Evaluation of Innovative Technology Experiences for Students and Teachers (ITEST) 2016-2017 Grant Activities

For

The Marine Advanced Technology Education (MATE) Center

August 2017

Submitted by:

SESRC

Social & Economic Sciences Research Center (SESRC)

Puget Sound Division
Washington State University

PO Box 43170
Olympia, Washington 98504-3170
Telephone: (360) 586-9292
Fax: (360) 586-2279

Evaluation of Innovative Technology Experiences for Students and Teachers (ITEST) 2016-2017 Grant Activities

For

The Marine Advanced Technology Education (MATE) Center

Candiya Mann and Yi Jen Wang

August 2017

WASHINGTON STATE UNIVERSITY

Social & Economic Sciences Research Center-Puget Sound Division
203 E. 4th Avenue, Suite 521
P.O. Box 43170
Olympia, WA 98504-3170
(360) 586-9292
Fax: (360) 586-2279

SPONSORSHIP

Support for this project was provided by the Marine Advanced Technology Education (MATE) Center and the National Science Foundation.

ABOUT SESRC

The Social and Economic Sciences Research Center (SESRC) at Washington State University is a recognized leader in the development and conduct of survey research.

SESRC-Puget Sound Division provides technical services and consultation to assist clients in acquiring data, understanding what data means, and applying that information to solving problems. The SESRC Puget Sound Division specializes in research design, data collection and analysis, using both qualitative and quantitative methods. The Division also provides interpretive reports, policy studies, presentations and consulting services directly to individual clients, organizations and consortia.

ACKNOWLEDGEMENTS

The authors would like to thank the management at the MATE Center for their leadership, coordination, and helpful suggestions. We are also grateful to the regional coordinators and their support staff, who made themselves available for interviews and provided valuable insights into the project. We appreciate Dr. Min Li's work improving our surveys and scoring rubrics. Finally, this project would not have been possible without the generous and thoughtful input contributed by the student participants, faculty/mentors, parents and judges/volunteers who completed surveys.

TABLE OF CONTENTS

TABLE OF TABLES AND FIGURES

EXECUTIVE SUMMARY

Evaluation of Innovative Technology Experiences for Students and Teachers (ITEST) Grant Activities

For the Marine Advanced Technology Education (MATE) Center

BY: CANDIYA MANN & YI JEN WANG

SOCIAL & ECONOMIC SCIENCES RESEARCH CENTER, PUGET SOUND OFFICE

WASHINGTON STATE UNIVERSITY

JULY 2017

In September 2013, the National Science Foundation (NSF) funded the Marine Advanced Technology Education (MATE) Center's proposal for an Innovative Technology Experiences for Students and Teachers (ITEST) grant. The MATE Center's ITEST program, titled *Scaling up Success: Using MATE's ROV Competitions to Build a Collaborative Learning Community that Fuels the Ocean STEM Workforce Pipeline*, leverages their extensive network of remotely operated vehicle (ROV) student competitions. The project's overarching goal is to encourage multi-year student participation in an effort to deepen student interest and learning and reinforce pathways leading to the STEM workforce.

The evaluation is based on multiple data sources, primarily surveys and interviews, and reflects the input of a variety of stakeholders, including students, teachers, parents, judges, volunteers, regional coordinators, and MATE management and staff. This report covers grant activities that took place between July 1, 2016 and May 15, 2017.

Findings

Project Goal 1: Increase middle and high school students' interest in STEM and STEM careers, as well as their knowledge of STEM and understanding of how science and engineering work together to solve real-world problems.

- **Increased Awareness of STEM Careers:** After building their ROV, over three-quarters of the students (77%, N=1,139) indicated that they knew more about careers in marine STEM.
- **Increased Interest in STEM Careers:** Three-quarters (75%, N=1,138) of the students stated that their ROV project made them more interested in pursuing a STEM career, and 90% of the teachers (N=236) observed an increase in their students' interest in pursuing a STEM career. Ninety-seven percent (96%, N=233) agreed that the ROV program provided a valuable venue to help prepare their students for STEM careers.

