Paper ID #20187

Elementary Engineering Student Interests and Attitudes: A Comparison across Treatments

Dr. Cathy P. Lachapelle, Museum of Science

Cathy Lachapelle leads the EiE team responsible for assessment and evaluation of our curricula. This includes the design and field-testing of assessment instruments and research on how children use EiE materials. Cathy is particularly interested in how collaborative interaction and scaffolded experiences with disciplinary practices help children learn science, math, and engineering. Her work on other STEM education research projects includes the national Women's Experiences in College Engineering (WECE) study. Cathy received her S.B. in cognitive science from the Massachusetts Institute of Technology and her Ph.D. in educational psychology from Stanford University.

Dr. Christine M. Cunningham, Museum of Science

Dr. Christine Cunningham is an educational researcher who works to make engineering and science more relevant, accessible, and understandable, especially for underserved and underrepresented populations. A vice president at the Museum of Science, Boston since 2003, she founded and directs Engineering is Elementary™, a groundbreaking project that integrates engineering concepts into elementary curriculum and teacher professional development. As of September 2016, EiE has served 12.6 million children nationwide and 118,000 educators. Cunningham has previously served as director of engineering education research at the Tufts University Center for Engineering Educational Outreach, where her work focused on integrating engineering with science, technology, and math in professional development for K-12 teachers. She also directed the Women's Experiences in College Engineering (WECE) project, the first national, longitudinal, large-scale study of the factors that support young women pursuing engineering degrees. Cunningham is a Fellow of the American Society for Engineering Education and was awarded the 2014 International Society for Design and Development in Education Prize. She holds B.A. and M.A. degrees in biology from Yale and a Ph.D. in Science Education from Cornell University.

Elementary Engineering Student Interests and Attitudes: A Comparison across Treatments (RTP)

Abstract

Students were asked to rate their interest and attitudes after participating in a school-year engineering curriculum unit, as compared to their remembered interest and attitudes of the summer before. The survey was developed from earlier research by the authors, as well as from surveys of science attitudes reported in the literature (Germann, 1988; Lichtenstein et al., 2008; Owen et al., 2008; Weinburgh & Steele, 2000). We report on the development and testing of the survey. We also report on our methods of analysis and results, looking at differences between the treatment groups as well as the moderating effect of student demographics.

Introduction

For decades, engineers and policymakers have expressed concern that too few students enter the engineering pipeline. This has led to many efforts to engage students in engineering in after school programs, summer programs, and more recently, in school curricula. The expectation is that, by engaging students in engineering, more will become aware of it as a possible career option, and some students will find a special affinity to engineering and pursue it. The hope often expressed is that such interventions will particularly increase the supply of new young engineers who are from demographics currently underrepresented in engineering, including women and minorities.

Given this goal, it is important to develop instruments capable of measuring change in student attitudes toward and interest in engineering. One important link in the validity argument that increasing student exposure to engineering will increase and diversify the population pursuing engineering careers, is that students become more interested in engineering and their attitudes more positive as they engage in engineering out-of-school-time experiences and curricula. To measure this assumption, we have developed an Engineering Interest and Attitudes (EIA) survey, drawing from earlier surveys used to measure student interest in and attitudes toward science.

Purpose of the Study

In this study, we examine survey responses from more than 10,000 students to examine the effects of treatment type and student demographics on student interests in and attitudes towards engineering.

Our research questions are as follows:

- To what extent do our engineering curriculum treatments increase children's interest in and attitudes towards STEM careers?
- Is there a difference in EIA outcomes between the two treatments?
- How do student demographic differences moderate the outcomes of implementation?

Method

Instrument Development

Our target population for assessment is elementary students. We searched the literature for instruments addressing interest in and attitudes towards engineering and science that were designed for elementary and middle school populations. We chose to draw from five instruments to develop our survey: one addressing attitudes towards engineering in elementary school, and the other four addressing attitudes towards science in elementary and middle school.

We reviewed the revised version of the Scientific Attitude Inventory (SAI-II), which was designed for use with middle and high school students (Lichtenstein et al., 2008). From this instrument, we chose items from the "I want to be a scientist" scale (Cronbach's α=.810) to test with our elementary students. We chose to test several items with the simplest phrasing from the Attitude toward Science in School Assessment (ATSSA), which was designed for use with high school students (Germann, 1988). From the modified Attitudes Toward Science Inventory (mATSI), developed for measuring changes in the attitudes of urban 5th grade students (ages 10 and 11) during an intervention, we pulled most of the 25 items for testing (Weinburgh & Steele, 2000). The Simpson-Troost Attitude Questionnaire (STAQ) was originally developed for use with high school students, and was reevaluated and shortened more recently using confirmatory factor analysis (Owen et al., 2008)—we pulled a number of the simpler items from this assessment as well, to repurpose for assessing elementary students' attitudes towards engineering. Finally, we incorporated most of the items from an earlier version of an engineering attitudes survey developed by the authors (2010). We revised all questions from the science attitude instruments to refer instead to engineering.

