DOCUMENT RESUME

ED 401 287 TM 025 779

AUTHOR Reeves, Thomas C.; And Others
TITLE New Approaches to Cognitive Assessment in Engineering Education.
PUB DATE Apr 96
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Measurement; College Freshmen; *Educational Assessment; *Engineering Education; Higher Education; Military Personnel; Pilot Projects; Problem Solving; *Self Evaluation (Individuals); *Thinking Skills
IDENTIFIERS Air Force Academy CO; *Problem Based Learning; Reflective Judgment

ABSTRACT

This paper describes the development, implementation, and effects of new approaches to cognitive assessment within an undergraduate engineering course at the U.S. Air Force Academy. The course, ENGR 110, "Introduction to Engineering," was designed to be a problem-based learning environment in which cadets worked in teams to solve problems integral to a "Mission to Mars," for example, getting to Mars, constructing a research site on Mars, and developing a renewable power source there. In addition to traditional knowledge and skill objectives, the course was focused on "higher order" outcomes such as "framing and resolving ill-defined problems," "communicating via multiple media," "exhibiting intellectual curiosity," and "developing a rich conceptualization of engineering." Forty-two freshman cadets participated in the pilot testing of the cognitive assessment methods used to assess student achievement with respect to these outcomes. Assessments included: (1) reflective judgment exercises; (2) self-assessment questionnaires; (3) concept maps; (4) focus groups; (5) e-mail journals; (6) observations; and (7) individual and group interviews with faculty. Results indicated statistically and educationally significant differences in "problem solving" between ENGR 110 students and two control classes of sophomore engineering students. Other results supported research and development. (Contains 2 tables, 2 figures, and 14 references.) (Author/SLD)

NEW APPROACHES TO COGNITIVE ASSESSMENT IN ENGINEERING EDUCATION

Thomas C. Reeves, James M. Laffey, & Mary R. Marlino

NEW APPROACHES TO COGNITIVE ASSESSMENT IN ENGINEERING EDUCATION

PAPER PRESENTED AT THE ANNUAL MEETING OF THE AMERICAN EDUCATIONAL RESEARCH ASSOCIATION (AERA), NEW YORK, NY, APRIL 8-12, 1996, SESSION 3.32

THOMAS C. REEVES, PH.D.
Instructional Technology
607 Aderhold Hall
The University of Georgia
Athens, GA 30602-7144
Voice: 706/542-3849
Fax: 706/542-4032
treeves@coe.uga.edu

JAMES M. LAFFEY, PH.D.
College of Education
304 Townsend Hall
The University of Missouri
Columbia, MO 65211
Voice: 314/882-5399
Fax: 314/884-5455
cilaffey@showme.missouri.edu

MARY R. MARLINO, ED.D.
Center for Educational Excellence
HQ USAFA/DFTE
U. S. Air Force Academy
Colorado Springs, CO 80840-5000
Voice: 719/472-2549
Fax: 719/472-3135
MarlinoMR%DFE%USAFA@dfmail.usafa.af.mil

(The contents of this paper are solely the responsibility of the authors, and as such do not represent an official position or statement of the U. S. Air Force or other federal agencies.)

ABSTRACT

This paper describes the development, implementation, and effects of new approaches to cognitive assessment within an undergraduate engineering course at the U. S. Air Force Academy (USAFA).* The course, ENGR 110, "Introduction to Engineering" was designed to be a problem-based learning environment in which cadets worked in teams to solve problems integral to a "Mission to Mars," e.g., getting to Mars, constructing a research site on Mars, and developing a renewable power source there. In addition to traditional knowledge and skill objectives, the course was focused on "higher order" outcomes such as "framing and resolving ill-defined problems," "communicating via multiple media," "exhibiting intellectual curiosity," and "developing a rich conceptualization of engineering." Several cognitive assessment methods were used to assess student achievement with respect to these outcomes. Results indicated statistically and educationally significant differences in "problem solving" between ENGR 110 students and two control classes of sophomore engineering students. Other results supported further research and development.

* This study was conducted by Dr. J. M. Laffey of the University of Missouri and Dr. T. C. Reeves of The University of Georgia under subcontract with the University of Colorado at Denver and the U. S. Air Force Armstrong Laboratory. Other members of the assessment team included Dr. Mary R. Marlino and Mr. Curtis Hughes from the Center for Educational Excellence at USAFA, Mr. Kevin Oliver, a doctoral student intern at USAFA from The University of Georgia, and the USAFA faculty involved in designing and implementing ENGR 110. The guidance and support of COL. M. L. Smith and his colleagues at USAFA, and Dr. C. L. Lynch, Reflective Judgment Associates, have been invaluable in this research.

DATE: 08/08/96

ERIC/TME Clearinghouse Report

Document Cover Sheet

TM #: 025779

Title: NEW APPROACHES TO COGNITIVE ASSESSMENT IN ENGINEER

Number of Pages: 20

Publication Date: 04/08/96

Document Level: 1

Notes:

INTRODUCTION

This paper describes the development, implementation, and effects of new approaches to cognitive assessment (Merluzzi, Glass, & Genest, 1986) within a new undergraduate Engineering course at the U. S. Air Force Academy (USAFA). In addition to traditional knowledge and skill objectives, the course is focused on "higher order" outcomes such as "framing and resolving ill-defined problems," "communicating with multiple media," "exhibiting intellectual curiosity," and "developing a rich conceptualization of engineering." Assessment in professional education, defined as the process of determining whether students have attained curricular goals, has traditionally focused on retention of knowledge and its application in severely limited contexts as measured by standardized tests (Wiggins, 1993). Increased interest in what has been labeled "alternative assessment" reflects frustration with traditional approaches to assessment as well as the sincere desire to assess the attainment of higher order educational goals that involve deep understanding and active use of knowledge in complex, realistic contexts (Herman, Aschbacher, & Winters, 1992). The interest in alternative assessment is reflected in the proliferation of terms such as 1) authentic assessment (Puckett & Black, 1994), 2) performance assessment (Wiggins, 1993), and 3) portfolio assessment (Knight, 1994). This project involved a form of alternative assessment called "cognitive assessment" to emphasize the focus on students' higher order thinking abilities, attitudes, and communication skills (Merluzzi, et al., 1986).