- **Increased Interest in STEM:** Over three-quarters of the students (76%, N=1,135) indicated that their ROV project made them want to learn more about STEM. Ninety percent (90%) of the teachers (N=236) and 88% of the parents (N=281) observed greater interest among the students in learning STEM.
- **Increased STEM Knowledge & Skills:** The great majority of the students reported increased skills and knowledge due to their ROV project in several subjects: engineering (90%, N=1,131), technology (89%, N=1,125), the competition theme (75%, N=1,114), science (79%, N=1,129), and math (61%, N=1,124). The majority of the teachers (98%, N=236) observed improvements in their students' STEM knowledge and skills. Parents reported that building an ROV contributed to improving their children's grades in engineering/robotics (47%, N=276), science (55%, N=279), math (48%, N=278) and computers (43%, N=276).
- **Increased 21st Century Skills:** Students reported that participating in the ROV project improved their problem solving (84% agreed or strongly agreed, N=1,124), teamwork (85%, N=1,113), critical thinking ((81%, N=1,122), leadership (73%, N=1,127), and organization skills (66%, N=1,124). Ninety-seven percent (97%, N=235) of the teachers observed increases in their students' skills in team building, problem solving, and/or critical thinking. Parents reported that their children were better problem solvers (96%, N=276), critical thinkers (94%, N=276), team members (95%, N=278), and/or leaders (86%, N=278).
- **Overall Rating of MATE Center Support:** After the competition season, 50% of the teachers (N=242) rated the support provided by MATE as excellent, and 38% provided a rating of good, for an overall positive rating of 88%.
- **Overall Opinions of ROV Program:** The ROV program was rated positively (excellent or good) by 87% of the students N=1,150), 98% of the teachers (N=242) and 96% of the parents (N=282).
- **Ability to Apply STEM to Real World Problems:** In the post-competition surveys, 83% of the students (N=1,132) indicated that participating in the ROV project helped them learn to apply STEM to real world problems, and 93% of the teachers (N=236) observed improvements in their students' abilities in this area, as did 92% of the parents (N=281).
- **Ability to Communicate Engineering Process and Designs to a Wide Audience:** Seventy-nine percent (79%) of the students **(**N=1,135) stated that participating in the ROV project helped them learn how to communicate their engineering design to other people. Eighty-nine percent (89%) of the instructors (N=236) observed improvements in their students' skills in this area.
- **Influence on Students' Educational and Career Paths:** ROV Competition Student Alumni Survey results include the following:
    - Among the 220 alumni who earned a college degree, 85% earned a degree in a STEM discipline.
    - Among the 236 current college and university students, 85% are studying towards a STEM degree.

- Among the employed alumni (N=320), 73% are currently working a STEM-related job, and 22% currently or previously worked a job related to ROVs or other underwater technologies.
- Two-thirds (67%, N=432) of the alumni credit the ROV competition with influencing their educational or career path "to a great extent" or "somewhat".
- The ROV competition played a role in alumni attaining employment (37%), admittance into educational programs/college/university (36%), internships (30%), awards (21%), and scholarships (21%).

- **Effect of Multi-Year Competition Participation:** Multi-year participants were statistically significantly more likely to report their participation in the ROV program resulted in higher levels of awareness of and interest in pursuing STEM careers, gains in interest in taking STEM courses, improvements in STEM knowledge and skills, increased 21st Century skills, and the receipt of awards, honors, and new educational and career opportunities.
- **Impacts among Underrepresented Groups:** According to the demographic data in the surveys (N=1,145), the students were over one-quarter female (31%); 34% were of minority backgrounds; 45% came from high poverty areas; and 3% reported that they had disabilities requiring accommodations. Statistically significant differences existed between the groups (gender, ethnicity, socioeconomic status, and disability status) in the following measures:
    - ***Awareness of and Interest in STEM Careers:*** Male participants were more likely than female participants cite the competition as influencing them towards a STEM career. Students living in low poverty areas were more likely to indicate that the competition led to gains in knowledge of and interest in STEM careers.
    - ***Interest in STEM Topics:*** Males were more likely to report increased interest in computer science and engineering courses. Students in low poverty areas were more likely to indicate that the competition increased their desire to take science, math, and hands-on classes or club activities. Students with disabilities were less likely to report increased interest in computer science courses.
    - ***STEM Skills and Knowledge:*** After the competition, male students and white students were more likely to report increased skills and knowledge in engineering. Students living in a lower poverty area were more likely to report gains in skills and knowledge in engineering and technology.