We chose to implement the survey as a post-only Likert-scale survey where students were asked to answer each question twice: once to the prompt "Last summer, I would have said:" and also to the prompt "Now I would say." We chose to implement the survey in this way knowing that the students in our study were likely to know little to nothing about engineering before engaging in the curriculum, and we had learned from prior experience with our earlier attitudes survey that children's responses before engineering instruction were likely to be highly unreliable. By asking about "before" and "now" after engineering instruction, we hoped to get more reliable data about students' prior views, through their retrospective assessment.

Pilot Testing the Instrument

To test the items we had gathered for suitability with an elementary population, we conducted interviews with 15 students in the target grade range (grades 3-5, ages 8-11), all from classrooms that had implemented engineering curricula. For each interview, we had students read each question aloud and talk about it. We asked them to explain any confusing aspects of the question, and to talk aloud about what they were thinking as they chose answers from the Likert scales. Based on these interviews, we dropped about half of the candidate questions, and revised the wording of others.

We then pilot tested the instrument with a convenience sample of 123 students from grades 3-5 classrooms that had implemented engineering thematic units. We examined the internal consistency reliability of items, and conducted an Exploratory Factor Analysis (EFA) to examine the relationships between items. Based on this analysis, we dropped several items that performed particularly poorly.

Data Collection

As part of a large-scale efficacy study of an elementary engineering curriculum, we collected post-surveys of students' interests in and attitudes towards engineering. Over two years, we collected surveys from almost 11,000 students in grades 3, 4, and 5. Students spanned a wide range of demographic groups, from rural, urban, and suburban areas of several geographically non-contiguous states. See Table 1 for the demographic breakdown of the sample.

Table 1: Student demographic breakdown of sample.

	%Male	%Undrp[1]	%FRL[2]	%EL[3]	%Gr3[4]	%Gr4	%Gr5	Total N
Comparison	50.3	37.1	46.3	6.5	23.8	31.4	44.7	5994
Treatment	51.4	31.7	43.8	5.8	32.3	36.7	31.0	4912
Total	50.8	34.6	45.1	6.2	27.6	33.8	38.5	10906
Final Sample after drops (due to incomplete surveys):								
Comparison	49.7	35.4	45.4	6.3	22.6	31.9	45.6	5385
Treatment	51.4	30.2	42.6	5.5	31.9	36.8	31.3	4417
Total	50.5	33.1	44.1	5.9	26.8	34.1	39.1	9802

[1]Percentage of students from underrepresented minority groups (Black, Hispanic, Mixed, Other).
[2]Percentage of students receiving Free or Reduced-Price Lunch.
[3]Percentage of students classified as English Learners.
[4]Percentage of students in Grade 3.

Students participated in one of two engineering curricula, the treatment and comparison. We recruited teacher volunteers for this study through their principals and superintendents. Teachers applied to participate as teams of 2-4 teachers from the same school. Only teachers from schools that had not implemented engineering curricula were accepted. Once accepted, we randomized at the school level into the treatment or comparison group.

The treatment curriculum is designed from a social constructivist theoretical framework, taking as given that students learn deeply through meaningful engagement in a discipline at a developmentally appropriate level (Sawyer, 2006). The treatment curriculum meets the criteria for project-based learning (PBL), where students focus on a main design challenge which engages them with key ideas in science and engineering. In the treatment, the central project is open-ended, students are engaged in the problem with a realistic context, and heavy scaffolding is provided to engage in engineering practices and reasoning. The comparison curriculum also includes hands-on challenges, but in contrast, the challenges are not motivated with a context, scaffolding is not provided, many challenges are not open-ended, and information is given through direct instruction.

