no statistically reliable results, the group of students using the interactive CD performed at least as well as the groups of students using the text and film and both student and teacher reaction to the use of the interactive CD was positive. We were encouraged by these facts. The lessons learned from the pilot study formed the basis for the study in the Spring of 1998. In this and the final two studies, we describe the methods and results for each assessment. We begin each with an introduction and conclude with a discussion.

STUDY 1 (SPRING 1998)

A number of changes in the design and implementation of the study were made. First, additional questions were added and the scoring criteria for each question were refined. Two graders were hired and trained in the use of the grading criteria for the study.

Also, the time for the study was extended to 1.5 hours. Because of this extended time, students in the Text Group were given more material,[5] students in the Film Group were given the opportunity to rewind and pause the film

[5] ''Confronting Death: Who Chooses, Who Controls?,'' a dialogue between Dax Cowart and Robert Burt published in the *Hastings Center Report* 28, no 1 (1998): 14–24.

(each had his or her own VCR), and students in the interactive CD Group were requested to use the Archive Section of the interactive CD after completing the Guided Inquiry. As with the previous study, all students were allowed to take notes and bring them to the test the following day.

In order to improve interrater reliability, we constructed strict scoring criteria for each question. For every question, we constructed a list of relevant statements that could be associated with that question. We gave students credit for a given statement if they wrote something equivalent to that statement or something that directly and obviously implied that statement. Each student's answer to that question received a score between 1 and 4. The student in the course who performed the best on this question received a score of 4; every other student received a score between 1 and 4 proportional to the number of questions they answered correctly. The following is an example of the scoring criteria that we used for Dr. White (Dax's psychiatrist):

Felt Dax was rational (mentally competent, fit to make decision) ___
Was against letting Dax die ___
Felt in essence that stopping treatment was the same as killing him (also, Dax was asking others to help him commit suicide) ___
Felt Dax's feelings of resentment were natural ___
  Described natural feelings as 'little boy feelings' ___
Felt Dax should postpone decision until he could do it himself ___
Felt Dax could have a reasonable life ___
  Felt Dax should accept a chance of happiness ___
Felt Dax should continue with more painkillers ___
Felt Dax could be helped with psychological counseling ___

Our criteria were designed to meet the following goals:

- The criteria would be explicit enough that a grader, with training, could score an essay without assistance from the investigators.
- The criteria for grading would be sufficiently objective for different graders to achieve a high level of interrater reliability.
- The statements that were awarded points were given roughly equal relative importance in the interactive CD, the texts, and the film.
- Essays displaying a greater understanding of the issues of the case would receive a higher number of points.
- This scoring system is unambiguous enough that one could attempt to replicate our results using the same system.

After the criteria were constructed, the graders, who also helped create the guidelines, were trained as follows. They practiced grading exams from the previous year. Each interrater disagreement was looked at and discussed. In some cases, the grading criterion was revised to make it less ambiguous.

To address the large within-group variability of the previous study, an in-class short essay exam similar in form (but different in content) to our assessment was used to account for individual differences between the students. Since the exam was similar in form to the assessment used in our study, it is likely that the exam is a measure of students' ability to perform on these types of examinations.

Method

Participants

There were 38 students enrolled in an introductory ethics course, taught by the first author, who participated in this study. Participation in this study fulfilled a course requirement.

Design

The design of this study was a one-factor three-level experiment. The experiment employed three groups: a text group, a film group, and an interactive CD group. Students were assigned to groups by matching verbal SATs and their score on an in-class essay exam given before our study commenced.

Procedure

The dependent measure of this study was the students' performance on an in-class short essay exam that was designed to probe the students' knowledge of the facts of the Dax Cowart case study, the perspectives of the key individuals associated with the case, and the morally relevant issues associated with the case. This is the same exam that was given to the students in the pilot study.

The night before the exam was given, the students were asked to learn about the Dax Cowart case study using either the interactive CD, text, or film medium. In this study and all subsequent studies, the first author observed all groups studying the material. In general, the students appeared to be working diligently on their task. They were given 1.5 hours to learn the material. All the students were allowed to take notes and bring them to the test the following day. They were given up to 1 hour to complete the exam.