Project Goal 2: Provide teachers with professional development, instructional resources, and mentors to support and sustain the delivery of STEM career information and learning experiences.

- **Increased Confidence Facilitating STEM Learning Experiences:** The percentage of workshop survey respondents who rated themselves as "very comfortable" facilitating STEM learning experiences for students rose between the pre (N=138) and post (N=143) workshop surveys in all STEM areas: science (pre: 56%, post: 79%), technology (pre: 36%, post: 52%), engineering (pre: 30%, post: 58%), and math (pre: 38%, post: 54%).

- **The MATE Community:** Among the post-competition surveys, 77% of the teachers (N=230) agreed that they felt they were part of a MATE community that provides support and relevant resources.
- **MATE Robotics Activities/Curriculum Incorporated into Courses and after-School Programs:** Seventy-one percent (71%) of the post-competition teacher survey respondents (N=272) incorporated building ROVs into an after-school club. Seventeen percent (17%) built ROVs as part of a course; 22% built ROVs as a voluntary activity; and 5% built ROVs in another venue.
    - Over three-quarters (76%) of the competition teachers (N=230) stated that they used MATE materials and resources to incorporate ROV building into their course or club, and over one-third (41%) modified their curriculum and teaching based on MATE resources.
    - In the one-year follow-up survey of the week-long, intensive 2016 Pufferfish Summer Institute, 19 participants built ROVs with students. A total of 109 ROVs were built with 1,687 students.
- **Classroom Mentors:** In several regions, the regional coordinator matched up college and high school students – in many cases, former ROV competitors themselves – with middle school ROV teams to work with them throughout the competition season. For 25% of the post-competition teacher survey respondents (N=230), a classroom/club mentor came to their site to help their teams. Among these teachers, over one-third (37%) indicated that the mentor helped them incorporate robotics into their course or club to "a great extent". The majority of the respondents (79%) indicated that their mentors were adequately prepared to help them and their students through the ROV design and building process.

Project Goal 3: Increase parental involvement in order to support and encourage students to pursue STEM education and careers.

- **Increased Parental Support of Their Children's Interest in STEM:** Ninety-two percent (92%) of the parents (N=276) indicated that participation in the ROV program changed how they envisioned their child's future, making it easier to picture their child with a STEM career.
- **Enhanced Online Resources:** The online Parent Resource Center was launched in the spring of 2015. It contains competition videos, frequently asked questions, background information, highlights, and contact information for the MATE Center, along with types of information that the MATE Center can provide upon request. Anecdotal feedback indicates that the parents found the resources helpful.
- **Regional Advisory Committees:** Advisory committees included participation from parents as well as industry representatives, professional organizations, government agencies, 6-12th grade educators, community college faculty, and university faculty. In 2016-2017, some regions held an annual advisory meeting, while others held monthly advisory meetings, bi-annual meetings or communicated on an ongoing basis with advisory members outside of formal meetings. The regional coordinators were responsive to their committees' recommendations.

Conclusions

The MATE Center successfully implemented the 2016-2017 year of ITEST grant activities. The 2017 MATE ROV Competition was held, with ITEST funding helping to support 16 of 20 18 US-based regional events. A total of 413 regional workshops were held for teachers and students, and 20 teachers attended the intensive (60 hour) Summer Institute professional development. The Center disseminated a suite of online instructional materials, including videos, PowerPoints, ROV kits, and an online course.

