

## **AC 2007-2837: ASSESSING PARTICIPANT ENGAGEMENT IN A MIDDLE SCHOOL OUTREACH PROGRAM**

### **Jessica Matson, Tennessee Technological University**

Jessica Matson is a Professor of Industrial Engineering at Tennessee Technological University. She received her BS from Mississippi State University and her MS and PhD from the Georgia Institute of Technology, all in industrial engineering. She has previously served on the faculty of Mississippi State University and the University of Alabama and as industrial engineering department chair at Tennessee Tech. She is a registered PE.

### **Kristine Craven, Tennessee Technological University**

Kris Craven received her BS, MS, and PhD in Mechanical Engineering from West Virginia University. She coordinated the Freshman Engineering program at WVU before joining the Basic Engineering faculty at Tennessee Technological University. She has taught Programming, Graphics, Thermodynamics, Dynamics, and Fluid Mechanics. Her interests include numerical modeling in heat transfer and fluid flow, and educational concerns.

### **Sally Pardue, Tennessee Technological University**

Sally Pardue is an Associate Professor of Mechanical Engineering at Tennessee Technological University. She began her academic appointment in August 1999 following four years as a Research and Development Engineer. Dr. Pardue received her PhD in Engineering from Tennessee Technological University in 1995. Her research interests are random vibrations, machine diagnostics, cavitation, and NDE of composite materials.

### **Corinne Darvennes, Tennessee Technological University**

Corinne Darvennes is a Professor of Mechanical Engineering at Tennessee Technological University and interim chair of the Industrial and Systems Engineering Department. She received her BS from the Université de Technologie de Compiègne, her MS from the Institute of Sound and Vibration Research, and her PhD from the University of Texas at Austin. She is interested in research in acoustics and ultrasonic NDE, and outreach activities.

### **Alison Wachs, Tennessee Technological University**

Alison Wachs, a senior industrial engineering major at Tennessee Technological University, will graduate in May 2007. She currently serves in leadership roles for the Institute of Industrial Engineers, Alpha Pi Mu, and Tau Beta Pi. She is a recipient of the Dwight D. Gardner Scholarship from IIE.

# **Assessing Participant Engagement in a Middle School Outreach Program**

## **Abstract**

Engineering A Future (EAF), an annual outreach program for girls in grades five through eight, has been held at Tennessee Technological University since 2003. The program is aimed at introducing these young women to careers in engineering. Sponsored by Tennessee Tech, the Society of Women Engineers, the American Association of University Women, the Tennessee Space Grant Consortium, Girl Scouts of Cumberland Valley, and local industry, participation has grown each year. In the day-long program, each group of eight to twelve middle-school girls rotates through four 45-minute, hands-on, engineering-focused activities. High school and college students serve as volunteer guides, leading their assigned group from one activity to the next.

In response to participant evaluations and adult volunteer suggestions, program changes have been made each year. Prior to 2006, there had been no formal attempt to recruit the help of high school students. However, with the first EAF participants reaching high school age, it was time to engage them in the day activities.

To provide a meaningful experience for the volunteer guides in EAF 2006, they were asked to serve as observers to study participant engagement. This paper provides an overview of the findings from the engagement study, including the impact of time of day on participant engagement, characteristics of activities with the highest engagement, correlation with end-of-day participant evaluations, and lessons learned.

## **Introduction**

Engineering A Future (EAF), a day-long outreach program for girls in grades five through eight, has been held annually at Tennessee Technological University since 2003. Sponsored by Tennessee Tech, the Society of Women Engineers (SWE), the American Association of University Women (AAUW), the Tennessee Space Grant Consortium, Girls Scouts of Cumberland Valley, and Cummins Filtration, participation has grown each year, from 73 girls in 2003 to approximately 250 in 2006. The program features an initial teambuilding activity followed by four 45-minute sessions, each providing a hands-on activity focused on a different engineering topic. Lunch and an engineering-related lunchtime activity are also part of the program. The girls are divided into groups of eight to twelve participants. Activity sessions are led by volunteers, who represent engineering faculty, SWE members, or engineers from local industry. Some activities are designed to accommodate fifth- and sixth-grade participants; others are designed for seventh- and eighth-graders. High school and college students, which include both engineering and education majors, serve as volunteer guides, leading their assigned group from one activity to the next. With College of Education faculty and AAUW members assisting in the registration process and numerous other miscellaneous tasks, this program incorporates a unique connection between local school educators, University students and faculty, and middle

school students. The annual EAF effort is seen as a way to foster cooperation between the education and engineering communities.