Measurement

Two graders, one being the second author, graded each question as follows. For each question, the grader was given a list of facts. If in the grader's judgment the participant's essay contained a statement equivalent to that fact or a statement that directly implied that fact, the student was given credit for mentioning that fact. The participants' score for each question was a number between 1 and 4, and was determined by the criteria previously discussed. The graders were unaware of the participants' experimental condition. In the event of interrater disagreement, the student's score on the Dax Cowart exam was the mean of the two grader scores. (The first question, which asked students to state the important facts of the case, was graded in the same manner, except it was graded out of 5 points. Because two of the figures were represented sparsely in all of the media, the answers concerning the nurse Leslie Kerr and the lawyer Rex Houston were combined and graded jointly on a 4 point scale).

Results

One participant from the film group was removed from analysis due to an extremely low score that was considered to be an outlier. Since we were attempting to demonstrate that students in the interactive CD group outperformed their counterparts in the text and film groups, including this outlier student in our analyses would only strengthen our conclusions. The interrater reliability was .94.

Participants' scores ranged from 20 to 38.5 (possible scores ranged from 12 to 49). Participants in the interactive CD group (mean score = 30.3) outperformed students in the text group (mean score = 28.8) and the film group (mean score = 27.2). An ANOVA revealed that this main effect was statistically reliable, $F(2, 36) = 3.82$, $p < .01$. Individual contrasts revealed that participants in the interactive CD group performed significantly better than participants in the film group, $t(23) = 3.05$, $p < .05$, but did not reveal that students in the interactive CD group performed significantly better than students in the text group, $|t| < 1$.

In our study, we felt that a large portion of the within-group variance may have been due to differing student ability. We wanted to account for this variance by factoring out individual student differences. Thus, we attempted to factor out effects based on students' verbal abilities (as measured by their verbal SATs) and by their test-taking ability. The latter was measured by the student's score on the first exam of the course (which was similar in form, but not in content, to the Dax Cowart exam). A linear regression was run to determine if there were significant

effects on the Dax Cowart exam score based on Verbal SATs (200–800), the student's score on the first exam of the course (treated as a 12-point test), and which group the student was in (interactive CD, film, or text).

The following table presents the results of the regression:

ANOVA Table

	DF	Sum of Squares	Mean square	*F*-value	*p*-value
Regression	4	222	55.6	6.150	.0018
Residual	22	199	9.0		
Total	26	421			
$R^2 = .528$	$R^2_{\text{adjusted}} = .442$				

Variable	Coefficient	*t*-value	*p*-value
Regression coefficients			
β_o (Intercept)	27.596	4.156	$<.0001$
β_1 (Exam score)	0.986	3.734	.0012
β_2 (Verbal SATs)	−0.004	−0.584	.5652
β_3 (text)	−5.844	−4.000	.0006
β_4 (film)	−4.858	−3.033	.0012

The model accounted for a significant portion of the variance in our data. Three individual contrasts were performed to see if β_3 β_4, and $\beta_3 - \beta_4$ were significantly different from zero. Using the Bonferonni adjustment and a significance level of .05, the critical *p*-value for comparisons was $p = .0133$.

The contrasts indicate that β_3 is significantly different from zero, indicating that students in the interactive CD group did significantly better than students in the text group. Further, the contrasts indicate that β_4 is significantly different from zero, indicating that students in the interactive CD group did significantly better than students in the film group. $\beta_3 - \beta_4$ was not significantly different from zero ($|t| < 1$), indicating that there was not a significant difference in performance between students in the text and the film group.

Discussion

This study provides strong and statistically significant evidence for the effectiveness of the Dax Cowart interactive CD. Analysis shows that students in the interactive CD group significantly outperform students in the text group and the film group in our dependent measure, the Dax Cowart score.

However, this study had several drawbacks. First, although the graders were blind to which group each student was in, one of the graders was the second

author, leading to the possibility of 'experimenter bias'. Second, the graders themselves worked jointly to construct the scoring criteria. Although the graders' interrater reliability of .94 was quite high, this might not be the case with a different set of graders.

