AC 2007-2435: DEVELOPING COGNITIVE AFFECTIVE BEHAVIORAL WORK SAMPLING METHODOLOGIES TO ASSESS STUDENT LEARNING OUTCOMES

Tuba Pinar Yildirim, University of Pittsburgh

Tuba Pinar Yildirim is a Ph.D. candidate in Industrial Engineering and Marketing Departments of University of Pittsburgh. Her research interests are engineering assessment, financial optimization, market segmentation and customer portfolio management. She holds an M.S. (University of Pittsburgh) and a B.S. in Industrial Engineering (Middle East Technical University). Address: 1048 Benedum Hall, University of Pittsburgh, Pittsburgh, PA 15261; telephone 412.400.8631; fax: 412.624.9831; e-mail: tpy1@pitt.edu.

Joel Townsend, University of Pittsburgh

Joel Townsend is a graduate student in Industrial Engineering Department of University of Pittsburgh. His research interests are engineering assessment, engineering ethics and public policy. He holds a B.S. in Industrial Engineering (University of Pittsburgh.

Mary Besterfield-Sacre, University of Pittsburgh

Mary Besterfield-Sacre is an Associate Professor of Industrial Engineering and the Fulton C. Noss Faculty Fellow at the University of Pittsburgh. Her research interests are in engineering education evaluation, in empirical mod-eling applications, and K12 district system improvements. In the area of assessment, Dr. Sacre has written numerous conference and journal papers and has given many workshops and pres-entations. Her research in this area has been funded by the NSF, DOE, Sloan Foundation, EiF, and the NCIIA. She is an associate editor of the Journal of Engineering Education.

Larry Shuman, University of Pittsburgh

Larry J. Shuman is Associate Dean for Academic Affairs, School of Engineering, University of Pittsburgh and Professor of Industrial Engineering. His areas of interest are improving the engineering education and the study of ethical behavior of engineers. As Associate Dean, he has introduced a many curricula innovations. He has been principle or co-principle investigator on over 20 sponsored projects funded by the NSF, HHS and DoT, the RW Johnson Foundation, and EiF. He is Editor of the new Advances in Engineering Education.

Harvey Wolfe, University of Pittsburgh

Harvey Wolfe Professor Emeritus in the Department of Industrial Engineering at the University of Pittsburgh. He is a Fellow of the Institute of Industrial Engineers (IIE) and served as President. He served as President of the Council of Industrial Engineering Academic Department Heads in 1999-2000. He is a co-author of Engineering Ethics: Balancing Cost Schedule and Risk - Lessons Learned from the Space Shuttle (Cambridge University Press, 1997).

Developing Cognitive, Affective, Behavioral Work Sampling Methodologies to Assess Student Learning Outcomes

Abstract

In this study, we develop and validate a work sampling methodology to assess processes that engineers usually engage in (i.e., working in teams, conducting design work, addressing ethical issues). To obtain in-depth measures for these process oriented student learning outcomes, 100 percent behavioral observation is typically used, but which is time consuming and expensive. Work sampling is a common industry practice used to observe physical activities, as it minimizes time to collect data, yet provides statistically similar results relative to 100% behavioral observation. In our research we have bridged the gap between common practices in industry and assessment in engineering education by extending sampling theories to the observation of intervals that can capture the cognitive, behavioral and affective domains for three student learning processes – teamwork, design, and ethical reasoning.

We designed an experiment to statistically compared 100% behavioral observation with work sampling. Four environments with two examples each were videotaped. Each tape was evaluated by two observer teams: one to conduct 100% behaviorally observation and the other to work sample. ANOVA tests were used to determine inter-rater reliability both within and between teams. Results suggest that work sampling can replace 100% behavioral observation for teamwork. Similar positive results have been obtained for design. For ethical reasoning, although a high reliability could be obtained between observers for 100% behavioral observation, work sampling was not a suitable replacement method. This paper describes the overall study, its overarching results with respect to the three outcomes investigated, and comments on various factors related to each outcome that may permit work sampling to be an effective alternative for some outcomes but not for others.