Internal Reliability and Exploratory Factor Analysis

All analysis was conducted in SPSS 23 (IBM Corporation, 2012). Before beginning analysis, we split our sample in half randomly so that we could conduct EFA and Confirmatory Factor Analysis (CFA) on separate samples. With the first half, we used Parallel Analysis to estimate the number of factors; then the Principal Axis Factoring (PAF) method of Exploratory Factor Analysis (EFA) because we expected our Likert-scale data to be non-normal. The ratio of sample size (5390) to expected factors (<10) is quite high (539:1) so we expect that the sample size is sufficient for this procedure, even if extracted commonalities are low (MacCallum, Widaman, Zhang, & Hong, 1999). We used parallel analysis to determine the number of factors to retain. We used the Oblimin rotation with the PAF because we expect the resulting factors to be correlated to some extent. We then examined pattern matrices for item loadings, crossloadings, and internal consistency reliability (by calculating Cronbach's alpha) to determine the best model fit and suitability of scales.

Confirmatory Factor Analysis

Using the second random half from our sample, we conducted CFA. The CFA analysis was conducted in Mplus 7.4 (Muthén & Muthén, 2015). We used Structural Equation Modeling (SEM) to compare the fit of four nested models suggested by our theory that the instrument measures interests, attitudes, and gender biases of students toward engineering, and also by the results of our EFA. All model tests were based on the covariance test and used Maximum Likelihood (ML) estimation.

Explanatory Modeling

Our goal is a causal explanation of the outcome variables, which is feasible given the design of the study as a randomized, controlled trial (RCT). The relationship between predictors and outcome variables was found to be linear, observations were independent, and our sample size is sufficient for adequate power, therefore we chose to use Multiple Regression to measure the effects of our predictor variables on the outcome variables. We also checked for homoscedasticity and normality of errors.

Predictor variables were chosen based on theoretical relevance to the constructs of attitudes towards and interest in engineering; all prior variables were also used in pilot studies of engineering attitudes and interest, and found to be significant. These variables and their descriptors are given in Table 2, and derived from the full sample after drops (due to incomplete surveys), N=10128. Missing data is excluded listwise. The "Enter" method in SPSS 23 is used, because predictor variables were chosen based on relevance and prior demonstration of utility.

Table 2: Descriptors for predictor variables.

	Description	Var. Type	N	Min.	Max.	Mean	S.D.
Year_ID	0 = Year 1; 1 Year 2 of study (Teachers more experienced)	Categorical	10906	0	1	.4484	.49735
S_Gender	0 = Male; 1 = Female	Categorical	10594	0	1	.49	.500
S_Undrep	1 = Underrepresented minority	Categorical	10557	0	1	.35	.476
S_Books	0 = Few (0-10 books) 1 = 1 shelf (11-25 books) 2 = 1 bookcase (26-100 books) 3 = Several bookcases (>100)	Ordinal	10137	0	3	2.00	.999
C_Grade	# = Grade of student	Categorical	10906	3	5	4.108	.8092
O_Treat	0 = Comparison; 1 = Treatment	Categorical	10906	0	1	.45	.498
Valid N	(listwise)		10128				

Results

Exploratory Factor Analysis and Scale Development

EFA was conducted on each of set of items (PRE and NOW) concurrently, to ensure a factor structure that worked with both the PRE "Last summer, I would have said" and NOW "Now I would say" sets of responses. Throughout the EFA process, we worked to find the best fit that was the same model for both the PRE and NOW sets of items, and also made the most sense thematically. There were differences in the pattern coefficients for each set, requiring some compromise. Where the pattern coefficients were different across the two sets, we chose the thematically most sensible placement. This led to the construction of a model that was not fully ideal for either set of items. However, the final model was close to the ideal for each set of items, and made sense given the theoretical framework.

Parallel analysis run with 99% probability cutoff on the first random half of the sample determined that 7 factors existed for both PRE and NOW items; however internal reliability analysis and EFA showed that 2 items (items 10 and 23, see Table 3 below) should be dropped due to low communalities, low inter-item correlation, and negative impact on reliability. Therefore we ran the analysis again without these items. Parallel analysis showed 6 factors; however the scree plot showed a gradual bend, with the possibility of 4-6 factors. EFA was run setting the number of factors to 4, 5, and 6 respectively.

Two factors describing gender were consistently the first extracted across all EFA models; however they negatively affected reliability for the full set of variables. Therefore we analyzed these separately for reliability and factorability, and when they proved viable, re-ran parallel analysis and EFA for the remaining variables. We repeated this process, removing the strongest factor or factors in each case, until all factors were identified. The final two factors had low Cronbach's α and low communalities; however these factors fit together thematically as "Desire to Learn" and had much better reliability when combined. EFA on the joined factor also showed a good fit. The final factor structure is described in Tables 3, 4, and 5.