Each year EAF program changes have been made in response to the prior year participant evaluations and adult volunteer suggestions. However, because the EAF theme and activities have changed annually, there has been little feedback to activity leaders on the girls' level of interest in the activity. In addition, prior to 2006, there had been no attempt to recruit the help of high school students. However, with the first EAF participants reaching high school age, it was time to engage them in the day activities.

At EAF 2006, the theme was "Lights, Camera, Action: Behind the Hollywood Scene." Thirty activity sessions representing twenty-seven unique activities were scheduled concurrently to accommodate approximately 250 participants. The hands-on activities included King Kong Construction Company, Mythbusters, Jurassic Park Mud, Acrobatic Animation, and Earthquakes: Rock-a-bye Village, among others.

Participants were divided somewhat randomly into twenty groups of fifth- and sixth-graders and ten groups of seventh- and eighth-graders. An effort was made to ensure that girls from the same schools were assigned to different groups. Each group was given a Hollywood-related name such as Producer or Make-up Artist and assigned to one of four schedules. The four schedules were designed to sequence lunchtimes to ease crowding in the university cafeteria. Groups with older participants were assigned to schedules with later lunchtimes. Approximately forty high school and college students served as volunteer guides. Some local high school and middle school educators who participate in a three-year Math and Science Partnership grant were also involved in 2006. Some of them were paired with engineers to design and present activities; others chose to serve as guides.

Prior to 2006, the engagement level of the volunteer guides had been inconsistent. Some guides would participate fully with the middle school students while others would sit in a corner during the activity sessions, not participating at all other than to guide their group from one activity to the next. To provide a meaningful experience for the volunteer guides in EAF 2006, they were asked to serve as observers to study participant engagement. The following sections describe the engagement study and its significant findings.

## **Engagement Study Design**

Engagement Research. Research findings on engagement and learning guided the development of the engagement survey/observation form. However, there are conflicting studies and findings on the best way to measure engagement and what engagement means. In this study, the term *engagement* describes the participants' level of interest and level of active involvement in EAF sessions.

Because the EAF activity involved middle-school girls in engineering activities, the survey form development considered findings from engagement research in two domains: engineering education<sup>2,4</sup> and K-12 education<sup>1,3</sup>. Research on methods to increase engagement in engineering education has been summarized by Smith, et al.<sup>4</sup> Methods include active and collaborative (or

cooperative) learning and student-faculty interaction. The EAF form was designed to capture the degree of active learning but not student-leader interaction. In an engineering education engagement study, Harris and Cox<sup>2</sup> reported on an observation system to measure instructional differences in engineering classrooms. The observation form developed for EAF included some aspects similar to the observation system of Harris and Cox but focused specifically on student activity rather than leader activity.

Fredricks, Blumenfield, and Paris's<sup>1</sup> research on K-12 school engagement defined three types of engagement: emotional, behavioral, and cognitive. The observation form developed for EAF focused on the emotional and behavioral aspects of engagement and not on the cognitive aspects. Lutz, Guthrie, and Davis<sup>3</sup> reported the development of a method for assessing third- and fourth-grade student reading engagement. They defined emotional engagement as student reaction based on physical cues; the EAF survey form provided an opportunity to assess emotional engagement with a question on overall student interest level. Behavioral engagement was defined by Lutz, Guthrie, and Davis as student participation level; the EAF survey form reflected behavioral engagement through categories for hands-on activity, listening to leader, asking questions, or unrelated activity. Cognitive engagement, or degree of learning, was not assessed on the EAF form.

Another factor influencing the development of the EAF survey form was the plan to use high school students as observers. Because the observers were not knowledgeable researchers and would have only minimal training, it was important to keep the form as simple as possible.

Survey Form and Observer Orientation. A two-page survey observation form was developed, along with an instruction sheet. Prior to the opening session on the morning of EAF, each observer received a folder with the survey form for the day's activities and the instruction sheet, which included a completed example form. A brief explanation was given, and observers had time to ask questions. The orientation session was conducted several times as the volunteer observers arrived at the check-in location. Figure 1 shows a section of the data collection form and some typical results. Also recorded on the form, although not shown in Figure 1, were the group name and the observer's name.

Observer Assignment. Observers were randomly assigned to a participant group based on their arrival times. Although the goal was to assign two observers for each group of girls, needs in other program areas caused reassignments, resulting in fewer than two observers per group. Of approximately 45 observer folders distributed, only 32 were completed and returned, and only 30 of those were usable. Despite having fewer observers than planned, all activity sessions and 27 of the 30 participant groups had at least one observation.