A great deal of inference is involved in measures that are used to assess higher order thinking skills and attitudes. Cognitive assessment measures (e.g., a concept map exercise) may lack the "face validity" held by more traditional measures of academic achievement. Although face validity is not as important as construct validity in a purely psychometric sense, the former is essential if alternative assessment procedures are to be accepted within traditional academic communities such as engineering schools and other contexts for professional education and training. Not being directly observable, rich conceptual knowledge can only be inferred from students' performance on a range of assessments. Cognitive assessment is difficult because it is relatively easy for students to mimic conceptual knowledge by repeating answers to previously presented questions rather than generating unique responses to novel problems. Faculty often unwittingly participate in this charade by over-emphasizing the need to get "the right answer" rather than striving to detect the thought processes students have used to arrive at answers. On the other hand, assessments based upon completely novel problems may be so far removed from the "situations" in which learning took place that the likelihood of transfer to professional practice is remote. Cognitive assessment walks the fine line between traditional testing procedures that assess low level knowledge and skills and highly inferential measurement scales that have little relationship to the content or context of any given course of study.

CONTEXT

Our research focused on developing the means to assess students' abilities to: 1) frame and resolve ill-defined problems, 2) communicate their thought processes, problem-solving approaches, and solutions via multiple media, 3) exhibit intellectual curiosity, and 4) develop a rich conceptualization of engineering. These assessments were designed within the context of a new course, "ENGR 110 - Introduction to Engineering," which is innovative in several ways. First, the course integrates concepts and skills from all five engineering disciplines represented at USAFA. Second, the course includes both traditional learning objectives and higher-order educational outcomes, e.g., 1) concepts, 2, basic laws, 3) model building and problem solving, 4) oral, written, and graphical communication skills, 5) independent learning and metacognitive skills, 6) intellectual curiosity, and 7) teamwork skills. Third, learning opportunities within ENGR 110 are situated within the context of a challenging scenario, viz., the establishment of a research station on Mars.

Three major modules are included the course: 1) getting to Mars, 2) siting and constructing a research station on Mars, and 3) operating and maintaining an energy plant there. At the end of each module, cadet teams make formal military briefings attended by their peers, instructors, and others from USAFA. The course also involves "hands-on" projects such as building and flying model rockets. All cadets have their own computers in their rooms with access to many electronic resources such as word-processing, spread-sheets, multimedia presentation software, e-mail, and the World Wide Web. ENGR 110 was first offered in the Fall Semester of 1995 to 42 freshmen cadets at the U. S. Air Force Academy (USAFA). Cadets participating in this prototype course were selected at random from an entering class of more than 1,000 cadets.

Our overall goal was to develop valid, reliable, and feasible assessment strategies for this innovative engineering course. With respect to validity, we were primarily concerned with construct validity, i.e., "Are we measuring what we claim to be measuring?" With respect to reliability, we focused on the ability of the assessments to be administered, analyzed, and interpreted by faculty members in a consistent manner. With respect to feasibility, we were primarily concerned with whether the assessments were implemented in a timely, accurate, and cost-effective manner without placing unrealistic demands on both students and faculty. In addition to these three primary factors, we attended to the effects of these cognitive assessments on both faculty and students as well as its integration into the academic and military culture of USAFA.

METHODS

Based upon a review of the research literature on cognitive assessment within the context of undergraduate engineering programs, the following assessment and research strategies were utilized: 1) Reflective Judgment Exercises, 2) Self-Assessment Questionnaires, 3) Concept Maps, 4) Focus Groups, 5) E-mail Journals, 6) Observations, and 7) Individual and Group Interviews with Faculty. Figure 1 summarizes the primary methods used to assess Framing and Resolving Ill-defined Problems (P), Communication (C), Intellectual Curiosity (IC), and Conceptual Understanding (U).

Method	Outcome	Notes
Reflective Judgment Interview	P	Pre and post Reflective Judgment interviews (Lynch, Kitchener, & King, 1994) were used for ENGR 110 and for a control group. Reflective Judgment Scoring Rules (Kitchener & King, 1985) were used and represent a measure of critical thinking. We used this measure to judge the students' generalization of the ability to frame and resolve ill-defined problems which was practiced during the course.
Self Ratings Questionnaires	P, C, IC, U	Students were asked to rate their own performance across the four outcomes.
Concept Maps	U	Students were asked to develop a pre and a post concept map of engineering (Novak & Gowin, 1984).
Focus Groups	P, C, IC, U	Three groups of five to eight students were interviewed in a focus group roundtable.
Journals	P, C, IC, U	Students were asked to keep journals of the engineering and inquiry efforts (Fellows, 1994). Question stems were used to support reflection on the entries. The journal also served as a log for recording out of class tasks, such as using the library or the Internet for research.
Observer Ratings	C, IC	Classroom observers recorded student performances of questioning, responding to questions, and reporting during class time.
Faculty ratings	P, C, IC, U	Faculty were asked to rate student performance across the four outcomes.

Figure 1. Cognitive assessment methods for ENGR 110.

SAMPLE

Forty-two freshman cadets participated in the initial pilot testing of the cognitive assessment measures within the context of the new ENGR 110 course. Eight engineering faculty participated in the implementation of the pilot course. Thirty-nine sophomores enrolled two sections of ENGR MECH 120, an introductory mechanical engineering courses, served as a control group. A central facet of the study was a quasi-experimental comparison of the cognitive outcomes of the new problem-based course (ENGR 110) and the traditional, instructor-led engineering course (ENGR MECH 120).

RESULTS

The results are divided into four sections related to the four cognitive outcomes listed above (Framing and Resolving Ill-defined Problems, Communication, Intellectual Curiosity, and Conceptual Understanding). Each section below includes a summary headline, data, discussion, and recommendations.

SECTION 1 – CADETS ENROLLED IN ENGR 110 INCREASED THEIR ABILITY TO FRAME AND RESOLVE ILL-DEFINED PROBLEMS.