Table 3: Items in final scales

Item#	Scale	Text of Item from the Instrument
3	Enjoyment	I enjoy studying engineering
6	Enjoyment	We learn about interesting things when we do engineering in school
8	Enjoyment	Engineering is fun
9	Enjoyment	When we do engineering, we use a lot of interesting materials & tools
13	Enjoyment	I am interested when we do engineering in school
1	Desire2Learn	It is important for me to understand engineering
2	Desire2Learn	Engineering helps me understand today's world
5	Desire2Learn	I would like to work with other engineers to solve eng. problems
18	Desire2Learn	I would enjoy being an engineer when I grow up
26	Desire2Learn	I would like to learn more about engineering
30	Desire2Learn	I really want to learn engineering
14	Value2Society	Engineers help make people's lives better
17	Value2Society	I know what engineers do for their jobs
21	Value2Society	Engineering is useful in helping to solve the problems of everyday life
22	Value2Society	We learn about important things when we do engineering in school
24	Value2Society	Engineering is really important to my country
25	Value2Society	I try hard to do well in engineering
4	MaleBias	Boys are better at engineering than girls
28	MaleBias	Girls have a harder time understanding engineering than boys
19	FemaleBias	Boys have a harder time understanding engineering than girls
20	FemaleBias	Girls are better at engineering than boys
15	DROP	Girls and boys are equally good at engineering
10	DROP	It is important to understand engineering in order to get a good job
23	DROP	Engineering is easy for me

Table 4: Internal consistency reliability for factors extracted with EFA from Random Half 1

Scale	Description	PRE α	NOW α	Min	Max
Enjoyment	Engineering is fun and interesting	.798	.814	0	20
Desire2Learn	Desire to engineer in the future	.770	.785	0	24
Value2Society	Value of engineering in the world	.793	.743	0	24
MaleBias	Boys are superior at engineering	.706	.752	0	8
FemaleBias	Girls are superior at engineering	.730	.757	0	8

Table 5: Pattern matrices for PRE & NOW; 3 scales (not GenderBias).

	Enjoyment		Value2Society		Desire2Learn	
	1 PRE	1 NOW	2 PRE	2 NOW	3 &4 PRE	3 & 4 NOW
EIA_1a					**.551**	**.463**
EIA_2a					**.511**	**.443**
EIA_3a	**.350**	**.431**			.404	
EIA_5a					**.311**	**.443**
EIA_6a		**.520**			-.531	
EIA_8a	**.452**	**.629**				
EIA_9a		**.437**			-.410	
EIA_13a	**.416**	**.626**				
EIA_14a			**.580**	**.645**		
EIA_17a			**.380**	**.382**		
EIA_18a	.610					**.725**
EIA_21a			**.560**	**.565**		
EIA_22a		.319	**.375**		-.351	
EIA_24a			**.670**	**.583**		
EIA_25a			**.404**			
EIA_26a	.733	.320				**.593**
EIA_30a	.754					**.676**

Extraction Method: Principal Axis Factoring. Items chosen for scales in bold.

Confirmatory Factor Analysis

We initially examined and compared four nested models. The models are briefly illustrated in Figure 1, and fit indices are presented in Table 6. Neither of the first two models (both PRE and NOW) was positive definite. Problems with the first PRE model could not be straightforwardly corrected, so Model 1 was rejected. Both versions of Model 2, however, could be corrected by setting items 4 and 20 (PRE model) and item 19 (NOW model) to variance=0.

Of Models 2, 3, and 4, Model 4 had the best measure of parsimonious fit, the Aikake information criteria (AIC), as well as the lowest χ^2—for both these indices, a lower value is superior. However, Model 4 also had the worst measures of absolute fit (largest RMSEA and SRMR) and comparative fit (lowest CFI and TLI). The rules of thumb for these measures is that the root mean square error of approximation (RMSEA) should be less than .05 (for best fit) and below .1 for good fit; the standardized root mean square residual (SRMR) should be less than .08; and the comparative fit index (CFI) and Tucker-Lewis Index (TLI) should exceed .95 if possible.

To explore ways to improve fit, we examined the model modification indices. These indicated that the greatest improvements for both PRE and NOW models would be achieved by (1) dropping the correlational relationship between each of the gender bias latent variables and the latent variable “Desire to Learn”, and (2) by cross-loading items EIA_1 and EIA_2 with the “Value to Society” latent variable. We modified the PRE and NOW Model 2 the model which retained the most latent variables, which greatly improved fit indices for the NOW model, and slightly improved fit indices for the PRE model. The indices for the new model, which we called 2m (for 2-modified), are shown in Table 6; the diagram for the model is given in Figure 2.