Evaluation results continue to show strong positive outcomes for students and teachers. Involvement in the ROV competition generated greater awareness of and interest in pursuing STEM careers; increased interest in studying STEM topics; improved STEM knowledge and skills; and increased teamwork, critical thinking, leadership, and problem solving skills. Participating in the ROV competition helped students learn how to apply STEM skills to real world problems. They also learned how to communicate their engineering process and design to a wide audience.

ROV competition student alumni survey results and National Student Clearinghouse match analysis suggest that the majority of ROV competition participants go on to study STEM topics, earn STEM degrees, and work in STEM fields. In fact, roughly one in five former participants have worked in a job related to ROVs or other underwater technologies. The majority of ROV competition alumni credit the ROV competition with influencing their educational and career paths, including playing a role in attaining internships, scholarships, admittance to educational programs, and employment.

These findings suggest that the MATE ROV Competition is effective in increasing the STEM workforce, especially related to underwater technologies.

INTRODUCTION

In September 2013, the National Science Foundation (NSF) funded the Marine Advanced Technology Education (MATE) Center's proposal for an Innovative Technology Experiences for Students and Teachers (ITEST) grant. The MATE Center's ITEST program, titled *Scaling up Success: Using MATE's ROV Competitions to Build a Collaborative Learning Community that Fuels the Ocean STEM Workforce Pipeline*, leveraged their extensive network of remotely operated vehicle (ROV) student competitions. The project's overarching goal is to encourage multi-year student participation in an effort to deepen student interest and learning and reinforce pathways leading to the STEM workforce.

As stated in the proposal, the goals are fourfold:

1. Increase middle and high school students' interest in STEM and STEM careers as well as their knowledge of STEM and understanding of how science and engineering work together to solve real-world problems.
2. Provide teachers with professional development, instructional resources, and mentors to support and sustain the delivery of STEM learning experiences and career information.
3. Increase parental involvement in order to support and encourage students to pursue STEM education and careers.
4. Track students longitudinally to document how participation impacts their education and career path.

This report covers grant activities that took place between July 1, 2016 and May 15, 2017. (NSF requested that this year's report be turned in early so it doesn't cover a full year.) The MATE Center's ITEST grant evaluation was performed by the Puget Sound Division of the Social and Economic Sciences Research Center at Washington State University.

Methodology

The evaluation connects each of the project goals with evaluation questions and expected outcomes of the project. These goals and evaluation questions are presented below.

Table 1: Project Strategies and Evaluation Questions

Project Goal	**Evaluation Questions**
1. Increase middle and high school students' interest in STEM and STEM careers as well as their STEM knowledge and understanding of how science and engineering work together to solve real-world problems. • Add SCOUT+ class • Support for students who want to continue competition at next grade/school • Mentoring from students/industry professionals • Career advice/videos	1.1. To what extent did the MATE robotics activities lead to an increase in the students' interest in and knowledge of STEM content and STEM careers? Did educators and parents observe an increase in the students' interest in STEM content and STEM careers as a result of the robotics activities? An increase in the students' STEM knowledge and skills and 21st Century workplace skills?
	1.2. How did the robotics activities affect students' ability to apply STEM knowledge and skills to finding solutions to real-world problems?
	1.3. How did the robotics activities affect students' ability to communicate their engineering process and designs to a wide audience (from engineers and technicians to the general public)?
	1.4. How did participation in the robotics activities influence students' educational and career paths?
	1.5. What effect did multi-year participation have on the above evaluation questions?
	1.6. Did the robotics activities create the same impacts among underrepresented groups (by gender, ethnicity, socio-economic status, disability) as were found among students who traditionally participate in these types of activities?

