Incorporating Engineering in Middle School Science

Miss Gabriella J Ducamp, University of Virginia

Gabriella J. Ducamp is a former elementary school teacher and current STEM Education PhD student at the University of Virginia.

Crystal Jean DeJaegher, University of Virginia

Incorporating Engineering In Middle School Science: A Pilot Study Teaching About Electricity

Gabriella J. Ducamp and Crystal J. DeJaegher

University of Virginia

Abstract

The overarching goals of this study are to introduce engineering concepts to middle school students through digital fabrication, and increase science competency while stimulating interest in STEM careers. This pilot study incorporates digital fabrication, engineering design, and visualizations into a comprehensive unit that integrates hardware, software, curricula, and a collaborative space. The engaging nature of these activities may improve student attitudes toward STEM disciplines and increase the likelihood that students will take advanced STEM coursework and choose STEM careers, as they feel the success of completing hands-on activities supported by visualizations in content that might otherwise be overwhelming. Fabrication activities are a vehicle that may help deliver science content about electricity in a group setting that is approachable for students who may be less likely to pursue STEM activities. Three eighth-grade classes (n=57) participated in hands-on science lessons that incorporate technology to deepen their understanding about electricity using modeling, simulations, and measurement probes. The researcher triangulated data, including student interviews, analysis of student work samples, pre-/post tests, and the STEM Semantics Survey. Data may help provide practical ways to enhance interest and skills necessary to encourage participation in STEM activities, courses, and careers. This paper focuses on student attitudes, as measured by the STEM Semantics Survey. The paired samples t-test failed to reveal a statistically significant difference between the mean overall pretest and overall posttest totals.

In order to meet growing demands for a competitive Science, Technology, Engineering, and Math (STEM) workforce, education must adapt to reflect the skills necessary to be successful in these fields and students must be encouraged to maintain interest in these disciplines. Digital fabrication offers the opportunity to bring activities that are more like STEM professions to students than traditional classroom instruction. Similarly, the engaging nature of these activities may improve student attitudes toward STEM disciplines and increase the likelihood that students will take advanced STEM coursework and choose STEM careers. Digital fabrication involves creating physical objects from a digital design. Though digital fabrication has been a mainstay of professional engineers for decades, the next generation of 2D and 3D personal digital fabricators makes digital fabrication in schools feasible for the first time (Bull & Groves, 2009). This can facilitate the introduction of engineering design in science classes, as required by the proposed Next Generation Science Standards (National Research Council, 2011), through rapid prototyping. The ease of creating, making, and revising ideas can encourage an iterative process that is an important part of engineering design. Fabrication activities are a vehicle that may help deliver science content about electricity in a group setting that is approachable for students who are less likely to pursue STEM activities. Data gathered from preliminary activities using digital fabrication may help provide practical ways to enhance interest and skills necessary to encourage participation in STEM activities, courses, and careers. This study may help the researcher better understand instructional practices that might be used to improve the use of technology in educational settings.

Method

Participants

Three eighth grade classes participated in hands-on science lessons taught by their teacher about electricity. Sixty-two students were enrolled in the classes. However, one student was absent throughout, and four either missed the pre- or post-test administration of the STEM Semantics Survey, so their data was not analyzed. All participating students (n=57) were 13-14 years old. Classes were indirectly tracked by ability, because advanced and remedial Language Arts and Math classes impact enrollment in these science classes. Therefore, section 1 tended to have advanced students (n=23), while sections 2 (n=17) and 3 (n=17) had balanced and remedial populations, respectively. The classes were at a suburban school in a small Mid-Atlantic city. 33 males and 24 females participated.

Research Design

Students participated in a science unit on electricity, taught by their normal instructor, and incorporating technology to deepen their understanding through modeling, simulations, and measurement probes. Before beginning the unit, students completed the STEM Semantics Survey, as a baseline measure of their attitudes toward STEM content and careers. The day after the unit was completed, students took the STEM Semantics Survey again, to measure changes of their attitudes toward STEM content and careers. During the unit, the teacher graded student work normally, based on participation and completed assignments. A different topic was addressed each day (static, current, circuits, motors & generators, and electromagnetism). Each class began with a pre-test. After a brief lesson from the teacher, students completed hands-on activities, such as making Play-Doh circuits or building speakers. Visualizations supported content by either showing videos to activate and deepen prior knowledge or show molecular processes that students cannot observe. Students recorded their processes using Cornell Notes daily, and read and answered questions in worksheet packets. The following day, students turned in their work and took a post-test matched to the prior day's content. Gains from pre-to post-test indicate gained content knowledge. Throughout the week, students were interviewed, as well, to get a deeper response to the similarities and differences between typical science class and this unit.