Data

Three types of data were collected to assess cadets' ability to frame and resolve ill-defined problems as well as to assess their perceptions of problem-solving as a major focus of the ENGR 110 course:

(1) The Reflective Judgment Exercise (RJE) provided evidence that freshmen cadets improved their abilities to frame and resolve ill-defined problems as a result of their participation in the ENGR 110 course.

(2) Cadets self-reported (through focus group sessions, end-of-course questionnaires, and e-mail surveys) increased awareness of problem-solving as a focus of the ENGR 110 course as well as improved abilities to frame and resolve ill-defined problems.

(3) During interviews, ENGR 110 faculty reported they had observed cadets engaging in problem-solving during the course.

1. Reflective Judgment Exercise (RJE)

The Reflective Judgment Exercise (RJE) (King & Kitchener, 1994) was administered to all ENGR 110 cadets (freshmen) as well as to a comparison group of ENGR Mech 120 cadets (sophomores) at the beginning and end of the Fall 1995 semester. The pre-test problem involved a "sortie" scenario and the post-test involved a "desert survival" scenario. The reflective judgment exercises were all scored by Dr. Cindy L. Lynch, a consultant from Reflective Judgment Associates, using a procedure which kept her blind to the cadets' membership in the 110 or Mech 120 courses. Dr. Lynch is a professional consultant who has trained USAFA personnel to administer and score the RJE. Table 1 presents the pre-test and post-test results for the cadets enrolled in two sections each of ENGR 110 and ENGR Mech 120.

On the pre-test, the ENGR Mech 120 sample of sophomores (N=39) in the comparison group achieved a mean score of 2.23, and the ENGR 110 freshmen (N=42) achieved a mean of 2.02. These differences are not statistically significant. Both the ENGR Mech 120 (N=22) and the ENGR 110 (N=35) cadets scored higher on the posttest, but only the ENGR 110 cadets showed statistically significant gains which can be attributed to the experiences of the course. The differences in the ENGR 110 scores are also educationally significant in that they represent a shift from generally "deficient" problem-solving to generally "satisfactory" problem-solving (see Figure 2 below). The score differences between the ENGR 110 and ENGR Mech 120 classes are also statistically significant. The RJE instruments are scored using a 1 to 5 scale in which "1" represents deficient (D) problem-solving, "2" represents D+ or S- scores, "3" represents satisfactory (S) results, "4" represents S+ or E- scores, and "5" represents excellent (E) problem-solving. The ENGR 110 cadets improved their RJE scores by approximately one standard deviation as illustrated in Figure 2.

Table 1

Mean Pre-Test and Post-Test Scores on the RJE for ENGR 110 and ENGR MECH 120

	Pre			Post		
Section	n	mean	sd	n	mean	sd
110	20	2.05	.6	17	3.0*	.8
110	22	2.0	.75	18	2.8*	.9
Mech	19	2.16	.6	12	2.25	.75
Mech	20	2.3	.7	10	2.5	1.2

* Statistically significant pre- to post-test differences $p < .00001$.

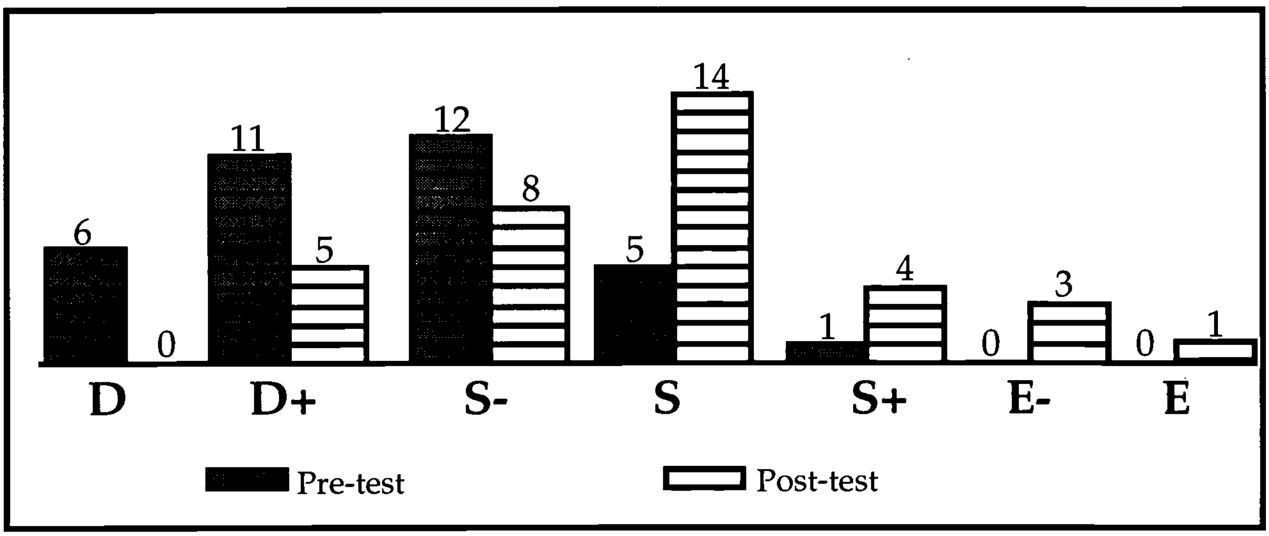

Figure 2. Pre-test and post-test scores for ENGR 110 cadets (n=35) with complete data.

Although the improvements in RJE scores for the ENGR 110 students are impressive, there were two limitations in the administration of these tests that should be kept in mind when interpreting the findings. First, there was a loss of subjects from pre- to post-test, especially in the ENGR Mech 120 classes (see Table 1). Second, pre-tests were given as an in-class assignment on the first day of the course and the post-tests were given as an extra credit question on the final exam.