Introduction

The engineering criteria has changed the motivation of engineering education accreditation from "what are you [the program] doing?" to "*what are your students doing*?" As a result, the need for solid, in-depth measurements has become a high priority. At recent engineering education conferences (e.g. *Best Assessment Processes in Engineering Education Symposiums, ASEE, FIE*) the number of evolving approaches for evaluating engineering programs, as well as methodologies for measuring various student outcomes is growing more rich. Yet, several troublesome issues still remain. First, most of these "assessment" methods had not been fully evaluated. Second, many focus on final products via performance appraisals particular to the outcome(s) using rubrics as the assessment tool. Third, many engineering administrators still voiced concerns about the costs associated with organizing, implementing and maintaining an effective assessment program, given limited resources of time, people (i.e. raters), and money.

Assessing non-sequential outcomes in engineering such as working in teams, development of designs or overcoming ethical dilemmas often require a methodological tool to examine behavior at various levels of the cognitive and affective domains (e.g. analysis, synthesis, evaluation, and valuation). While such a tool has been needed for professional work, the recent movement towards outcomes-based assessment in engineering education has highlighted the need for valid measures for similar student outcomes.

Unfortunately, until such a measurement tool is available, we typically must rely on common examinations (i.e. P.E. exam), surveys, or 100 percent observations. The former two are not fully capable of assessing higher levels of the cognitive and affective domains and can only examine the outcome at a single point in time. Rather, in-depth assessment methods, such as behavioral observation [1-4] are desirable because they enable us to investigate outcomes "in action" and evaluate the individuals' ability to function in the higher level learning domains; unfortunately, this assessment method requires considerable time and resources. To be an effective method the evaluator must: determine the educational parameters involved and the time period to be observed, conduct 100 percent observation of the environment and/or record it on audio/video tape, and then transcribe the observations prior to analysis.

Addressing these concerns, we developed the *Cognitive, Affective, Behavioral Sampling* (CABS) methodology[5] that utilizes work sampling to replace 100 percent observation. The use of work sampling seems reasonable as it has been successfully adopted in measuring physical activities in the work place [6-8]. Specifically, through CABS, we developed and validated work sampling as an alternative, cost effective evaluation tool that takes advantage of probability theory to "sample" the observable environment significantly reducing the time and costs necessary for the observation without the loss of quality information. By deriving an interval that statistically is as accurate as 100 percent observation, we were able to measure cognitive and behavioral work efficiently, much the same as we currently measure physical work. Only limited education-based [9] using related work sampling methods for 100 percent observation have occurred to date.

Our main intention in developing the CABS methodology was to bridge the gap between common practices in industry and assessment in engineering education and professional practice by extending sampling theories to the observation of intervals that can capture the cognitive, behavioral and affective domains. Though our work is focused primarily on engineering, and particularly engineering education, the derived precepts can work in multiple education related fields.

This paper provides an overview of our work in developing the CABS methodology for three process-oriented student learning outcomes – teamwork, design, and ethical reasoning. We first provide a short overview of the approach we have used to determine if work sampling results are statistically similar to behavioral observation. We then provide a summary of how the work sampling methodology could serve as a replacement for 100% behavioral observation for some of the outcomes. A discussion about the potential reasons for how work sampling may be used for some outcomes and not others is then provided.

Methodology

Our study, funded through a grant with the Department of Education's Fund for the Improvement of Post Secondary Education (FIPSE) program, was to develop a statistically-valid work sampling methodology for the observation of teamwork, design and ethical reasoning. The testing of CABS was accomplished through a large experimental design. Videotaped data of engineering students in team oriented project-based learning environments was used. These environments included: a single lab project, two short intensive projects, and a senior capstone design project, all from undergraduate engineering courses. A special workroom was created that facilitated the video taping of all activity that occurred in the room by a fixed light-activated camera; whereby all members of the particular team, who volunteered for the study, were directed to use for all project work. A description of these environments is given in Table 1.