Hypotheses:

Ho1: μ pretest = μ posttest

An independent-sample t-test will measure whether there is a significant difference between pre- and post-test STEM Semantics Survey Scores. The researcher hypothesizes that attitudes will have changed as a result of the intervention. Therefore, the null hypothesis states that there is no difference between the pre-and post-test scores on the STEM Semantics Survey.

Ho2: μ male= μ female

Gender is a known factoring STEM career paths. The researcher predicts that gender will affect changes in attitude toward STEM content and careers. Therefore, the null hypothesis states that the gender will not affect score on the STEM Semantics Survey for science, math, technology, engineering, career, or overall. Another way of stating this is: Ho1: aj=0 for j =science, math, technology, engineering, career, or overall.

Ho3: μ class 1= μ class 2 = μ class 3

This means that the class section does not affect score on the STEM Semantics Survey for science, math, technology, engineering, career, or overall. Another way of stating this is Ho3: Bk = 0 for k=science, math, technology, engineering, career, or overall

Measure

The STEM Semantics Survey identifies the STEM content and career interests of elementary and middle school students. It is unique in its young target audience and focus on the broad STEM area, rather than specific disciplines or all careers (Tyler-Wood, Knezek, & Christendsen, 2009). The STEM Semantics Survey is a twenty-five item, one-page multiple-choice instrument that takes only a few minutes to complete. Students respond to adjective pairs, choosing how they feel about an object on a Likert-type scale. There are five pairs of words for each of five categories: science, math, engineering, technology, and a career in STEM. Reliability ranges from "respectable" to "excellent" using Cronbach's Alpha, and construct validity is affirmed since test items correlate with their intended construct (Tyler-Wood, Knezek, & Christendsen, 2009).

Procedure

Students participated in a typical science class unit on electricity, taught by their normal instructor, and incorporating technology to deepen their understanding through modeling, simulations, and measurement probes. The teacher graded the class as normal, based on participation and assignment completion. Before the unit began and after the unit concluded, students took the STEM Semantics Survey to measure possible changes in attitude toward STEM content and careers in relation to the implementation. In order to "clean up" the data before analysis, the values of the survey were made consistent. In order to encourage students to reflect on each pair in the STEM Semantics Survey, some values are switched. For example, a 7 might be a very positive reflection of science in one question (ex. "fascinating"), but a very negative one (ex. "unappealing") in the next item. Therefore, all of the values were first converted so that bad = 1, and good = 7. Next, the five responses within each category were consolidated, so that the sum of all five adjective pair scores from each category (science, technology, engineering, math, and STEM careers) could total up to 25 points be recorded. A sum of scores from each category contributed to the overall score, a measure of whether the intervention broadly impacted student attitudes toward STEM content and careers. A repeated-measures design was considered, but with only two data points per participant, this design is less than ideal. Therefore, the total from the pre-test was subtracted from the post-test total for each category to determine gains as a result of the intervention.

First, a paired samples t-test will determine whether there are significant differences between pre- and post-test scores for each participant. The paired samples t-test compares the means of two variables by computing the difference between the two variables for each participant and measuring whether the average difference is significantly different from zero. This test is commonly used, but has limitations because it can only compare differences between two groups and can only examine effects of one independent variable on one dependent variable. It also only indicates whether there is a significant difference between the two groups, with no indication as to where those differences lie. Since there are more multiple independent variables and dependent variables in this intervention, further analyses are planned. The Analysis of Variance (ANOVA) can test hypotheses that the t-test cannot.

Next, a Two-Factor Randomized-Block ANOVA will be run to assess the impact of gender and class section on changes in science, technology, engineering, math, career, and overall attitude. Females are often less likely to complete STEM programs (Hill, Corbett, Rose, 2010), so the primary independent variable/treatment variable of this study is gender. However, a nuisance variable was noted when the study began. The class section that students are in serves as a confounding variable, hence the randomized blocks. Students at the test site are tracked into advanced, normal, and remedial math and language arts classes. Due to this scheduling, their science performance may be indirectly affected. Thus, an interaction of class and gender is considered in the analysis. This blocking serves to reduce residual variation and increase power. The dependent variables are the change from pre-to post-test on the STEM Semantic Survey for science, technology, engineering, math, career, and overall categories.

The Two-Factor-Randomized Block ANOVA has five assumptions: independence, normality, homogeneity of variance, compound symmetry, and additivity.