2. Self-Reports

2a. Focus Groups

Three focus group sessions were held during the semester with groups of cadets in ENGR 110. Each session was held at the end of one of the three major "Mars Mission" modules that guided the project-based learning activities in the course, i.e., "getting to Mars," "constructing a research station on Mars," and "establishing a renewable power source on Mars." Cadets were rewarded for their participation in focus groups with pizza and ice cream. The sessions were recorded and subsequently transcribed. The focus group sessions provide evidence that cadets saw the ENGR 110 course as different from other courses in terms of its emphasis on applying knowledge to solve ill-defined problems and in the requirements placed by faculty on original thinking and constructing their own solutions rather than memorizing the instructor's solution. Example statements include:

I think it's good we learned to think critically... there's not just one way to get to the right answer, and there's not just one right answer.... we always have to be thinking of what we can do next.

The class makes me think more than other classes do. There's no answer in the back of the book. There is no book!

I think the best part is the idea of going to Mars, using the equations for something. Not just learning bland equations, applying them. [You] use your intellect to think through a problem that hasn't been presented to you, like what would happen on Mars with this equation... think through it yourself and make sure.

The homework is different... it's mainly ill-defined problems. I wasn't expecting this and it made me nervous early on, but I have been working at it and getting good results.

I think I've gotten better at that (ill-defined problems) too ... in the beginning I wasn't very good at solving ill-defined problems. I think I improved upon that as I went through the first section.

This course more or less goes against everything most of us learn in high school - usually high school kids get the right answer and if you don't get it you are wrong. It's simply that way. There's one way to start and one way to finish and there's no different paths to choose. I don't think the real world is that way. I think the real world is full of ill-defined problems, and therefore I think the course was beneficial... it helped us to do that.

2b. Questionnaires

At the beginning and end of the semester, cadets in ENGR 110 completed self-assessment surveys. These instruments used a Likert scale with 6 response options ranging from "Strongly Disagree" to "Strongly Agree." On the post-course questionnaire (n=40), 85% of cadets agreed (15% disagreed) with the statement, "My abilities to frame and resolve ill-defined problems have improved as a result of this course." Although this finding in and of itself may not seem especially important, it complements the other data we found through our triangulation approach to assessing whether students improved in their ability to frame and resolve ill-defined problems as a result of their participation in this prototype engineering course.

2c. E-Mail Surveys

Following each of the three course modules, cadets used e-mail to respond to a set of questions about the course and the specific project just completed. After the first project, cadets were asked an open-ended question about what they had learned from the project. Eighteen of the 37 cadets (48%) identified that they had learned that engineering problems are ill-defined, complex, and require reasoning, and/or that engineering requires making assumptions, experimentation, and learning from errors. For example, one student stated:

I learned that engineering is not a subject that is as defined as a class such as Calculus. There are so many ill defined problems to be considered that the predictions that we make are not as sure as a simple algebra equation. For example, our predictions on the trajectories of our rocket were not accurate because we did not consider drag.

Following the final module, cadets were again asked what they had learned. Twenty-five of the 32 respondents (78%) identified recognizing that problem solving was a key part of engineering as what they had learned. For example, one cadet stated:

I learned a lot about thinking about problems when there is not a really good answer.

Twelve cadets focused on the complexity of doing real world engineering. Nine recognized that engineering requires going beyond formulas and equations. Four mentioned learning how to apply technical information to solve ill-defined problems. The pattern of evidence in these three e-mail surveys clearly supports the overall conclusion that students learned that framing and resolving ill-defined problems was a major part of the engineering process.

3. Faculty Reports

At several times during the semester and in a group interview near the end of the course, faculty and course observers were asked if they felt ENGR 110 contributed significantly to the ability of cadets to frame and resolve ill-defined problems. There was clear consensus that the presentations made by the cadets three times during the semester provided evidence of effective problem-solving. To make an effective presentation, the cadets had to pull together complex aspects of the problem, decide how to focus their presentation, and respond to questions. Faculty also commented, however, that cadets often failed to consider all aspects of the problems with which they were confronted. For example, the faculty observed that significant factors that would add to the complexity of the solutions proposed by cadets were often over-looked. The consensus of the faculty was that, although cadets had improved in their ability to recognize problems as ill-defined, there was much room for improvement in terms of the actual problem-solving strategies they used.

Discussion

In the absence of any single reliable, valid, and feasible method of assessing this higher-order outcome of "framing and resolving ill-defined problems," the assessment team employed a triangulation approach. Triangulation usually involves using multiple measures to converge on a more accurate estimate of the "true" value of a variable (Mark & Shotland, 1987). The triangulation approach assumes that errors in one type of measure are canceled out by errors in another type. In ENGR 110, we used the RJE, focus groups, questionnaires, e-mail surveys, and faculty observations to assess changes in cadets' perceived and actual ability to frame and resolve ill-defined problems.

The overall results of the approaches used to assess "framing and resolving ill-defined problems" are very positive. Cadets in ENGR 110 clearly developed an awareness of ill-defined problems in the work of engineering. The RJE provides evidence that the cadets also demonstrated improved capabilities to recognize problems as ill-defined and to utilize better approaches to solving ill-defined problems. The improvement on the RJE scores is both statistically and educationally significant, and a strong case can be made that these results stem from the experiences cadets had in ENGR 110. Cadets in this new course moved from an average score of deficient to an average score of satisfactory (see Figure 2), an increase described by Dr. Lynch as "very surprising."

This improvement in framing and resolving ill-defined problems evidenced by the RJE was also seen systematically in the ENGR 110 activities and the briefings cadets made to report these activities. Overall, we interpret these findings as representing a developmental shift from deficient to satisfactory problem-solving. The major factors in these developments appear to be the ability of the cadets to go beyond the facts given in the written case and their ability to identify assumptions other than those specified in a given problem statement. The improved problem-solving ability indicated by the RJE results corresponds to instructional practices in ENGR 110 which required cadets to think about whether their answer or solution to a problem was sufficient or "best." These results also correspond with evidence we found in the self report data that cadets improved their understanding that engineering problems are complex and that there may not be just one "right" answer to a given problem.

Recommendations

The results of the ENGR 110 course to build cadet capacity for framing and resolving ill-defined problems are noteworthy, and important progress was made in developing assessment strategies for measuring these learning outcomes. The self-report data describe cadet beliefs and interpretations of their own achievements which are corroborated by the RJE scores. Next steps in the assessment process should be to continue the self report and RJE assessment approaches to look for consistency and patterns of achievement. In the coming semesters, efforts should be made to keep any comparison groups in tact for pre- and post-tests. Most importantly, efforts should be made to insure similar circumstances for administering the pre- and post-tests.