2. Provide teachers with professional development, instructional resources, and mentors to support and sustain the delivery of STEM career information and learning experiences. • Curriculum continuum • Progression of ROV kits • Professional development workshops • Regional workshops • Regional teacher-leaders • Increase preparedness of mentors	2.1. Are teachers more confident delivering STEM learning experiences? Delivering career information and outlining career pathways?
	2.2. Do teachers feel they are a part of a larger MATE community that provides support and relevant, necessary resources?
	2.3. Do teachers incorporate MATE robotics activities/curriculum into courses and afterschool programs? Are the courses and/or curriculum adopted by school districts?
	2.4. Are teachers able to access classroom mentors as needed? Do the classroom mentors help them successfully incorporate robotics activities into the course? Are the classroom mentors adequately prepared?
3. Increase parental involvement in order to support and encourage students to pursue STEM education and careers. • Parent online resources/listserv • Regional parent advisory committees	3.1. Did the MATE robotics activities lead to an increase in the parents' support of their children's interest in STEM careers? 3.2. Did the enhanced parent online resources lead to an increase in the parents' ability to provide assistance and support for their children's involvement in the MATE robotics activities? 3.3. Did the regional parent advisory committees provide feedback and advice to help improve the competitions and ensure that the program is inclusive of all participants?

DATA SOURCES

The evaluation relies upon multiple sources of data. The data collection includes input from a variety of stakeholders, including students, teachers, parents, judges/volunteers, regional coordinators, and MATE staff. Below are descriptions of each of the data sources. All of the surveys were developed in collaboration with MATE staff and regional coordinators.

Student Follow-up

The evaluation includes several student follow-up efforts: 1) the Washington State Follow-up, 2) The ROV Competition Student Alumni Survey, and 3) the National Student Clearinghouse Data Match.

WASHINGTON STATE FOLLOW-UP

Background: As part of the ITEST proposal, the Washington State Education Research and Data Center (ERDC) agreed to match the 2006-2013 ROV competition program participants from Washington with the ERDC P-20 data and also to create a comparison group. They agreed to return high school, postsecondary, and workforce data. The economics section of the Washington State Office of Financial Management (OFM) Forecasting division agreed to do the statistical analysis of the ERDC data.

This analysis will explore two main research questions:

1) To what extent are the ROV competition participants the same/different from the general population of students, and

2) What is the impact of the ROV competition on the probability of attending college, studying STEM, persisting, and completing college degrees?

2016-2017 Update: Over the past 3 years, MATE and the evaluator have coordinated with ERDC to come to an agreement on work and budget. The evaluator has sent student participant data to ERDC several times for matching, and ERDC did preliminary matches in 2014-2015 and 2015-2016. Unfortunately, as of the end of this grant year, June 2017, we do not have a signed contract, and no substantial work has taken place. We attribute this lack of progress to the fact that the analyst within ERDC at the time of the proposal, who was a strong champion for this research project, passed away unexpectedly shortly after the proposal was funded. ERDC has been extremely busy and has had limited personnel to assign to our project. We will continue to negotiate with ERDC and hope to complete this analysis in 2017-2018.

ROV COMPETITION STUDENT ALUMNI SURVEY

In June of 2015, we conducted the ROV Competition Student Alumni Survey. The goal of the alumni survey was to answer the questions: “Where are they now?” and, “To what extent did their involvement with MATE influence their trajectory?” The survey included questions about their higher education, employment, internships, scholarships and other opportunities that opened due to their involvement with the ROV competition.

NATIONAL STUDENT CLEARINGHOUSE DATA MATCH

Background: The National Student Clearinghouse (NSC) is a nationwide source of higher education information. Colleges and universities, numbering over 3,400 institutions enrolling over 96% of college students, share their enrollment data with NSC. The NSC database includes over 130 million students. See http://www.studentclearinghouse.org/ for further information about the NSC.

2015-2016 Progress: A total of 3,974 former competition participants (aka "alumni") were identified for matching with National Student Clearinghouse (NSC) data in the May of 2017. These alumni have birthdates that indicate that they are at least 18 years of age as of May 2017, and their addresses suggest that they live within the United States. Two-thirds of the alumni (66%, or 2,633) were found in the NSC database.