Data Collection. After their orientation, observers went to meet their participant groups. Beginning with the teambuilding activity at the opening session, each observer recorded what her assigned participants were doing, e.g., hands-on activity, listening, asking questions, or unrelated activities, at each five-minute interval during the activity session. The observers had been instructed to record the primary activity or activities taking place at the five-minute interval and to leave the interval blank if they missed an observation time. For example, a session might begin with a brief explanation by the activity leader. After five minutes, if one girl was asking



session. The overall engagement rating was determined by scoring the level of interest ratings with High = 3, Medium = 2, and Low = 1. The average overall engagement rating by the observers was 2.53 out of a maximum of 3. The tallies showed a total of 822 questions asked with 44% of the girls asking a question. Of the activity observations recorded at five-minute intervals, 59% of the observation opportunities for all participants were categorized as hands-on activities. In recording when the hands-on activity occurred, the survey forms indicated that the average time was 10.4 minutes after the beginning of the session.

Table 1. Observer Data Summary Results

Characteristic	EAF 2006 (Opportunities)	Observer Total (Recorded)
Activity Sessions	30 different sessions, each repeated 4 times, 120 total opportunities	30 different sessions 105 total session observations
Participant Groups Observed	30 groups assigned	27 groups
Activity Engagement Score	Maximum = High = 3	Observer Assessment Average = 2.53
Activity Total Questions	Not Applicable	822 recorded in tallies
% of Girls Asking Questions	0-100%, range	44% recorded in tallies
% of Hands-on Observations	0-100%, range	59% recorded
First recorded Hands-on Observation	5 minutes, minimum	10.4 minutes (average)

Both Excel and Minitab were used for statistical analysis of the data. Based on EAF experience from prior years, anecdotal comments from volunteers in EAF 2006, and outreach research literature, a number of hypotheses were tested, as listed in Table 2 and explained in the following subsections.

The Lunch Effect. Activity leaders from previous years had lamented the lunch effect, or the seemingly lower level of engagement of participants in the afternoon following lunch. Thus, the first test was to compare the activity ratings for participant groups having a later lunch schedule with ratings for groups with an earlier schedule. Fourteen groups were scheduled to complete three activities before lunch so that lunch was delayed until either 12:30 or 1:00p.m. The remaining sixteen groups were scheduled for lunch at 11:30a.m. or noon with the third activity following lunch. As shown in the hypothesis test results in Table 3, observers of groups who had their third activity scheduled before lunch rated their engagement for that activity lower than did observers whose groups had lunch before their third activity. However, at a 0.05 level of significance, the difference is not statistically significant, and it must be concluded that there is no difference in the engagement ratings.

Table 2. Hypothesis Tests

H <sub>O1</sub> :	Observer engagement ratings for activity 3 will be equal for those groups who completed the activity before lunch and those who completed the activity after lunch.
H <sub>A1</sub> :	Observer engagement ratings for activity 3 will be lower for those groups who completed the activity before lunch.
H <sub>O2</sub> :	Observer engagement ratings for activities 3 and 4 will be equal to ratings for activities 1 and 2.
H <sub>A2</sub> :	Observer engagement ratings for activities 3 and 4 will be lower than ratings for activities 1 and 2.
H <sub>O3</sub> :	Observer engagement ratings will not vary for activities with higher versus lower percentages of hands-on activity.
H <sub>A3</sub> :	Observer engagement ratings will be higher for activities with higher percentages of hands-on activity.
H <sub>O4</sub> :	Observer engagement ratings will not vary for activities with higher versus lower percentages of participants asking questions.
H <sub>A4</sub> :	Observer engagement ratings will be higher for activities with higher percentages of participants asking questions.
H <sub>O5</sub> :	Observer engagement ratings will not vary for activities with earlier versus later observations of hands-on activity.
H <sub>A5</sub> :	Observer engagement ratings will be higher for activities with earlier observations of hands-on activity.
H <sub>O6</sub> :	Observer engagement ratings on the survey forms will not differ from end-of-day evaluations by participants.
H <sub>A6</sub> :	Observer engagement ratings on the survey forms will differ from end-of-day evaluations by participants.

Table 3. Activity 3 Lunch Test

t-Test: Two-Sample Assuming Equal Variances		
<i>Activity 3</i>	<i>After Lunch</i>	<i>Before Lunch</i>
Mean	2.529	2.364
Variance	0.265	0.455
Observations	17	11
Pooled Variance	0.338	
Hypothesized Mean Difference	0	
df	26	
t Stat	0.737	
P(T<=t) one-tail	0.234	
t Critical one-tail	1.706	
P(T<=t) two-tail	0.468	
t Critical two-tail	2.056	

Another way to test the lunch effect is by comparing the observer ratings of engagement for activities 3 and 4 with the observer ratings of engagement for activities 1 and 2. If the perceived lunch effect reflects fatigue, one would expect lower levels of engagement in the later activities. Table 4 shows that the result of this hypothesis test is similar to the first test. There are no significant differences in the observer engagement ratings for earlier versus later activities. In fact, the later activities have a higher mean score and lower variance for the engagement ratings. With the random assignment of girls to participant groups, this result may reflect greater levels of comfort with new acquaintances as the day progressed.