Perhaps more importantly, efforts should be made to look for feasible ways of assessing the ill-defined problem-solving which takes place within the context of ENGR 110 activities. The skills, abilities, and habits of mind which make for successful and consistent framing and resolving of ill-defined problems go well beyond what is judged by the RJE. A pilot assessment strategy should be developed to assess problem-solving skills (including reflective judgment) within the context of an authentic task presented to the cadets in ENGR 110 during the Spring 1996 semester. For example, during the time that cadets are working on the problem of establishing a renewable power source on Mars, a crisis scenario could be introduced in which a major assumption of the proposed solution was undermined. Observations and post-problem interviews could be used to assess cadets' abilities to frame and resolve this type of realistic problem.

SECTION 2 – AS MEMBERS OF PROJECT TEAMS, CADETS INCREASED THEIR ABILITY TO COMMUNICATE TECHNICAL INFORMATION AND PROBLEM SOLUTIONS THROUGH FORMAL BRIEFINGS

Data

Two types of data were collected to assess cadets' communication skills:

(1) Cadets self-reported (through focus group sessions, end-of-course questionnaires, and e-mail surveys) improvements in their ability to communicate through formal briefings in the ENGR 110 course.

(2) A second indicator of communication skills came from observations of classroom briefings and instructor judgments of improved communication skills.

1. Self- Reports

1a. Focus Groups

The three focus group sessions conducted at the end of each major module of the course provided evidence that cadets perceived that their own abilities to communicate as well as the abilities of their peers improved substantially over the semester. Cadet comments showed an increased appreciation for the value of communicating through military briefings. In addition, it was clear that the cadets recognized that they improved their communication skills because of opportunities in the course such as the feedback they received from faculty and the assignment to provide multiple briefings.

Example statements from the focus groups include:

(Dialog from Focus Group)

S1 - *I learned about briefing. I think that there is a difference between presentations and briefing. I never gave a briefing before in high school.*

S2 - *I think that was the best thing I learned too.*

S3 - *Yes, I think so also.*

S4 - *I think that's what I learned the most too.*

S1 - *We're kind of forced to use transparencies, which I had never used before in high school, and I used PowerPoint for the first time.*

I liked doing presentations. I know we are going to have to do that throughout our Air Force careers and in our lives. I had no idea about the way it was supposed to be done before this. It was hard. We were up many times til 2:00 AM in the morning the night before, but that just because we are procrastinators. I liked that part a lot.

Yeah, I got a lot better [at briefings]. I though the material was more difficult. Like the power thing, we had a heck of a time finding hydro and geothermal. I thought that was very difficult.

1b. Questionnaires

On the end of course self-assessment questionnaire, 87% of cadets agreed (13% disagreed) with the statement, "My abilities to communicate have improved as a result of this course."

1c. E-Mail Surveys

Following each of the three course projects, cadets used e-mail to respond to a set of questions about the project. Following the final project, students were asked what they had learned from ENGR 110. Responding to this open-ended question, fourteen of the 32 respondents (44 %) identified working well on a team and communicating with team members. Eight stated (25%) that they learned how to brief effectively.

2. Observations and Faculty Judgments

We did not have an instrument in place to measure change from the first presentation made by the cadets to the last, but both faculty and observers noted improvements such as a) the increased use of and improvement of computer slides, b) better organization, and c) improved poise and confidence. At the end of the semester, faculty and course observers were asked if they felt the course contributed significantly to the communication skills of the students. There was clear consensus that the cadets' presentations had improved substantially. Particular mention was made of the improved use of graphics and data displays. Faculty also cited improved papers and remarked about the quality of the 3-D representations in the civil engineering section. Observers who observed most of the ENGR 110 classes as part of the assessment team reported that the improvements in the cadet briefings were obvious to them.

Discussion

The dialog presented above came from the last focus group session of the semester. Cadets clearly believed that they had improved their ability to present briefings over the semester. ENGR 110 afforded cadets several opportunities to present briefings, opportunities to observe others present briefings, and a chance to build concrete skills (e.g., skills in using Microsoft PowerPoint presentation software). In addition, opportunities to collaborate in teams in terms of briefing preparation, in-depth discussions, and informal reviews were supported by the course activities. Most importantly, cadets received feedback from the faculty that shaped subsequent briefings through class discussions and individual guidance. Cadets recognized the value of improving their ability to present technical briefings as critical to their future in the Air Force, and took obvious pride in their improvement over the semester. For many cadets this improvement appeared to be the most significant outcome of the course. The briefings provided very compelling, high stakes learning environments (especially when attended by ranking visitors). Cadets appeared to take the briefings very seriously and no doubt learned from their errors and the errors of other.

Recommendations

The improvement in student communication skills for technical briefings appears to be a successful outcome of the ENGR 110 course. However, more targeted ways of measuring improvements and the overall level of achievement are needed, including a set of criteria to measure the quality of briefings and an instrument for consistently rating briefing performances. The instrument should include ratings of organization, use of presentation software features, command of technical issues, and responsiveness to questions and criticism. In addition to faculty and observers, the other cadets in the class should rate the performance of each team conducting a briefing. Having cadets rate the performance of their peers will provide valuable learning opportunities as well as additional assessment data. Of course, peer ratings should be anonymous and have no bearing on grades.

The briefings provide considerable "natural" feedback for how well cadets presented ideas, but because multiple presentations were packed into such a short time frame, valuable opportunities to probe the cadets' thinking and problem-solving approaches were probably lost. Faculty would adopt probing questions to ask at the conclusion of presentations to encourage deeper planning and reflection. For example, questions similar to the ones on the second part of the Reflective Judgment Exercise (RJE) could be asked at each debriefing. Typical questions might be "How confident are you in your proposed solution?" and "What additional information do you wish you had had before making this briefing?" This type of questioning would encourage cadets to build habits of reflection within the context of their course assignments.