Table 1. Teamwork Environments in the CABS Project

Project Name	Environment	Description
T-shirt Design	Single Lab	Focuses on the design of an anthropometrically developed T-shirt. Two teams of three people each worked on the project for 2-5 hrs.
Delta Design	Short Project	Delta Design is a game in which students collectively design a habitat suitable for residents of the imaginary planet Delta. Two teams of four people each worked on the project for 4-5 hrs.
Development of a Regression	Short Project	Development of a regression model for a particular subject of interest. The project was completed as part of a course. Two teams of four people each worked through semester.
Product Realization	Capstone Project	Working in multidisciplinary design teams from engineering and business, students take a product from concept to business plan. In doing this, they address issues of market analysis, design, manufacturing design, and production planning. Two teams of five people worked on the project for a semester.

Two student teams per project were taped. The tapes were then experimentally observed with two sets of raters observing each team. For each project one set of raters 100%-observed one student team and then work sampled the other team to validate the previous observation. Each rater set included at least two raters to ensure inter-rater reliability, which was tested at α=0.05 for teamwork (and at a lowest level of $\alpha = 0.1$ for engineering design and ethical reasoning). Raters, both 100% behavioral observation and work sampling, viewed the same tapes using the same protocols for observation. Specifically, both sets of raters watched the tape and indicated which attribute (to be discussed) was being conducted by the subject. As the subject conducted their activities (i.e. related to the outcome), the observers recorded the time when the subject change to a new attribute. Note, only one attribute could occur at any point in time. For the work sampling raters, the tape was forwarded to a pre-specified time where the rater observed the subject until a decision could be render about the observed attribute. For our research, several intervals were tested (10 second, 20 second and floating). Floating intervals worked best for the raters resulting in observations that took less than 10 seconds. In terms of time, 100% behavioral observation took more than twice the time of the actual tape (i.e. raters often stopped the tape to rewind and re-review); whereas work sampling took less time than it takes to view the entire tape.

The attributes or characteristics of the process that the observers recorded through their observation for teamwork and design outcomes were established by a comprehensive literature review and input from experts in the field. For the ethical reasoning outcome, a rubric (based on literature and expert opinion) was used [10]. Through pilot studies, attributes for each outcome studied were established based on the literature and/or rubric. Eight teamwork attributes (see Table 2) were developed are based on McGourty's work on teamwork [11].

Table 2. Teamwork Attributes and Definitions

Teamwork Attribute	Description
Working Together	Working together to accomplish goal/project; participating in the development of ideas; all good forms of communication.
Disruption/Non productive Activities	Distractions by a member that prevent team from accomplishing goal; team member leaves meeting for unknown reasons; does not actively participate in team project development (i.e., sleeps during meetings, under the influence of drugs and/or alcohol, ignores team members, being idle, staring off into space).
Coming to Conclusions	Uses ideas/comments to reach intermediate or final agreement; point when team or member decides on action and moves on – can later come back and change action, which then becomes another conclusion.
Reporting Results of Independent (Subgroup) Work	Point where subgroup member discusses to other team members what they found; discussing/showing results of independent/subgroup work.
Team Management	Things that are done to keep the team working well together; discourages side conversations and/or getting off track during discussions; uses meeting time efficiently; doing these activities described above, but not being the object of these.
Working Individually	Performing calculations by hand or using calculator; working on computer; reading/writing; thinking/staring at work.
Researcher Cannot Tell	Cannot see team member (back is to camera or member is completely out of view), cannot hear member to place action in category.
Other	Outside distractions.

Upon several iterations testing different attributes found in literature, the engineering design attributes (see Table 3) were adopted from Mehalik's recent meta-analysis of over 40 journal articles of empirical studies characterizing design [12]. These were modified with help from the author.