Table 4. Activities 1 and 2 Versus Activities 3 and 4 Test

t-Test: Two-Sample Assuming Equal Variances		
	<i>Activities 1 &amp; 2</i>	<i>Activities 3 &amp; 4</i>
Mean	2.492	2.560
Variance	0.358	0.292
Observations	59	50
Pooled Variance	0.328	
Hypothesized Mean Difference	0	
df	107	
t Stat	-0.622	
P(T<=t) one-tail	0.268	
t Critical one-tail	1.659	
P(T<=t) two-tail	0.535	
t Critical two-tail	1.982	

Significant Factors Related to Engagement. The next three hypotheses in Table 2 are related to questions about activity characteristics that are associated with perceived engagement. Numerous research findings point to hands-on activity as encouraging engagement. Observer data resulted in two types of information related to hands-on activity. Recording the number of girls participating in hands-on activity at five-minute intervals allowed each activity to be assessed for the percentage of “hands-on” content versus listening or other activity. In addition, the first observation of hands-on activity was available from the survey forms. Presumably, activities with earlier hands-on participation might be considered more engaging.

Stepwise regression analysis was used to assess the relevance of various factors as predictors of engagement rating. Twelve factors were included for consideration, as shown in Table 5. Figure 2 provides the results from Minitab. Of the twelve potential factors, only four were found statistically significant and brought into the regression model. Those were (1) the percentage of hands-on activity, (2) the participant group, (3) the activity number (i.e., whether it was the first, second, third, or fourth activity of the day, and (4) the group schedule (i.e., whether it was an early lunch or late lunch schedule, corresponding somewhat to grade level of participants).

Table 5. Factors Considered in Stepwise Regression

Factors	Significant
Group	X
Number of Girls in Group	
Schedule Assignment	X
Observer	
Activity Number in Schedule (1, 2, 3, or 4)	X
Activity	
Percentage of Activity Judged “Hands-on”	X
Interval at Which Hands-on Activity Was First Recorded	
Number of Questions Asked	
Percentage of Girls Asking Questions	
Activity Start Time	
Rating for Teambuilding Activity	

Response is Rating on 12 predictors, with N = 104				
Step	1	2	3	4
Constant	2.171	2.380	2.202	2.422
HO%	0.61	0.56	0.52	0.49
T-Value	2.78	2.58	2.41	2.27
P-Value	0.007	0.011	0.018	0.026
Group Name		-0.0126	-0.0129	-0.0140
T-Value		-1.84	-1.91	-2.07
P-Value		0.069	0.059	0.041
Activity No.			0.089	0.086
T-Value			1.78	1.74
P-Value			0.078	0.086
Group Category				-0.081
T-Value				-1.53
P-Value				0.128
S	0.556	0.550	0.544	0.540
R-Sq	7.03	10.03	12.80	14.83
R-Sq(adj)	6.11	8.25	10.19	11.39
Mallows C-p	3.5	2.1	1.0	0.8
PRESS	32.7358	32.2766	31.8633	31.8460
R-Sq(pred)	3.47	4.83	6.05	6.10

Figure 2. Stepwise Regression Results from Minitab

Even with the four factors selected for their statistical significance, the statistics in Figure 2 indicate that the resulting regression model would not be a good predictor of engagement rating. The best model would probably be one with two factors and Mallows C-p value of 2.1; this model includes the percentage of hands-on activity and the participant group (group name). However, the adjusted  $R^2$  for the model is only 8.25%, which means that almost 92% of variation is unexplained by the two factors. “Best fit” regression modeling with Minitab supported this conclusion.

Correlation with End-of-Day Evaluations. The last hypothesis in Table 3 concerns how well the observers’ engagement ratings reflected the true level of engagement of the participants. One measure of engagement that reflected the middle school girls’ opinions was captured in an end-of-day evaluation of the activities in which the girls participated. The end-of-day evaluations were scored similarly to the observer ratings. For a given activity, a “Great” rating was scored as 3, “Good” as 2, “OK” as 1, and “Poor” as 0. There was a correlation of 0.623 between the girls’ average end-of-day ratings of activities and the observers’ average ratings (with High = 3, Medium = 2, and Low = 1) for those same activities. A paired sample test comparing the end-of-day ratings with their corresponding observer ratings of engagement is shown in Table 6. Although a moderate degree of correlation exists, the hypothesis test shows a statistically significant difference in the results between the observers and the participants.