SECTION 3 – DESPITE SUBSTANTIAL BARRIERS, THERE IS EVIDENCE THAT CADETS WENT BEYOND THE EXPECTATIONS OF ENGR 110 AND EXHIBITED INTELLECTUAL CURIOSITY IN THE COURSE.

Data

Two types of data were collected to assess cadets' exhibition of intellectual curiosity:

(1) Cadets self-reported (through focus group sessions, end-of-course questionnaires, and e-mail surveys) their beliefs with respect to opportunities to exhibit intellectual curiosity and the ways in which they exhibited this attribute.

(2) A second indicator of intellectual curiosity came from faculty observations and reports about cadets and their work.

1. Self-Reports

1a. Focus Groups

Focus group sessions held at the end of each module provided evidence that students became interested in the challenge of going to Mars. In fact, the project orientation and the Mars Mission were often mentioned as among the most positive aspects of the ENGR 110 course. Example statements include:

(speaking about classmates) Mostly it's how much they enjoy it, I mean, the idea of being in there, thinking we're going to Mars, just the excitement of human exploration.

In my opinion, I think it [Mars Mission] was a good idea. Sometimes we'll be in calculus or other subjects, and the real-world application does not seem to fit. You're just chugging away at problems and have no actual idea what they relate to. It seemed like every problem we were given [in ENGR 110] had some bearing on our mission to Mars which in the future could be possible.

Also in the focus groups, cadets occasionally reported behaviors that we interpreted as evidence that they had gone beyond the requirements of the course and exhibited intellectual curiosity. For example, cadets said:

I read a book on Interstellar Planetary Trajectories and got a lot out of that, but it was something the class kind of got me into, but it wasn't something that I really got out of the class.

I've found so much out there that is just out there for the taking. I've gone to Georgia Tech [on the Web] and hooked up to their library, tried to go to M.I.T. and they wouldn't let me in. I mean you can just go anywhere and find out so much if you have the patience to use it.

During focus group sessions, cadets also described substantial barriers to developing and pursuing intellectual interests beyond the minimum requirements of a given course. Students identified time, logistics, and fatigue as constraints on their ability or willingness to invest in additional intellectual pursuits. For example, one cadet said:

....I just think the class takes a lot of time. It's not so much it's difficult, and also working in groups, you know, being freshmen, we can't exactly walk over to the next guy's room at 8:00 PM. We've got to set up these lab room meetings, we have to run there and certain things come up, and then certain stuff like training sessions, so we couldn't work on Friday nights. Just group work, that takes a lot of time. You get there at 8:00 PM, your partner doesn't think it starts until 8:30 PM. Just stuff comes up like that.

1b. Questionnaires

At the beginning and end of the semester, cadets in ENGR 110 completed self-assessment questionnaires. Most of the items on these instruments used a Likert scale with 6 response options ranging from "Strongly Disagree" to "Strongly Agree." On the pre-course questionnaire, 93% of the freshmen cadets agreed with the statement, "If I find a topic interesting, I will pursue it on my own beyond what is needed to do well in a course." On the end of course questionnaire, 62% of students agreed (38% disagreed) with the statement, "I went beyond the expectations and assignments for this course." The majority of cadets fell within the middle groups with only two cadets giving a response of "strongly agreeing."

1c. E-Mail Surveys

Following each of the three course projects, cadets responded via e-mail to questions about the projects. After the first project students were asked what they had done beyond the instructor's requirements. Eleven cadets stated "little to nothing." Ten mentioned additional research or reading. And twelve described additional work to improve the products required for their course assignments. Following the final project, students were again asked the question about what they had done beyond the requirements for ENGR 110 for project 3. At this point, only six cadets stated they had done nothing. Twenty-six cadets described additional work they had done to advance their understanding or improve their products for the course. Examples included additional effort on slides for their presentations, using the Internet for research, and meeting more often as a group.

2. Faculty Judgments

At several times during the semester and in a group interview near the end of the course, faculty and course observers were asked if they felt the course contributed significantly to raising intellectual curiosity among the students. There was no clear consensus on this question, but several examples of intellectual curiosity were cited. Examples included bringing in outside resources to the class discussion and the instance of a student who had researched the legal aspects of the Mars Mission. The faculty also reported that they perceived that the cadets were strongly interested in the Mars Mission as the overall context for this course.

Discussion

While it is clear that the Mars Mission captured the interest of the cadets in ENGR 110, we have only limited evidence that cadets were so intellectually engaged and curious that they went beyond the requirements in this course. Self-report data indicated that two-thirds of the cadets believed that they had gone beyond the requirements of the course at some time during the semester, but it is unclear whether this was motivated by genuine intellectual curiosity or by the desire to improve their grades.

On the other hand, some might argue that in light of the mental, emotional, and physical demands placed upon first year USAFA cadets, it may be unrealistic to expect that cadets will perceive many opportunities to exhibit intellectual curiosity or that they will follow through on these curiosities in a systematic manner. The fact that both the self-report data and the faculty observations provided some, albeit modest, evidence of intellectual curiosity among the cadets in ENGR 110 should not be dismissed. Much work remains in the effort to assess intellectual curiosity. First, a clearer definition of what counts as intellectual curiosity within the context of engineering education at USAFA is needed. As noted above, one of the difficulties of assessing this variable is that many of the behaviors which can be taken as evidence of intellectual curiosity could also be interpreted as simply "extra work" undertaken in pursuit of a higher grade. More systematic processes for assessing this important outcome are required.

Recommendations

One area where it may be fruitful to examine the intellectual engagement of cadets is in their use of technology, especially their voluntary exploration of Internet resources related to the content and activities of the ENGR 110 course (Reeves & Okey, 1996). The easy availability of the World Wide Web (WWW) for USAFA cadets overcomes some of the logistical limitations to pursuing interests in the non-electronic world. A prototype WWW has been developed to support the ENGR 110 course. It may be feasible to track when cadets access this home-page as well as when they attempt to link beyond the resources provided on the ENGR 110 home-page to other WWW resources related to the course content and activities.