Curriculum and Online Instructional Resources

PRE-POST KNOWLEDGE TESTS

In the ITEST proposal, the MATE Center proposed creating a complete curriculum, tied to standards, with pre-post knowledge tests corresponding to each module. In 2014-2015, the MATE Center changed focus in response to user feedback in an intensive set of 103 interviews conducted as part of an NSF I-Corps grant. Rather than a complete curriculum, PIs determined that teachers preferred online resources that they could incorporate into their own curricula.

With this information regarding teacher preferences, the MATE Center has changed focus from designing a complete curriculum to designing a well thought out menu of online curriculum modules and supporting resources, such as instructional videos, PowerPoint presentations, and other activities.

In 2016-2017, MATE began transitioning their educational resources to Google Slides, which allow for quick and easy updating. They have continued their transition to the Canvas Learning Management System. At this point, five courses are hosted on Canvas, and 400 educators have enrolled in at least one of these courses. There is one course aligned with each competency/ROV kit level, plus the *Diving into Underwater Sensors and Arduino* course. These courses include quizzes and worksheets.

We are working with Dr. Min Li to design, pilot, and assess the psychometric qualities of the courses' pre- and post- tests. In 2017, the evaluation plans to incorporate the pre-post test results.

Canvas offers the capability for educators to clone these courses and offer them to their students. MATE will be able to monitor student performance on the cloned courses. The cloned courses will provide valuable data for the evaluation because they solve the logistical problem of asking teachers to provide their students' test data to MATE. With this approach, the teachers will administer the assessments, and the data will be available to MATE seamlessly.

ROV Competitions

Background: At the ROV competitions, input was solicited from as many stakeholders as possible, including students, teachers, parents, and judges/volunteers. The competition surveys were primarily administered as paper surveys in a "scannable" format; there was a web option as well. Data entry was completed by scanning the surveys and entering the written comments by hand. Data analysis was performed with the Statistical Package for the Social Sciences (SPSS). Student and parent surveys were offered in Spanish as well.

In 2016-2017, in order to complete analysis for this report early at NSF's request, the timeframe of this report is July 1, 2016 through May 15, 2017. Because of this timeframe, the analysis does not include the international competition, which took place in June 2017.

POST-COMPETITION SURVEYS: STUDENTS

At the ROV competitions, students were asked to complete surveys. The survey protocol was a modified version of the student survey that has been administered to more than 7,657 students over the past nine years at regional and international ROV competitions. The survey covered the following topics: awareness and interest in ocean STEM careers, increased desire to take STEM courses due to involvement in the program, awards/honors received as a result of competition experience, and self-assessment of change in STEM knowledge.

POST-COMPETITION SURVEYS: TEACHERS

Teachers also completed surveys at the ROV competitions. The survey protocol was a modified version of the faculty/mentor survey that has been administered to more than 1,593 respondents over the past nine years at ROV competitions. The survey addressed topics such as the value of the competition, incorporation of competition into course curriculum, interest in participating in future competitions, assessment of change in their students' STEM knowledge and skills, 21st Century skills, interest in STEM careers, and related topics.

POST-COMPETITION SURVEYS: PARENTS

In contrast to the student and teacher surveys, which have been conducted for years at MATE ROV competitions, the 2010 competition season was the first time parent input was solicited. Parents responded enthusiastically and seemed to appreciate the opportunity to provide input. Parent surveys addressed the topics of parental support of their children's interest in STEM and STEM careers, the value of the competition, and changes they have observed in their children since they became involved in the program. In 2016-2017, 283 parents and other family members (such as grandparents) completed the survey.

POST-COMPETITION SURVEYS: JUDGES

In the 2011 competition season, input was solicited for the first time from industry representatives serving as judges at the competitions. This survey collects information on the judges' experience at the competition, whether they feel it was a worthwhile use of their time, the skills of the students they observed, and their opinions on the usefulness of the competition in preparing future employees. In 2016-2017, 128 judges and 83 volunteers completed the survey.