1. Independence is the assumption that one data point does not influence another. In this case, students completed the STEM Semantics Survey on their own, in a quiet room, so there is no indication that the attitudes of one student would affect another in this study. Therefore, independence has not been violated throughout the study.
2. Normality refers to a probability distribution of the data that is known to have properties such as being symmetrical (no skew) and kurtosis (clustering near the tails of the distribution) of zero. Examining normal Q-Q plots gives an indication of normality. Freidman's ANOVA is a non-parametric test of whether more than two related groups differ that could compensate for non-normality. However, there are only two treatment groups in this study, so it cannot be performed if the assumption of normality is violated.

3. Homogeneity of variance is the assumption that the variance of a given variable is stable at all levels of another variable. A significant result of Levene's test indicates that the groups are significantly different. Unequal sample sizes (as is the case in this study) increase the chance of incorrectly rejecting the null hypothesis. If the assumption of homogeneity of variance is violated, a non-parametric test like the Kruskall-Wallis test may be used to transform the data.
4. Compound symmetry is a condition that holds true when both the variances across conditions are equal and the covariances between conditions are also equal. Since the ANOVA isn't robust to this violation, if compound symmetry is violated, levels of A can be limited, F tests can be adjusted, or a multivariate ANOVA can be used. It is most practical to use a procedure such as the Geyser Greenhouse or Hyunh & Feldt to accommodate for violation of compound symmetry/sphercity.
5. Tukey's test of additivity indicates whether there is an interaction term or residual additivity. If there is an interaction term, power goes down since the type II error for factor A increases.
6. In addition to checking the above assumptions, Multiple Comparison Procedures may be used. The MCPs selected depend on whether spericity has been met. However, they do not apply to the research questions in this project, so they are not run.

I hope to learn whether student attitudes improved toward STEM subjects and careers overall, whether females' STEM attitudes are similar to males', and whether attitudes of students at different academic tracks improved as a result of the intervention.

Results

The paired samples t-test failed to reveal a statistically significant difference between the mean overall pretest (M=98.74 and s=9.90) and overall posttest (M=100.42 and s= 10.88) totals on the STEM Semantics Survey, t(56)=-1.343, p=.185, α=.05. Therefore, we fail to reject the null hypothesis. There is insufficient evidence to support a change in overall attitude toward STEM content and careers as a result of the intervention. However, other factors such as gender, academic track, and changes in attitude toward specific disciplines will be examined.

T-Test

[DataSet1] J:\private\BlockDesignData.sav

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pretotal	98.7368	57	9.90405	1.31182
	Posttotal	100.4211	57	10.87748	1.44076

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pretotal & Posttotal	57	.588	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Pretotal - Posttotal	-1.68421	9.46640	1.25386	-4.19599	.82756	-1.343	56	.185

The two-factor randomized-block ANOVA will be run six times; once each for changes in attitude in science, technology, engineering, mathematics, STEM career, and overall.

Science: In an attempt to measure whether the intervention affected attitude (from pre- to post-test) as measured using the STEM Semantics Survey, I compared the gender and a nuisance variable of gender*class to gains in science. The results of the two-factor randomized-block ANOVA reveal no significant impact of gender on science gains: F(1,51) = 1.674 , p= .202.

The results of the two-factor randomized-block ANOVA reveal no significant impact of the nuisance variable gender*class on science gains: $F(4,51) = .310$, $p = .870$. This means that the null hypotheses for gender and class are not rejected. I continue to explore whether any assumptions have been violated.

Tests of Between-Subjects Effects

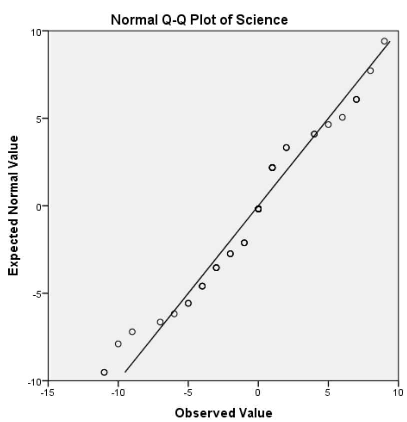
Dependent Variable: Science

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	60.571[a]	5	12.114	.619	.686	.057	3.097	.208
Intercept	9.052	1	9.052	.463	.499	.009	.463	.102
Gender * Group	24.222	4	6.055	.310	.870	.024	1.238	.114
Gender	32.745	1	32.745	1.674	.202	.032	1.674	.246
Error	997.464	51	19.558					
Total	1076.000	57						
Corrected Total	1058.035	56						

a. R Squared = .057 (Adjusted R Squared = -.035)

b. Computed using alpha = .05

1. **Independence** is not violated, as it is unlikely that student answer choices affected other students, given the surveys were administered to individual students in a quiet room.
2. **Normality** is not violated, as examined on the Q-Q plot. Expected and observed points are linear.