Standardized parameter estimates for Model 2m are given in Table 7, with PRE / NOW specified in each cell. All model parameters were highly significant (p<.001) except V2SOC BY EIA_2b for the NOW model (p<.05), and except for the three parameters with variance set to zero (MBIAS BY EIA_4a and FBIAS BY EIA_20a in the PRE model, and FBIAS BY EIA_19b in the NOW model). Model parameters, with the exception of the aforementioned parameters, also explained substantial item variance (R^2 = .13 to .91). Disattenuated correlations between the factors are presented in Table 8.

Figure 1: Four nested models to compare

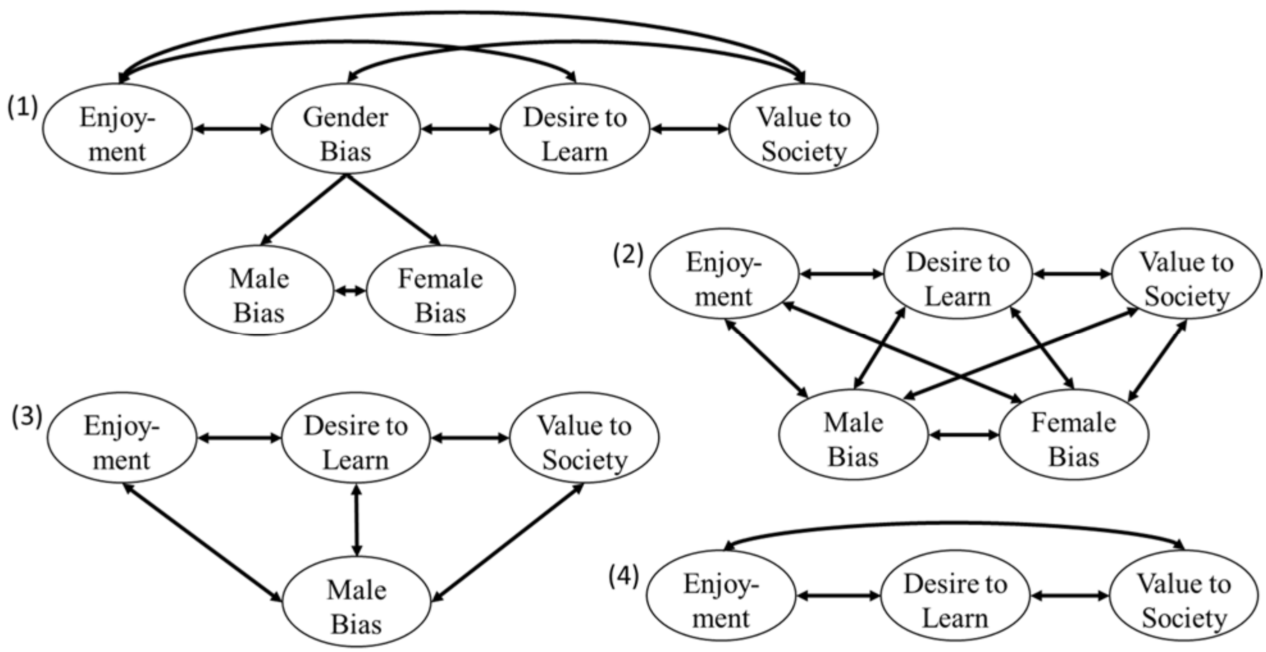

Table 6: Fit indices for four nested models

Model	χ^2	df	AIC	CFI	TLI	RMSEA	SRMR
1 PRE	2075.6	181	352581	.949	.941	.044	.030
2 PRE	2071.5	179	352580	.949	.941	.044	.030
3 PRE	1958.2	146	316293	.948	.939	.048	.030
4 PRE	1886.4	116	278724	.946	.937	.053	.032
1 NOW	2885.3	181	316423	.928	.916	.052	.042
2 NOW	2884.3	179	316426	.928	.915	.052	.042
3 NOW	2770.7	146	279937	.924	.910	.057	.045
4 NOW	2660.9	116	243132	.920	.906	.063	.048
2m PRE	2012.0	181	352517	.951	.943	.043	.030
2m NOW	2135.0	180	315674.945	.948	.939	.044	.034