STUDY 2 (SPRING 1999)

The purpose of this study was to replicate the results of Study 1 with a different pair of graders unaffiliated with this project. Also, the grading criteria were revised with the goal of achieving higher construct validity. In particular, (a) the list of statements associated with each question was expanded, and (b) for each question every statement received a weight proportional to its importance – that is, stating a more important detail received more credit than stating a less important detail. The following section from the grader sheet on 'Perspectives of Important Figures in the Dax Cowart Case' provides a sample of the criteria that we used:

DR. WHITE

Felt Dax was sane (lucid, mentally competent, and the like will do) (+5)
Was against letting Dax die (+3)
Felt in essence that stopping treatment was the same as killing him (also, Dax was asking others to help him commit suicide) (+1)
Felt Dax's feelings of resentment/helplessness were natural (+1)
  Described Dax's natural feelings as 'little boy feelings' (+1)
Felt Dax should postpone decision until he could do it himself (+1)
Felt Dax could have a reasonable life (+1)
  Felt Dax should accept a chance of happiness (+1)
Felt Dax should continue with more painkillers (+1)
Felt Dax could be helped with psychological counseling (+1)

Methods

Participants

There were 28 students enrolled in an introductory ethics course, taught by the first author, who participated in this study. Participation in this study fulfilled a course requirement.

Design
The design was identical to the previous study.

Procedure
The procedure was identical to the previous study.

Measurement
A pair of graders, otherwise unaffiliated with the study, scored each exam. The graders were not told the purpose of the study and the graders were blind to the design of the study. The graders scored the exam using criteria similar to the previous study. A student's score was a number between 0 and 4, proportional to how much credit he received for his responses by the graders, such that the student who received the most credit earned a '4'. (Again, the first question was an exception to this criterion and was graded out of 5 points.)

Results
One student was dropped from the analysis because he took the Dax Cowart exam a full week after working on the Dax Cowart interactive CD. Six other students were not included in the analysis since they either did not have scores for their homework assignments or their Verbal SATs were not available. Thus, 21 students were included in analysis.

The interrater reliability was .84, which represents a high level of overall interrater reliability.

Participants' scores ranged from 12.0 to 29.9 (the possible range of scores was 0–49). We replicated the previous studies' results in that the students in the interactive CD group (mean score = 22.6) outperformed students in the text group (mean score = 21.1) and students in the film group (mean score = 18.6).

We performed a linear regression using the students' Dax Cowart exam score (the dependent variable in the preceding analysis) as the dependent variable and the students' average homework assignment score, the students' verbal SATs, and the medium used by the student as the independent measures. The regression did not predict a statistically significant amount of the variance ($R^2 = .164$, $F < 1$). A priori t-tests did not show that a student could be expected to perform significantly higher being in the interactive CD group than the text group ($|t| < 1$) or the film group ($t(16) - 1.099$, $p > .2$).

Discussion
Although we failed to achieve statistically significant effects in this study, it must be stressed that these results still support our hypothesis. Includ-

ing the pilot study, this was the third consecutive study in which the students learning about the Dax Cowart case using the interactive CD outperformed students learning about the case via text or film. The probability of such a result occurring by chance is only 1 in 27, or less than 5%.

There were several factors that may have contributed to the failure to achieve a statistically significant result in Study 2. First, this study had very few participants in each group. Second, the take-home exam score was not an effective predictor of the score on the Dax Cowart assessment (it accounted for less than 10% of the variance). This exam score made it difficult to factor out the within-group variance. In Study 1, the in-class essay exam was a much better predictor of the score on the Dax Cowart assessment (it accounted for 17% of the variance). We believe that this difference arose because that students had a 'take home' exam rather than an 'in-class' written exam. The latter proved to be a more effective predictor of student performance than the former. Finally, some students appeared to spend much of their time answering the first question of the exam. The inefficient use of time by some students may have added unnecessary random variance to the data.