3. **Homogeneity of Variance** is violated, as tested by Levene's Test. In this case, .005<.05, revealing a significant difference between groups. Therefore, a non-parametric test could

be performed, such as the Kruskall-Wallis. However, the Kruskall-Wallis only functions for one-way ANOVAs. Since this study has a nuisance variable, it can't be run.

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Science

F	df1	df2	Sig.
3.892	5	51	.005

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Gender * Group + Gender

4. **Compound symmetry** There is no option to test compound symmetry in SPSS, unless there is a repeated measures design. Since this is not a repeated measures design, it is not able to be assessed.
5. **Additivity** Tukey's Test of Additivity reveals significant additivity. Therefore there, there is no interaction effect. $F(1,55) = 16.381$, $p = .000$

ANOVA with Tukey's Test for Nonadditivity

			Sum of Squares	df	Mean Square	F	Sig
Between People			28.158	56	.503		
Within People	Between Items		6.395	1	6.395	14.264	.000
	Residual	Nonadditivity	5.761[a]	1	5.761	16.381	.000
		Balance	19.344	55	.352		
		Total	25.105	56	.448		
	Total		31.500	57	.553		
Total			59.658	113	.528		

Grand Mean = 1.6579

a. Tukey's estimate of power to which observations must be raised to achieve additivity = -2.166.

Mathematics: In an attempt to measure whether the intervention affected attitude (from pre- to post-test) as measured using the STEM Semantics Survey, I compared the gender and a nuisance variable of gender*class to gains in math. The results of the two-factor randomized-block ANOVA reveal no significant impact of gender on science gains: $F(1,51) = .257$, $p = .615$. The results of the two-factor randomized-block ANOVA reveal no significant impact of the nuisance variable gender*class on science gains: $F(4,51) = 1.947$, $p= .117$. This means that the null hypotheses for gender and class are not rejected. I continue to explore whether any assumptions have been violated.

1. **Independence** is not violated, as it is unlikely that student answer choices affected other students, given the surveys were administered to individual students in a quiet room.
2. **Normality** is not violated, as examined on the Q-Q plot. Expected and observed points are linear.
3. **Homogeneity of Variance** is not violated, as tested by Levene's Test. In this case, .736>.05, revealing no significant difference between groups.
4. **Compound symmetry** There is no option to test compound symmetry in SPSS, unless there is a repeated measures design. Since this is not a repeated measures design, it is not able to be assessed.
5. **Additivity** Tukey's Test of Additivity reveals significant additivity. Therefore there, there is no interaction effect. $F(1,55) = 16.381$, $p= .000$

Engineering: In an attempt to measure whether the intervention affected attitude (from pre- to post-test) as measured using the STEM Semantics Survey, I compared the gender and a nuisance variable of gender*class to gains in engineering. The results of the two-factor randomized-block ANOVA reveal no significant impact of gender on science gains: $F(1,51) = .995$, $p= .323$. The results of the two-factor randomized-block ANOVA reveal no significant impact of the nuisance variable gender*class on science gains: $F(4,51) = 2.296$, $p= .072$. This means that the null hypotheses for gender and class are not rejected. I continue to explore whether any assumptions have been violated.

1. **Independence** is not violated, as it is unlikely that student answer choices affected other students, given the surveys were administered to individual students in a quiet room.
2. **Normality** is not violated, as examined on the Q-Q plot. Expected and observed points are linear.
3. **Homogeneity of Variance** is violated, as tested by Levene's Test. In this case, $.671>.05$, revealing no significant difference between groups.
4. **Compound symmetry** There is no option to test compound symmetry in SPSS, unless there is a repeated measures design. Since this is not a repeated measures design, it is not able to be assessed.
5. **Additivity** Tukey's Test of Additivity reveals significant additivity. Therefore there, there is no interaction effect. $F(1,55) = 16.381$, $p= .000$

Technology: In an attempt to measure whether the intervention affected attitude (from pre- to post-test) as measured using the STEM Semantics Survey, I compared the gender and a nuisance variable of gender*class to gains in tech. The results of the two-factor randomized-block ANOVA reveal no significant impact of gender on science gains: $F(1,51) = .112$, $p= .740$. The results of the two-factor randomized-block ANOVA reveal no significant impact of the nuisance variable gender*class on science gains: $F(4,51) = 2.255$, $p= .076$. This means that the null hypotheses for gender and class are not rejected. I continue to explore whether any assumptions have been violated.