Figure 2: Diagram of Latent Variables / Factors, Model 2m

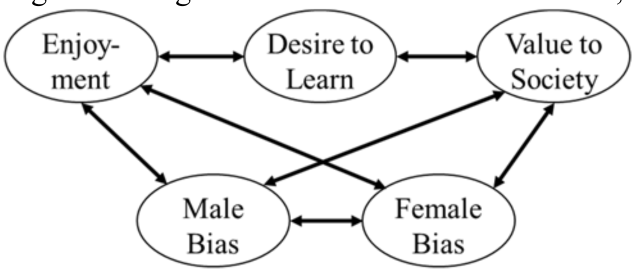

Table 7: Standardized parameter estimates for Model 2m: PRE / NOW.

Item	ENJOY	DESIRE	V2SOC	MBIAS	FBIAS	h^2
EIA_1		.31 / .13	.27 / .46			.30 / .30
EIA_2		.22 / .04*	.31 / .49			.25 / .27
EIA_3	.68 / .72					.47 / .51
EIA_4				1** / .67		1** / .45
EIA_5		.55 / .59				.31 / .35
EIA_6	.64 / .65					.41 / .43
EIA_8	.74 / .79					.55 / .63
EIA_9	.60 / .53					.36 / .28
EIA_13	.72 / .76					.52 / .57
EIA_14			.62 / .53			.38 / .28
EIA_17			.60 / .50			.36 / .25
EIA_18		.65 / .63				.43 / .39
EIA_19					.59 / 1**	.35 / 1**
EIA_20					1** / .63	1** / .40
EIA_21			.65 / .56			.43 / .32
EIA_22			.70 / .66			.49 / .43
EIA_24			.59 / .52			.35 / .27
EIA_25			.61 / .57			.38 / .32
EIA_26		.75 / .81				.56 / .66
EIA_28				.56 / .91		.31 / .83
EIA_30		.71 / .81				.51 / .66

Note: For all values except those marked * or **, p<.001. *: p<.05. **: Estimate set to 1.

Table 8: Disattenuated correlations between factors

Factor	ENJOY	DESIRE	V2SOC	MBIAS	FBIAS
ENJOY	1.00				
DESIRE	.93** / .89**	1.00			
V2SOC	.84** / .77**	.81** / .66**	1.00		
MBIAS	.02 / -.07**	-Set to 0-	-.01 / -.16**	1.00	
FBIAS	-.00 / -.07**	-Set to 0-	-.04* / -.17**	-.14**/ .20**	1.00

* p<.01; **p<.001

Given the slight changes to our factors (adding items EIA_1 and EIA_2 to the factor "Value to Society"), we recalculated internal consistency reliability for Random Half 1, Random Half 2, and the full dataset. Cronbach's α is given in Table 9 below.

Table 9: Internal consistency reliability for factors post-CFA (each sample)

Scale	Min	Max	Random Half 1		Random Half 2		Full Sample	
			PRE α	NOW α	PRE α	NOW α	PRE α	NOW α
Enjoyment	0	20	.798	.814	.812	.820	.805	.817
Desire2Learn	0	24	.770	.785	.782	.786	.776	.785
Value2Society	0	32	.812	.780	.820	.772	.816	.776
MaleBias	0	8	.706	.752	.717	.757	.711	.754
FemaleBias	0	8	.730	.757	.741	.772	.736	.765

Multiple Linear Regression Analysis

For each of our five factors (scale variables), we calculated a multiple linear regression to predict the NOW scale dependent variable (DV) based on the following independent variables (IV's—see Table 2): the related PRE scale variable, Year_ID, S_Gender, S_Undrep, S_Books, C_Grade, and O_Treat. Significant regression equations were found for all five dependent variables, shown in Tables 10 through 15.

For the first three variables—Enjoyment, Desire2Learn, and Value2Society, ratings increased substantially from PRE to NOW, showing an increase in interests and attitudes. For both MaleBias and FemaleBias, ratings decreased from PRE to NOW, showing a decrease in gender biases.

The rating of engineering Enjoyment was .254 higher for each point marked on the PRE, .436 points more for female students, and was .214 higher for each point on the Books scale (where students rated how many books they had at home). The increase in enjoyment was .544 less for underrepresented minority students, was .154 less for each grade year, and .664 less with participation in the treatment curriculum. Both treatment and comparison interventions led to increased scores, but comparison led to slightly higher scores than treatment. Similarly, girls' scores increased more than boys' scores, younger students more than older students, higher-SES students (as measured by Books) more than lower-SES students, and white/Asian students' scores more than Black/Hispanic scores.