STUDY 3 (FALL 1999)

We conducted a third study with the goal of producing statistically reliable evidence that students learning about the Dax Cowart case study using the Dax Cowart interactive CD can outperform students learning about the case study using text or film. To avoid the previous problems of a small sample size, we conducted this study using students in a large, general introductory philosophy course.

Methods

Participants

There were 68 students enrolled in a general, introductory philosophy course who participated in this study. The instructor of the course was unaffiliated with this study. Students in this course were given the option of participating in this study or writing a paper. Students' performance on this study did not affect the grade they received in the course.

Design

The design was identical to that of previous studies, except that students were placed into groups by their last name. Students whose last names fell into the first third of the alphabet were placed into the text group. Students whose last names fell into the middle third of the alphabet were placed into the film group. Students whose last names fell into the last third of the alphabet were placed into the interactive CD group. We assumed that this placement would provide a relatively random distribution of students. Analysis reported in the 'Results' section demonstrates that each group of students was performing equivalently in the course.

Procedure

The procedure was identical to the previous studies except participants were not permitted to take notes while learning about the Dax Cowart case study, the last question (Part IV) was dropped from the exam, and the exam was structured so as to make obvious to the participants how much time should be spent on each question. The latter was achieved by creating appropriately sized blank spaces between the questions and indicating to students how much time to spend on each question.

Measurement

The scoring of the exams was identical to the measurement used in the previous study.

Results

The interrater reliability was .71, which represents a suitable degree of interrater reliability.

Participants' scores ranged from 2.0 to 28.6 (possible scores ranged from 0 to 45). We replicated the previous studies' results that students in the interactive CD group (mean score = 17.7) had a higher mean score than the film group (mean score = 12.2) or the text group (mean score = 14.6). An ANOVA reveals that this effect is statistically reliable ($F(2, 65) = 6.83$, $p < .01$). Further, individual contrasts were performed using Fisher's PLSD which revealed statistically significant effects for individual contrasts between the interactive CD group and the text group and between the interactive CD and the film group with the significance level being .05.

We also compared each group's performance on an in-class essay exam given prior to this experiment. An ANOVA failed to yield a statistically significant

Mean group score by question in Study 3			
Medium	CD	Film	Text
Question 1	3.0	3.0	3.3
Question 2	8.8	6.2	7.6
Question 3	5.8	3.0	3.7

effect on essay score based on the medium used to learn the case study, $F(2, 65) = 1.33$, $p > .25$. This indicates that a student's course performance and the group he was placed in were relatively independent of each other.

We further analyzed the data by question. Table 2 presents the mean score of each group on each question.

As can be seen, students in the interactive CD group did not outperform their counterparts on the first question – the question where students were asked to list the facts of the case. An ANOVA yields no significant main effect by medium for mean group score on the first question, $F < 1$, $p > .5$. The second and third questions probe students' knowledge about the perspectives of individuals associated with the case and their understanding of the issues involved with the case, respectively. Each of these questions contained multiple parts. The second question contained six parts and thus a student could gain maximum 24 points for this question. The third question contained four parts and thus a student could gain 16 points maximum. The interactive CD group outperforms the other groups on the second question, producing a marginally reliable effect, $F(2, 65) = 2.76$, $p = .07$. The interactive CD group significantly outperforms the other groups on the third question. An ANOVA yields a significant effect by medium for mean group score on the third question, $F(2, 65) = 9.95$, $p < .001$. Further, Fisher's PLSD reveals statistically significant effects for individual contrasts between the interactive CD group and the text group and between the interactive CD group and the film group, with the alpha-level of significance being .05.

Discussion

This study provides solid evidence that students in the interactive CD group outperform students in the text group and film group. These results are entirely

consistent with the results of the other reported studies. The fact that a statistical difference was obtained in this study, but not in the second study, may be attributed to the larger sample size used in this study.

The mean scores in this experiment were lower than in the previous experiments we conducted. We hypothesize three causes for this result. First, the last question was dropped from the assessment, so there were 4 fewer points that students could obtain. Second, the population taking this class was in a general introductory philosophy class, not an ethics class. We used this group of students in order to achieve a sufficiently large sample size to obtain statistically reliable results (less than 40 students enroll in the ethics course per semester). However, the students in the general philosophy class may have been less motivated and less able to answer questions pertaining to an ethics case study. Third, unlike the previous experiments, performance on the assessment did not count toward a course grade.