Table 3. Design Attributes and Definitions

Engineering Design Attribute	Description
Project Management	Leading the group in an activity. Such as: Planning and organizing (timeframes, meeting points, etc.), Choosing a method of approach, Explaining the problem solving strategy to the group, Deciding on the problem solving strategy, Decomposing a large problem.
Research	Any basic research activity of a single person related to the design. Such as: Performing research using a computer, journal articles, published materials, etc., Discussing known, preexisting designs, Literature review, Mathematical investigation.
Evaluation of Current Design	Any single/ group activity accomplished in order to evaluate and critically look at how a *complete* design performs. Such as: Designing tests based on the constraints, Testing if constraints are satisfied, Discussing test results, Performing calculations.
Review	Any group activity related to rethinking the problem and the constraints on the problem. Such as: Evaluating the big picture of the problem, Discussing the problem constraints (the basic idea, how they can be changed, etc.), Discussing/ talking about the problem without talking about the design
Working Alone	A single person's real thinking and innovation phase, where all the learning and referencing is done and actual designing start. Such as: Thinking individually about the problem, Brainstorming, Narrowing down alternatives, Actual designing, change on the current design, or innovation.
Group	The phase where there is a simultaneous work of the group on the project, or a discussion

Engineering Design Attribute	Description
Discussion of Current Design	related to it. Such as: Brainstorming, Suggesting new ideas, Actual designing, change on the current design or innovation, Discussing alternative solutions, Discussing feasibility, infeasibility, practicality of ideas, Engaging in a conversation related to the project.
Critical Decision Making	The phase in which there is critical decision making by the group members. Such as: Debating/ choosing between viable, feasible options, Selecting a design to be implemented, Deciding if project is complete, Debating the validity of claims, Deciding on the standard.
Commodity Work	Thinking related to the project is complete, but the project needs to be implemented by a commodity work. Such as: Building a model from a blueprint, Acquiring the supplies, Anything related to the final submission of the project, Labor intensive non-thinking activity.
Negative Impact	Non-progressive activities. Such as: Taking a break, leaving the table, Using cell phones, Talking off task, Eating.
Waiting	Non-progressive activity of the subject where s/he is idle, but at the same time not distracting anyone.
Cannot Tell	Researcher cannot tell what the subject is doing, or there is a technical problem that makes it impossible to accurately observe the subject.

Finally the observable attributes of ethical reasoning were adopted from P-MEAR rubric [13-14], as shown in Table 4. As with design, the authors of the rubric helped to modify the attributes to use in behavioral observation.

Table 4. Ethical Decision Making Attributes

Attribute	Basic Definition
Recognition of Dilemma	Subject recognizes one of the key ethical dilemmas. This task is a single subject attribute (i.e. only reflective on the original speaker of the statement and not upon the rest of the group)
Information	Subject is reading or speaking of material that is already present in the documentation given. This does not include any analysis into the case study. This can be a single subject or multiple subject task. (i.e. it is reflective on other subjects if they are listening or actively participating in the conversation).
Analysis	Subject is analyzing the facts in terms of how they relate to the ethical issue. This is a multiple subject task.
Perspective	Subject brings a new perspective in to the discussion. This will mainly pertain to outside examples that may seem relevant in understanding the case. This is a single subject task.
Resolution	Subject is talking about group's overall conclusion or analysis. This attribute is only referenced to the subject speaking of the resolution and not of those listening. This can be a single subject or multiple subject task.
Negative Impact/Not On Task	Subject is showing actions that have a negative impact on the project. This category can be single or multiple subject task.
Waiting	Subject is waiting (but not negatively impacting) on another member to perform some task. This is a single subject attribute.
Do Not Know	Subject may not be in visual view or not heard via audio, or the viewer is completely unsure of the subject's categorical status.

All the attributes form a set of observable, mutually exclusive and exhaustive categories that describe the various (observable) states of the process. While deriving these attributes, our researchers had to modify some of the definitions, specify the behaviors that correspond to the attributes and provide examples. Often, such as in ethical decision making, we also identified visual and audio cues that can help to identify the behaviors correctly and facilitate the training process of new observers.