The rating of Desire2Learn engineering was .480 higher for each point marked on the PRE, .316 higher for students of more experienced teachers, females marked their scores .229 points higher than male students, and we see .160 higher scores for each point on the Books scale (where students rated how many books they had at home). The increase in Desire2Learn was .394 less for minority students, was .197 less for each grade year, and .460 less with participation in the treatment curriculum. This pattern of changes is very similar to that of Enjoyment, except that teacher experience also had a positive effect.

Table 10: Regression equations for dependent variables

DV	Regression Equation	Significance	R^2
Enjoyment	F(7, 9841) = 172.398	p<.000	.109
Desire2Learn	F(7, 9800) = 637.908	p<.000	.313
Value2Society	F(7, 9730) = 161.685	p<.000	.104
MaleBias	F(7, 9517) = 1407.740	p<.000	.509
FemaleBias	F(7, 9674) = 1514.115	p<.000	.523

Table 11: Descriptives and Coefficients for DV EnjoyNOW

Model	Mean	S.D.	Unstd. Coefficients		Std. Coeff.	t	Sig.
			B	Std. Error	Beta		
(Constant)	16.271	4.114	14.238	.253		56.229	.000
EnjoyPRE	9.788	4.979	.254	.008	.308	32.074	.000
Year_ID	.448	.497	.020	.079	.002	.260	.795
S_Gender	.490	.500	.436	.079	.053	5.536	.000
S_Undrep	.340	.473	-.544	.086	-.063	-6.303	.000
S_Books	2.010	.994	.214	.041	.052	5.219	.000
C_Grade	4.108	.809	-.154	.049	-.030	-3.149	.002
O_Treat	.450	.497	-.664	.080	-.080	-8.309	.000

Table 12: Coefficients for DV DesireNOW

Model	Mean	S.D.	Unstd. Coefficients		Std. Coeff.	t	Sig.
			B	Std. Error	Beta		
(Constant)	16.774	4.676	11.926	.256		46.590	.000
DesirePRE	11.277	5.394	.480	.007	.554	65.516	.000
Year_ID	.448	.497	.316	.079	.034	4.008	.000
S_Gender	.500	.500	.229	.079	.025	2.907	.004
S_Undrep	.340	.473	-.394	.086	-.040	-4.566	.000
S_Books	2.020	.993	.160	.041	.034	3.891	.000
C_Grade	4.107	.809	-.197	.049	-.034	-4.021	.000
O_Treat	.450	.497	-.460	.080	-.049	-5.754	.000

Table 13: Coefficients for DV Value2SocietyNOW

Model	Mean	S.D.	Unstd. Coefficients		Std. Coeff.	t	Sig.
			B	Std. Error	Beta		
(Constant)	26.831	4.989	21.978	.315		69.878	.000
Val2SocPRE	17.296	6.695	.192	.007	.258	26.682	.000
Year_ID	.449	.497	.274	.096	.027	2.843	.004
S_Gender	.500	.500	.390	.096	.039	4.044	.000
S_Undrep	.340	.473	-1.212	.105	-.115	-11.493	.000
S_Books	2.020	.992	.560	.050	.111	11.145	.000
C_Grade	4.109	.808	.173	.060	.028	2.884	.004
O_Treat	.450	.497	-.502	.098	-.050	-5.141	.000

Table 14: Coefficients for DV MaleBiasNOW

Model	Mean	S.D.	Unstd. Coefficients		Std. Coeff.	t	Sig.
			B	Std. Error	Beta		
(Constant)	2.333	2.674	1.810	.125		14.482	.000
MBiasPRE	2.735	2.671	.640	.008	.639	80.526	.000
Year_ID	.449	.497	-.106	.039	-.020	-2.741	.006
S_Gender	.500	.500	-.511	.042	-.096	-12.179	.000
S_Undrep	.330	.472	.435	.043	.077	10.212	.000
S_Books	2.020	.990	-.173	.020	-.064	-8.520	.000
C_Grade	4.115	.8073	-.161	.024	-.049	-6.649	.000
O_Treat	.450	.497	-.134	.039	-.025	-3.422	.001