A second area which may represent intellectual curiosity is the willingness of the cadets to take on the challenge of ill-defined problems. Most cadets are not prepared for the responsibility of defining problems, and they freely admitted to us that it is safer and easier to get involved in traditional school problems (where the instructor defines the problem and the cadets' job is to find the answer that the instructor already knows). Cadets described transforming to the type of problem-solving work required in ENGR 110 as challenging and risky. Perhaps a measure of intellectual curiosity can be found in the enthusiasm cadets bring to these types of problems or their willingness to take on the added responsibility and risk when given a choice.

It must be admitted that our methods of recognizing and assessing these behaviors were fairly crude and indirect during the first semester. Additional emphasis will be placed on the intellectual curiosity outcome during the Spring 1996 semester. The importance of this pursuit is highlighted by the following quote from Albert Einstein [see http://stripe.colorado.edu/~judy/einstein.html]:

> The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery every day. Never lose a holy curiosity.

SECTION 4 – CADETS IN ENGR 110 ENRICHED THEIR CONCEPTUALIZATION OF "ENGINEERING," ESPECIALLY AS A TYPE OF WORK WHICH INCLUDES SOLVING ILL-DEFINED PROBLEMS.

Data

Two types of data were collected to assess cadets' conceptualization of "engineering:"

(1) Cadets self-reported (through focus group sessions, questionnaires, and e-mail surveys) that the conceptualization of "engineering" has been enriched by the ENGR 110 course.

(2) A second indicator of the cadets' conceptualization of "engineering" was derived from concept maps created by cadets at the beginning and end of the course.

1. Self-Reports

1a. Focus Groups

The focus group sessions provided evidence that the experiences they had in ENGR 110 influenced the way cadets thought about engineering as a career choice. Example statements include:

I think just the fact that they incorporated the three different fields of engineering, that was a good idea. Coming in here I wasn't sure what I wanted to major in, and now I know what the different fields deal with, instead of jumping into them without any concept of what I'd be dealing with.

But I didn't like engineering, I had no intentions of going into engineering, but now I want to major in engineering.

During the focus group sessions students spoke of their understanding of the nature of work that engineers perform and a growing confidence in their ability to do that work. Example statements include:

My impression of this class is it's a view of engineering from a different viewpoint, be able to work in teams. I guess the thing is to solve ill-defined problems which I guess is what we're doing.

I think I've become better, cause I remember last year I took Physics and the problem was pretty straightforward, you just had an equation and you just plug in numbers, and at the beginning of the year I was trying to do that. I found out pretty soon you can't just do that all the time. You had to figure things out for yourself, and I think I learned a lot from that.

In the beginning you might just stare at it (problem), you learn as you go through it. You just start somewhere, and if its not right you go in another direction and start over and try something else and then eventually you might see a pattern. So I think it does help, and I have improved.

1b. Questionnaires

On the end of course self assessment questionnaire, 82% of students agreed (18% disagreed) with the statement, "My understanding of engineering has improved as a result of this course." Most students strongly agreed (scores of 5 and 6) while only a very few students strongly disagreed.

1c. E-Mail Surveys

Following each of the three course projects, cadets used e-mail to respond to a set of questions about the project. After the first project students were asked what they had learned from project 1. Eighteen of the 37 (49%) respondents identified learning that engineers work on complex problems which are ill-defined in nature. Following the final project, cadets were again asked what they had learned from ENGR 110. The majority of cadets (25 of 32 or 79%) mentioned in their response the recognition that engineers work on complex projects requiring problem-solving in ill-defined areas which go beyond simply applying formulas and equations.

2. Concept Maps

Prior to and at the end of the course, cadets completed concept maps of their view of engineering. A comparison group of ENGR MECH students also completed pre- and post- concept maps. The concept maps were scored for depth of structure by using the Novak and Gowin (1984) scoring system which provides a quantitative formula for counting links, cross links, hierarchical structure, and nodes. Table 2 presents the pre- and post- mean scores for cadets in ENGR 110 and ENGR Mech 120. The results are difficult to interpret because one ENGR 110 class scored very high on the initial map.

Table 2

Mean Pre- and Post- Scoring of Concept Maps for ENGR 110 and ENGR Mech 120

	110 - 2		110 - 3		Mech - 6		Mech - 7	
	mean	sd	mean	sd	mean	sd	mean	sd
Pre	46.1	19.7	38.1	12.5	39.2	16.0	39.8	11.5
Post	40.2	12.7	47.8 *	11.6	44.7	12.4	52.2 *	13.4

* Statistically significant pre- to post- differences $p < .05$.

Concept maps were also assessed for qualitative differences in the types of nodes and links used. All concepts and links included in the maps were coded into groupings of like terms. The following categories emerged: problem-solving, problem-solving processes (analyze, design, feasibility study), equations (laws, theorems), engineering terms (e.g., sheer, strain, statics), and course or field descriptors (civil, astro). However, because of problems with the administration of the concept maps (described below), additional analysis was not deemed to be worthwhile.

Discussion

In the focus group and e-mail journals, students described views of engineering consistent with the emphasis on ill-defined problems in the ENGR 110 course. We learned in the focus groups that many of these cadets arrived at USAFA with fairly well-developed views of engineering which came from interactions with relatives who are engineers and other prior experiences. If the cadets did not already recognize the nature of engineering work as ill-defined, complex, and integrated, they certainly increased this appreciation by the conclusion of the ENGR 110 course.

Unfortunately, we cannot place much faith in the concept maps that were collected before and after this course. The pre-measure was given to cadets to complete as a homework assignment over the course of the first weekend of the semester. Due to poor scheduling, the post-concept maps were included as optional tasks on the final exam. The great differences in the administration of what amounts to a less than robust instrument for assessing complex mental models meant that the pre- and post- data are suspect. Further, we have no explanation for the statistically significant differences on the pre-measure between one section of ENGR 110 and the other sections of 110 and ENGR Mech 120 (see Table 2).

Recommendations

Despite problems with the data collected during the Fall semester, concept maps seem to have great potential for representing cadets' conceptual understanding of engineering and their mental models of engineering processes. The quantitative and qualitative approaches available to analyzing concept maps are appropriate to our goal of assessing enriched conceptualizations of engineering. Our ability to realize this potential will depend on our ability to control the conditions under which the students create the concept maps. These conditions seem fragile and our results from the fall semester indicate that we failed to adequately establish the necessary conditions.