Students in the interactive CD group outperformed their counterparts, but only on the second and third questions. In this study, students using the interactive CD did no better than other groups in reciting the facts of the case. (Although there is an interactive analysis of the facts provided in the first menu option of the Guided Inquiry section, this section was not used in our studies due to time constraints on the assignments. Use of this feature of the interactive CD might have had an advantageous impact on student performance with question 1 as well.)

SUMMARY AND CONCLUSION

In conjunction, these studies provide strong evidence that students learning about the Dax Cowart case using an interactive multimedia CD will acquire more knowledge about the case than students using traditional text and film representations of the case. These studies have replicated this main result three consecutive times. In two of these studies, we were able to obtain statistically reliable results. Finally, the effectiveness of the interactive CD was demonstrated with two populations – students in an ethics class and students in a general philosophy class.

We wish to emphasize the importance of these results for the ongoing conversation about the effectiveness of interactive multimedia and to highlight the role that the media played in *critical reasoning*. In this study, students in the interactive CD group outperformed students in the text and film groups

Mean group score on Dax Cowart assessment			
Medium	CD	Film	Text
Study 1	30.3	27.2	28.8
Study 2	22.6	18.6	21.1
Study 3	17.7	12.2	14.6

with regard to (a) *understanding* the complex perspectives and positions of the principals in the case and (b) *analyzing* the case with respect to its morally relevant details. And they were able to do this in a way that allows us to predict, with statistical assurance, that they will out perform students who use text or film as alternate representations of the case.

Given these outcomes, it is possible that the conclusions of the study may well reach beyond the particular interactive CD involved. By leveling the field in regard to assignment, content, and student background, the study could indicate some *general advantages* of interactive multimedia. The hypothesis that students who participate in *reflective engagement* with real and compelling content (e.g., the Guided Inquires section of the interactive CD) may have significant learning advantages over those who only read text or view videos seems to be supported by this study – at least in the realm of case analysis. In this regard, the study might place discipline-based, pedagogically sound interactive multimedia ahead of text and linear video in terms of student comprehension and retention. But this takes us into the Methods / Medium controversy (Clark, 1994b; Fisher, 2000; Kozma, 1994) and beyond the scope of the current discussion.

REFERENCES

Andersen, D., Cavalier, R., & Covey, P. (1996). *A Right to Die? The Dax Cowart Case*. New York: Routledge.

Cavalier, R. (1996). Evaluating evaluation in light of discipline-specific computational turns. *Journal of Computing in Higher Education*, *8*, 85–96.

Cavalier, R., & Covey, P. (1996). *A Right to Die? The Dax Cowart Case CD-ROM Teacher's Guide*. New York: Routledge.

Clark, R.E. (1994a). Media will never influence learning. *Educational Technology Research and Development*, *42*, 21–29.

Clark, R.E. (1994b). Media and Method. *Educational Technology Research and Development*, *42*, 7–10.

Clouser, K.D. (1993). Hastings Center Report, Special Supplement, *25*, S11.

Covey, P. (1993). The crucible of experience. In: R. Heeger and T. van Willigenberg (Eds.), *The Turn to Applied Ethics* (pp. 55–72). Kampen, Netherlands: Kok Pharos Publishing House.

Ehrmann, S. (1996). Asking the right question: What does research tell us about technology and higher learning? *Change Magazine* (March / April). Vol. no. 27 pp. 20–28.

Fisher, S. (2000). Medium, method, and message: *Why we can* measure the pedagogic effectiveness of instructional technology. Paper presented at the Computing and Philosophy Conference August 2000.

Kliever, L. (ed.) (1989). *Dax's Case: Essays in Medical Ethics and Human Meaning*. Dallas, TX: SMU Press.

Kozma, R.B. (1994). A Reply: Media and methods. *Educational Technology Research and Development*, *42*, 11–13.