Table 15: Coefficients for DV FemaleBiasNOW

Model	Mean	S.D.	Unstd. Coefficients		Std. Coeff.	t	Sig.
			B	Std. Error	Beta		
(Constant)	2.461	2.688	1.289	.118		10.939	.000
FBiasPRE	2.523	2.599	.716	.008	.692	90.730	.000
Year_ID	.444	.497	.046	.038	.009	1.214	.225
S_Gender	.500	.500	.189	.041	.035	4.635	.000
S_Undrep	.340	.472	.332	.042	.058	7.942	.000
S_Books	2.020	.990	-.090	.020	-.033	-4.500	.000
C_Grade	4.110	.808	-.158	.024	-.047	-6.660	.000
O_Treat	.450	.497	-.066	.039	-.012	-1.723	.085

The rating of engineering Value2Society was .192 higher for each point marked on the PRE, .274 higher for students of more experienced teachers, .390 points higher for female students, .560 higher for each point on the Books scale, and .173 higher for each grade year. The rating of Value2Society was less by 1.212 points for minority students compared to non-minorities, and .502 points less with participation in the treatment curriculum as compared to the comparison curriculum. Here we see that underrepresented minority students show a much smaller increase in scores than non-minority students. Unlike the earlier scales, older students are likely to give larger increases in ratings than younger students.

MaleBias and FemaleBias both decreased slightly, according to student self-report, after the intervention. Students with higher PRE scores and underrepresented minority students tended to show smaller decreases or even increases in MaleBias, but all other variables (including treatment) were associated with greater decreases in MaleBias. FemaleBias was similarly affected by higher PRE scores and minority status, but it was also effected by student gender, with female students scoring .189 points higher on the NOW scale than male students. Books, Grade, and Treatment were all associated with lower FemaleBias.

Discussion and Conclusion

In this paper we discussed the development of the Engineering Interests and Attitudes (EIA) measure for elementary school students. We discussed the development of factor models for the instrument, and showed that the scales we developed from EFA and CFA are reliable and can be used to measure changes in student enjoyment of engineering, desire to learn engineering, and value of engineering to society. It can also be used to measure the level of student gender bias in engineering. We expect that researchers and curriculum developers will want to use this instrument with students ages 8-11 to measure changes in student interests and attitudes after participation in engineering activities, programs, and curricula.

In examining our data set using the instrument, we found that students who participated in an engineering curriculum showed greater enjoyment, desire to learn, and valuation of engineering. The increase in attitudes was slightly larger for students in the comparison curriculum. However, students in the treatment curriculum showed larger decreases in gender biased attitudes. The interests and attitudes of girls increased more than that of boys in our data set, but students from minority groups underrepresented in engineering showed smaller increases than non-minority students. These are mixed results for our treatment curriculum, which we will investigate further and attempt to explain with qualitative research.

In future work, we will check further for interaction effects of variables. We will incorporate mediating variables into our models, and use hierarchical modeling methods for increased power. We will also further explain the evidence for validity of the instrument.

References

Cunningham, C. M., & Lachapelle, C. P. (2010, June). *The impact of Engineering is Elementary (EiE) on students' attitudes toward engineering and science*. Paper presented at the American Society for Engineering Education Annual Conference & Exposition, Louisville, KY.

Germann, P. J. (1988). Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal of Research in Science Teaching*, *25*(8), 689–703.

IBM Corporation. (2012). IBM SPSS Statistics for Windows (Version 21.0). Armonk, NY: IBM Corporation.

Lichtenstein, M. J., Owen, S. V., Blalock, C. L., Liu, Y., Ramirez, K. A., Pruski, L. A., … Toepperwein, M. A. (2008). Psychometric reevaluation of the scientific attitude inventory-revised (SAI-II). *Journal of Research in Science Teaching*, *45*(5), 600–616. https://doi.org/10.1002/tea.20244

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84.

Muthén, L., & Muthén, B. (2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Owen, S. V., Toepperwein, M. A., Marshall, C. E., Lichtenstein, M. J., Blalock, C. L., Liu, Y., … Grimes, K. (2008). Finding pearls: Psychometric reevaluation of the Simpson–Troost Attitude Questionnaire (STAQ). *Science Education*, *92*(6), 1076–1095.

Sawyer, R. K. (2006). Introduction: The new science of learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 1–16). Cambridge, UK: Cambridge University Press. Retrieved from internal-pdf://Sawyer 2006 New science of learning-3484481024/Sawyer 2006 New science of learning.pdf

Weinburgh, M. H., & Steele, D. (2000). The Modified Attitudes toward Science Inventory: Developing an Instrument to Be Used with Fifth Grade Urban Students. *Journal of Women and Minorities in Science and Engineering*, *6*(1), 87–94.