1. **Independence** is not violated, as it is unlikely that student answer choices affected other students, given the surveys were administered to individual students in a quiet room.
2. **Normality** is not violated, as examined on the Q-Q plot. Expected and observed points are linear.

3. **Homogeneity of Variance** is not violated, as tested by Levene's Test. In this case, .910<.05, revealing no significant difference between groups.
4. **Compound symmetry** There is no option to test compound symmetry in SPSS, unless there is a repeated measures design. Since this is not a repeated measures design, it is not able to be assessed.
5. **Additivity** Tukey's Test of Additivity reveals significant additivity. Therefore there, there is no interaction effect. $F(1,55) = 16.381$, $p= .000$

Career: In an attempt to measure whether the intervention affected attitude (from pre- to post-test) as measured using the STEM Semantics Survey, I compared the gender and a nuisance variable of gender*class to gains in STEM career interest. The results of the two-factor randomized-block ANOVA reveal no significant impact of gender on science gains: $F(1,51) = .099$, $p= .755$. The results of the two-factor randomized-block ANOVA reveal no significant impact of the nuisance variable gender*class on science gains: $F(4,51) = .562$, $p= .692$ This means that the null hypotheses for gender and class are not rejected. I continue to explore whether any assumptions have been violated.

1. **Independence** is not violated, as it is unlikely that student answer choices affected other students, given the surveys were administered to individual students in a quiet room.
2. **Normality** is not violated, as examined on the Q-Q plot. Expected and observed points are linear.
3. **Homogeneity of Variance** is not violated, as tested by Levene's Test. In this case, .507>.05, revealing no significant difference between groups.

4. **Compound symmetry** There is no option to test compound symmetry in SPSS, unless there is a repeated measures design. Since this is not a repeated measures design, it is not able to be assessed.
5. **Additivity** Tukey's Test of Additivity reveals significant additivity. Therefore there, there is no interaction effect. $F(1,55) = 16.381$, $p= .000$

Overall: In an attempt to measure whether the intervention affected attitude (from pre- to post-test) as measured using the STEM Semantics Survey, I compared the gender and a nuisance variable of gender*class to gains overall. The results of the two-factor randomized-block ANOVA reveal no significant impact of gender on science gains: $F(1,51) = .370$, $p= .545$. The results of the two-factor randomized-block ANOVA reveal no significant impact of the nuisance variable gender*class on science gains: $F(4,51) = .993$, $p= .420$. This means that the null hypotheses for gender and class are not rejected. I continue to explore whether any assumptions have been violated.

1. **Independence** is not violated, as it is unlikely that student answer choices affected other students, given the surveys were administered to individual students in a quiet room.
2. **Normality** is not violated, as examined on the Q-Q plot. Expected and observed points are linear.
3. **Homogeneity of Variance** is not violated, as tested by Levene's Test. In this case, $.390>.05$, revealing no significant difference between groups.
4. **Compound symmetry** There is no option to test compound symmetry in SPSS, unless there is a repeated measures design. Since this is not a repeated measures design, it is not able to be assessed.

5. **Additivity** Tukey's Test of Additivity reveals significant additivity. Therefore there, there is no interaction effect. $F(1,55) = 16.381$, $p= .000$

Discussion

Limitations The intervention implemented in this study took place over one week, and was administered in one class subject only. Attitudes toward several disciplines take time and experience to develop, and this study may not have been delivered long enough to make a significant difference. Since both gender and class served as independent variables (creating six groups), the sample size for this study may not have been large enough to produce statistical power. The post-test was administered the Monday following the intervention, which may have affected student attitude.

Implications & Conclusions

Participating in a one-week science intervention about electricity, incorporating digitally-fabricated objects, visualizations, and hands-on activities, does not significantly affect attitude toward science, mathematics, engineering, technology, or careers in STEM fields. Gender and the nuisance factor of class section do not significantly affect attitude toward science, mathematics, engineering, technology, or careers in STEM fields. Further research into student content knowledge and interviews may reveal factors that made this intervention successful or unsuccessful.

References

Hill, C., Corbett, C., & St Rose, A. (2010). *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. American Association of University Women. 1111 Sixteenth Street NW, Washington, DC 20036.

National Research Council (NRC). 2011. A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: National Academies Press.