Table 6. End of Day Evaluations Compared with Observer Ratings of Engagement

t-Test: Paired Two Sample for Means		
	<i>End of Day Evaluation</i>	<i>Observer Rating</i>
Mean	2.240	2.532
Variance	0.160	0.168
Observations	27	27
Pearson Correlation	0.623	
Hypothesized Mean Difference	0	
df	26	
t Stat	-4.320	
P(T<=t) one-tail	0.000	
t Critical one-tail	1.706	
P(T<=t) two-tail	0.000	
t Critical two-tail	2.056	

This test result does not mean that the observers' assessments were incorrect, however. It may be that the observers were actually more accurate than the end-of-day evaluations in assessing the participants' engagement. The observers recorded their assessment at the time the activities occurred whereas the participants completed the evaluation after a significant amount of time had passed for some activities. Also, the participants were asked to rate how well they liked the activities rather than to rate how well they were engaged. The results indicate that further study is warranted.

### Lessons Learned

The primary "lesson learned" related to the conduct of the engagement study was that a brief orientation on the hectic morning of EAF was not a sufficient training period. To improve the data collection effort, the observers need practice time with the form and a calmer introduction to their task. For EAF 2007, the two-week period prior to EAF will be used to conduct training sessions for high school observers.

The form itself can also be improved to make it easier for the observers to complete. Many observers had difficulty with the "Asking Questions" observation category on the form, and the form would be simpler without this column. Adding an activity end-time to the form is recommended because it was not clear whether blank observations at the end of the activity observation period indicated that the observation times were missed or that the activity had ended. Removing the transition interaction observations is also recommended; there was little contribution to results from these data. Because research findings point to the importance of student-leader interaction and one-on-one interaction, another variable that might be added to the form is the number of leaders and activity assistants. A category for type of activity, e.g., problem-solving or demonstration, would also help in assessing the factors that promote engagement.

In addition, the large number of variables (30 observers, 30 activity sessions, 30 participant groups, and 4 schedules, among other factors) introduced a large amount of variability into the statistical analysis. A more focused study that controls some of the variability might produce

stronger results. EAF 2007, which will have only 2 schedules and 15 activity sessions, will permit a more focused study.

Finally, a “lesson learned” based on the results of the observers’ ratings of engagement is that intuitive results need to be tested. Although activity leaders believed that the participants were not as engaged after lunch, the observers saw no decline in the engagement level.

## **Conclusion**

This paper has described an effort to engage high school and college volunteers in a middle school outreach program by assigning them as observers to assess the engagement of the middle school participants. Although the results do not prove that the engagement ratings are valid, the following findings were significant.

- The “lunch effect” that was expected to result in decreased engagement in the afternoon was not observed by the volunteers.
- The most significant factor related to engagement rating was the percentage of hands-on activity within a session. Factors that were not significant included the number of questions asked and the percentage of girls asking questions. Although questions may indicate engagement in some situations, questions can also indicate confusion and frustration in others.
- Although there was a moderate correlation between the observers’ engagement ratings and end-of-day evaluations by the participants, these averages were found to be significantly different. Which statistic represents a better measure of engagement is unclear without further study.

A major goal in the study was to engage the high school and college volunteers themselves to get them more invested in the EAF activity. In their end-of-day evaluations, 73.9% of these volunteers indicated that they would consider engineering as a major, and 98.6% indicated that they would volunteer for EAF again. Based on these results, the goal of volunteer engagement was met with great success.

## **Bibliography**

1. Fredricks, Jennifer A., Phyllis C. Blumenfield, and Alison H. Paris. (2004) “School Engagement: Potential of the Concept, State of the Evidence,” *Review of Educational Research*, 74(1), 59-109.
2. Harris, Alene H. and Monica Farmer Cox. (2003) “Developing an Observation System to Capture Instructional Differences in Engineering Classrooms,” *Journal of Engineering Education*, 92(4), 329-336.
3. Lutz, Susan L., John T. Guthrie, and Marcia H. Davis. (2006) “Scaffolding for Engagement in Elementary School Reading Instruction,” *The Journal of Educational Research*, 100(1), 3-20.
4. Smith, Karl A., Sheri D. Sheppard, David W. Johnson, and Roger T. Johnson. (2005) “Pedagogies of Engagement: Classroom-Based Practices,” *Journal of Engineering Education*, 94(1), 87-101.