Some of the conditions we must establish are:

• Cadets need to understand how to complete a sufficient concept map, e.g., many students failed to label links.

• Instructions for completing the maps must be consistent with scoring systems, e.g., in the Novak and Gowin system hierarchical maps are rated more highly than flat maps.

• The time to complete the map needs to be constant across pre- and post- as well as treatment and control administrations of the instrument.

• Cadets need to be motivated to provide their best performance across pre- and post- as well as treatment and control administrations.

Several strategies to be followed in analyzing concept maps in the future are:

• Continue the Novak and Gowin scoring system, but look for alternative mapping and scoring strategies, including scoring systems which do not emphasize hierarchical structures.

• Further implement the content analysis by working with faculty to develop categories and criteria for analyzing concept maps.

• Ask faculty to generate concept maps which could be used as models for judging cadet maps.

• Investigate computer-based concept mapping (e.g., Pathfinder).

CONCLUSIONS

Critics of authentic assessment complain that the knowledge/skills measured via portfolios and performances prohibit easy comparisons among students and that such assessments lack generalizability to other contexts. The capacity to discriminate different achievement levels among students is especially important at USAFA where a cadet's grade point average is the basis for many highly prized promotions, honors, and privileges. In this context, the difficulties encountered in assessing students based on team projects were considerable. It is difficult to convince faculty that much knowledge is heavily contextualized and its generalizability is severely limited (Brown, Collins, & Duguid, 1989). Faculty often over-emphasize the measurement of decontextualized knowledge, assuming that the less contextualized their tests are, the more likely they are to assess generalizable knowledge and skills. Poor transfer across most parts of the curriculum and to performance contexts indicate that this assumption is very weak.

Making assessments within the ENGR 110 course as authentic as possible within the temporal and financial constraints of USAFA was another major challenge in this study. Engineering faculty have used authentic learning assignments and performance assessments for years, but rarely have they been integrated into a course as extensively as they are in ENGR 110. Ideally, assessments should replicate as faithfully as possible the conditions under which ultimate performance will occur in the hope that the challenge and motivation engendered by these assessments will be so appealing that the customary disdain students express toward the assessment process will decrease (Baker & O'Neil, 1994). This might be described as the "Holy Grail" of assessment in professional education.

In the final analysis, few will dispute the fact that the nature of educational outcomes at USAFA and many other professional schools is undergoing profound change. Clearly, there is increased attention to outcomes related to higher order thinking skills and attitudes while at the same time there is renewed interest in more authentic assessments of traditional knowledge and skills. The challenge of defining and assessing new and traditional educational outcomes for engineering education has been and continues to be enormous. We believe that the research described in this paper is an original and important response to that challenge.

REFERENCES

Baker, E. L., & O'Neil, H. F. (1994). Performance assessment and equity: A view from the USA. *Assessment in Education, 1*(1), 11-26.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-41.

Fellows, N. (1994). A window into thinking: Using student writing to understand conceptual change in science learning. *Journal of Research in Science Teaching, 31*(9), p.985-1001.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.

Kitchener, K. & King, P. (1985). *Reflective judgment scoring manual*. (Available from C. L. Lynch, Reflective Judgment Associates, HCR 75, Box 91, New Concord, KY 42076.)

Knight, M. E. (1994). *Portfolio assessment: Application of portfolio analysis*. Lanham, MD: University Press of America.

Lynch, C., Kitchener, K. & King, P. (1994). *Developing reflective judgment in the classroom: a manual for faculty*. (Available from C. L. Lynch, Reflective Judgment Associates, HCR 75, Box 91, New Concord, KY 42076.)

Mark, M. M., & Shotland, R. L. (Ed.). (1987). *Multiple methods in program evaluation*. San Francisco, CA: Jossey-Bass.

Merluzzi, T. V., Glass, C. R., & Genest, M. (Eds.). (1986). *Cognitive assessment*. New York: New York University.

Novak, J. & Gowin, D. (1984). *Learning how to learn*. New York: Cambridge University Press.

Puckett, M. B., & Black, J. K. (1994). *Authentic assessment of the young child*. New York: Macmillan.

Reeves, T. C., & Okey, J. (1996). Alternative assessment for constructivist learning environments. In B. Wilson, (Ed.). *Constructivist learning environments* (pp. 191-202). Englewood Cliffs, NJ: Educational Technology.

Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco, CA: Jossey-Bass.

TM025779

AERA April 8-12, 1996

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: New Approaches to Cognitive Assessment in Engineering Education	
Author(s): Thomas C. Reeves, James M. Laffey, & Mary R. Marlino	
Corporate Source: The University of Georgia	Publication Date: April 1996

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] **Sample sticker to be affixed to document**

Check here
Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document [ ]

or here
Permitting reproduction in other than paper copy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Thomas C. Reeves	Position: Professor
Printed Name: Thomas C. Reeves, Ph.D.	Organization: The University of Georgia
Address: Instructional Technology 607 Aderhold Hall Athens, GA 30602-7144	Telephone Number: (706) 542-3849
	Date: April 30, 1996

CUA

THE CATHOLIC UNIVERSITY OF AMERICA
Department of Education, O'Boyle Hall
Washington, DC 20064
202 319-5120

February 27, 1996

Dear AERA Presenter,

Congratulations on being a presenter at AERA[1]. The ERIC Clearinghouse on Assessment and Evaluation invites you to contribute to the ERIC database by providing us with a written copy of your presentation.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality.

Please sign the Reproduction Release Form on the back of this letter and include it with **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can drop off the copies of your paper and Reproduction Release Form at the **ERIC booth (23)** or mail to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to: AERA 1996/ERIC Acquisitions
The Catholic University of America
O'Boyle Hall, Room 210
Washington, DC 20064

This year ERIC/AE is making a **Searchable Conference Program** available on the AERA web page (http://tikkun.ed.asu.edu/aera/). Check it out!

Sincerely,

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

[1]If you are an AERA chair or discussant, please save this form for future use.