The ratings obtained from work sampling were compared to the 100% behavioral observations results. To accomplish this, analysis of variance was employed using an α level of 0.05 for teamwork; and an α level of 0.1 for engineering design. Unfortunately, for ethical reasoning viable rater reliability for work sampling could not be established; and hence, no statistical comparisons could be made between work sampling and behavioral observation. Table 5 provides an example of teamwork in which the two work sampling observer ratings were compared to the targeted 100% behavioral observation. This target and confidence interval was established by averaging the statistically similar behavioral observations of two raters (independent of the work sampling raters). The two work sampling rater observations were then compared to the target and confidence interval. For this particular example, one work sampling observation did not fall into the desired target range (highlighted in gray). For a full in-depth description of the methodology see Besterfield-Sacre et. al.[15].

Table 5. Work Sampling Results for T-Shirt Design Environment

T-Shirt Design: Group 1, Subject B						
		Observer 3	**Observer 4**	**Target**	**Lower Limit**	**Upper Limit**
	Observations	46	45			
Teamwork Attributes	1	0.72	0.59	0.70	0.56	0.83
	2	0.13	0.00	0.06	0.00	0.13
	3	0.04	0.00	0.02	0.00	0.06
	4	0.00	0.06	0.02	0.00	0.06
	5	0.02	0.00	0.04	0.00	0.09
	6	0.09	0.30	0.07	0.00	0.15
	7	0.00	0.05	0.05	0.00	0.11
	8	0.00	0.00	0.05	0.00	0.11

Feasibility of Work Sampling on Various Cognitive Outcomes

Our results indicate that the use of work sampling to replace behavioral observation is successful for two of the three outcomes evaluated. Teamwork attributes were observable with little confusion as they could be easily discriminated through audio or visual differences. Further, training of raters required less than 10 hours; and consistency between and across observer pairs were maintained regardless of the project environment. Besterfield-Sacre et al. provide more information on teamwork results of CABS project [16]. A positive aspect about observing teamwork is that this type of outcome, as we have measured through our attributes, appears to be independent of the nature of the task. From our particular research, observations of teamwork activities do not appear to be influenced by the type of assignment.

Although engineering design could be work sampled, the procedures leading to observation were more complicated compared to teamwork. First, during the development of visual and audio attribute "cues", more revisions were necessary compared to teamwork. Second, training of the raters took 3 to 6 weeks (30 to 60 hours), considerably more time than for teamwork. One reason for the difficulty in training was related to the nature of the environment observed (i.e. raters needed to be familiar with the design problem and the processes that students may use to solve the problem). Finally, our experiments were not able to prove a statistical similarity between work sampling and 100% behavioral observation for ethical decision making. Although, rater reliability was easily achieved for 100% behavioral observation, statistical rater reliability could not be achieved for work sampling. For this outcome, we concluded that the observation personnel needed to be experts in ethics in order to correctly categorize interactions among subjects who are discussing ethical dilemmas and responsibilities.

Throughout the research the authors noted three factors or areas that may influence the feasibility of work sampling for the particular outcomes: the nature of the outcome related to the degree of cognition, attitudes and behavioral, the nature of the experiment (i.e. the projects, subjects tested, problem solving approaches), and the approach to analyses (i.e. number and type of observers, how much training was involved, and the definitions of the attributes), as shown in Table 6.

Table 6. Factors that Potentially Affect the Feasibility of Work Sampling

Type	Factor	Teamwork	Design	Ethics
Success of Work Sampling?		Successful approval of hypothesis	Successful approval of hypothesis after substanial training	Failure to prove
Nature of the Outcome	**Degree of Cognition**	Behavioral in nature	Both behavioral and cognitive in nature	Cognition, behavioral and attitudinal in nature
Nature of the Experiment	**Tapes/ Environments/ Projects**	Works for any environment: Delta design, Regression, T-shirt Design, Product Realization,	Works for limited environments: Delta design, T-shirt Design, Product Realization, HF Hand-tool (tested)	Works for environments focused on ethics
	Subjects Tested	Degrees: ME/ EE/ IE/ BIOE; Sophomores, juniors, seniors, not educated about the project before experiment, mixed in gender	IE/ BIOE, Sophomores, juniors, seniors, not educated about the project before experiment, mixed in gender	IE, CE, BIOE Juniors, seniors, half of the subjects educated in engineering ethics, mixed in gender
	Process of Student Problem Solving	Not applicable	Product oriented	Resolution; not necessarily solution
Nature of the Analysis	**Training/days to learn**	2 weeks or 3 half hour sessions	1 month, roughly 6 half hour sessions	2-3 weeks, roughly two 45 minute sessions
	Embeddedness of Definition/ Cues	Primarily visual; some audio cues	Varying visual and audio cues	Varying visual and audio cues
	Observers	15 students: graduate and undergraduate level, mixed in gender	11 students: graduate and undergraduate level, mixed in gender	3 students: graduate and undergraduate level, mixed in gender

As table 6 shows, the success of work sampling fluctuated due to the different variables present in each of the three outcomes. The nature of the outcome range from primarily behavioral in nature (Teamwork), to cognitive and behavioral (Design), to cognitive, behavioral and attitudinal domains (Ethics). For the teamwork outcome, the raters indicated that the process was easily observable. This is possibly due to the fact that teamwork attributes could be easily discriminated by the use of audio and visual cues. As a result, teamwork displayed high inter-rater reliability for both 100% behavioral observation and work sampling. Design and ethics did not achieve such inter-rater reliability possibly due to the nature of the outcome which involves discussion about engineering concepts, models, etc. (cognition). This discussion required raters to interpret audio cues from the subjects that potentially increased observer variation (and hence lower rater reliability). This invariably led to longer training periods of the raters. With regards to ethical reasoning, observation also included interpretation of subjects' attitudes about the case, which often got "heated" among subject participants. This is one reason raters had a difficult time work sampling the outcome. Context about the current discussion required understanding cognitive aspects of the problem and the problem solving process, but this also had to be done in the context of subjects' attitudes about the dilemma.

The processes used to "solve" the different types of projects greatly differed. As mentioned, as we defined through the attributes, observation of teamwork abilities was primarily concerned with the interactions a subject had with the other group members. In design, observations were further coupled with the design process. As a result, observers were observing not only the interacting behaviors but also the information and knowledge exchanged between subjects. For example, the categories of "Evaluation of the Current Design" and "Critical Decision Making" go beyond interaction with other group members, but are important parts involved in understanding the design process. Similarly to design, making ethical decisions required information or knowledge interaction between subjects but was further complicated with attitudinal perspectives. Here, "Recognition of Dilemma", "Analysis", "Perspective" and "Resolution" attributes were often marred with subjects' attitudes about the dilemma. Given that some groups observed had a course in engineering ethics allowed this to be teased out as they had a better understanding of the important aspects of the ethics case and hence were less "emotional" about the ethical decision making process.

In regards to the differences among observers, similar conditions existed for all three outcomes. The ratio of male observers to females was close to one, and observers were mostly undergraduate students directed by a graduate student. Subjects involved in the experiments also ranged in their level in the program, in majors and also in whether they had education related to the topic (i.e. engineering ethics training).

Conclusions and Future Research

This research has focused on the feasibility of work sampling to replace 100% behavioral observation for three process oriented outcomes: teamwork, engineering design, ethical decision making. Through our experimentation and as defined by our attributes, work sampling can be readily used in place of 100% behavioral observation for teamwork; and with proper training of the design problem and observation, it can be used for engineering design. However, at this time, work sampling does not appear to be a viable alternative for ethical reasoning.

The implications of this study are two-fold. First, the research has provided a cost effective alternative to assessing student capabilities associated with teamwork and design. This methodology offers educators an opportunity to measure extensive and time consuming student activities using fewer resources. Second, it is possible to extend this approach to industrial environments providing human resource analysts and managers a different performance evaluation mechanism. The framework presented may be helpful in analyzing the relationships between team members allowing engineering managers to better understand how teams spend their project time. Further, this approach potentially can be used to show which attributes a team of engineers may spend more time than intended and how they traverse to other attributes. An extension of this study suggests that a Markov chain can be used to analyze how subjects transition between the various attributes [17].

References

1. Lincoln, Y. S. and E. G. Guba, *Naturalistic Inquiry*. Newbury Park, CA, SAGE Publications, 1985.
2. Miles, M. B. and A. M. Huberman, *Qualitative Data Analysis*. Beverly Hills, Sage Publications, 1984.
3. Bucciarelli, L. L. *Designing Engineers*. The MIT Press, Cambridge, MA, 1994.
4. Brereton, M.F., et al. *An Exploration of Engineering Learning*. in *ASME - Design Theory and Methodology*. 1993.
5. Besterfield-Sacre, M., E. Newcome, L. Shuman, and H. Wolfe, "Extending Work Sampling to Behavioral and Cognitive Concepts," *Industrial Engineering Research Conference,* Houston, TX, May 16 – 18, 2004 (CD-ROM - 6 pgs.).
6. Aft, L., *Work Measurement and Methods Improvement*, John Wiley & Sons, Inc., New York, 2000.
7. Wolfe, H. and JP Young, "Staffing the Nursing Unit: Part I – Controlled Variable Staffing," *Nursing Research*, Summer 1965a.
8. Wolfe, H. and JP Young, "Staffing the Nursing Unit: Part II – Multiple Assignment Technique," *Nursing Research,* Fall 1965b.
9. Hendricks, G.C. et. al., 1974, "Effects of Behavioral Self Observation on Elementary Teachers and Students," Technical Report No. 121, Stanford Center for Research and Development in Teaching, Stanford University, CA.
10. Shuman, LJ, ME Besterfield-Sacre and BM Olds, "Ethics Assessment Rubrics," in Mitcham, C., L. Arnhart, D. Johnson and R. Spiers, editors, the *Encyclopedia of Science, Technology, and Ethics*, Macmilliam, Spring 2005.
11. M.E. Besterfield- Sacre, L.J. Shuman, H. Wolfe, C.J. Atman, J. McGourty, R.L. Miller, B.M. Olds, and G.M. Rogers, Defining the Outcomes: A Framework for EC 2000, IEEE Transactions on Engineering Education, 43(2), pp. 100-110, (2000).
12. Mehalik, Matthew M., and Schunn, Christian (2006). "What Constitutes Good Design? A Review of Empirical Studies of Design Processes," International Journal of Engineering Education, 22 (3), pp. 519 – 532.
13. Shuman, L., M. Besterfield-Sacre, H. Wolfe, M. Sindelar, R. Pinkus, B. Olds, and R. Miller, "Using Rubrics to Assess Students' Ability to Resolve Ethical Dilemmas," *Industrial Engineering Research Conference*, Atlanta, GA, May 15 – 17, 2005 (CD ROM – 6 pgs.).
14. Shuman, LJ, ME Besterfield-Sacre and BM Olds, "Ethics Assessment Rubrics," to be published in the Mitcham, C., L. Arnhart, D. Johnson and R. Spiers, editors, the *Encyclopedia of Science, Technology, and Ethics*, Macmilliam, Spring 2005.
15. Besterfield-Sacre, M., L.J. Shuman, H. Wolfe, R.M. Clark, and T.P. Yildirim, "Development of a Work Sampling Methodology for behavioral Observations: Application to Teamwork," submitted to the *Journal of Engineering Education*, February 2007.
16. Reference 15.
17. Yildirim, T.P., Besterfield- Sacre, M., Shuman, L., Clark, R., Wolfe, H. and Bailey, M. (2007). "A New Perspective for Analysis of Social Processes: Applying Markov Chains to Teamwork Processes", submitted to the *Journal of Engineering Education*, March 2007.