PSYCHOMETRIC ANALYSIS: VALIDATION OF SCORING RUBRICS, IMPROVING INTERNAL CONSISTENCY OF POST-COMPETITION SURVEYS

2015-2016 Update: In 2016-2017, the MATE Center continued the valuable collaboration with Dr. Min Li, Associate Professor at the University of Washington's Department of Education. In this grant year, Dr. Li focused on validating the competition scoring rubrics, establishing inter-rater reliability of competition scoring, and assessing whether last year's changes to the student and parent post-competition surveys improved the internal consistency. (Analysis confirmed that the changes had successfully improved internal consistency.) Please see below for Dr. Li's summary of work conducted in 2016-2017, findings, and plans for 2017-2018:[1]

***

The psychometric analysis to validate various instruments developed and used in this project focused on two main goals. First goal pertains to employing a process for validating competition scoring rubrics and determining and/or establishing methods for ensuring inter-rater reliability of competition scoring such that they may be used as an indicator of student learning. Our second goal is to verify the internal consistency of existing surveys by evaluating the technical quality of the questionnaire constructs.

Related to the goal of validating the competition scoring rubrics, we analyzed the scores assigned by judges using both the Cronbach's alpha as the indicator of the internal consistency and the Generalizability (G) coefficients calculated based on the G theory. For the second goal of verifying the internal consistency of the existing surveys/questionnaires, we performed the exploratory and confirmatory factor analysis with the survey items some of which had been refined and re-organized based on our previous analysis and identified the areas that survey items could be refined. The data for running the psychometric analysis were collected from the round of 2016 competition. The analytic procedures and syntax files can be easily applied with the collected data from previous rounds, including the samples of both ranger and explorer participants, to cross validate the validity issues that we have found from previous years' data collection.

---

[1] Note that the table and figure numbers in Dr. Li's summary are separate from the numbering of the figures and tables elsewhere in the report.

In what follows we organize the psychometric findings into two sections corresponding to the two goals described.

Validation of Competition Scoring Rubrics

Participants' products to the ROV competition are scored by volunteering judges into four types of scores: (1) technical documentation that focuses on the technical aspects of the report that a team submits electronically; (2) mission that assesses the performance of the ROV demonstrated by the team; (3) poster display that a team summarizes necessary information about their ROV project; and (4) product presentation that the team interactively reports the ROV project to a group of audience, including judges and other audience.

Prior to the statistical analysis, we performed the content analysis of the scoring rubrics by reviewing the focus of each rubric and carefully comparing them against the competition manual. The manual is supposed to include all the instructions to the participants which need to explicitly communicate the scoring rubrics and expectations. Applying this guideline, we flagged parts that can be unclear or confusing for the students and then modified, simplified, and clarified these aspects of the competition manual to streamline the instructions and align them closely to each of the scoring rubrics. In addition, we made the language of the instructions consistent to the scoring rubrics and comparable across the four forms of products since several key words in the rubrics were changed from last round. We anticipated that this revision of articulating the scoring rubrics would ultimately improve the inter-rater reliability of the rubrics. Importantly, the formatting of the rubrics had been updated as rating scales with clearly defined descriptions for each level of performance.

We chose to analyze the technical quality for only three of the four rubrics: Technical Documentation, poster display, and product presentations. The mission rubric was believed to yield a much higher reliability because it directly evaluates whether the vehicle performed the tasks or it did not, which less likely involves subjective judgments compared to other three rubrics. The Poster rubric and product presentation rubric in our prior analysis were found to yield a reliable scoring process as using two raters led to absolute Generalizability coefficients larger than .89 for any pair of judges. In this report, we further evaluate their psychometric quality to determine whether judges who might differ from previous years were able to apply the rating scale format reliably. In addition, we chose to evaluate the amount of measurement error for the Technical Documentation rubric which was found more challenging for judges in our last year's report. For each of these rubrics, experienced engineers were assigned into triads or pairs and each team then was randomly assigned to rate a group of eight to ten projects.

We employed both the classical test theory (CTT) and Generalizability theory to examine the inter-rater reliabilities. For the CTT, we computed the Cronbach's alphas. For the G theory, we used the random model of a project x judge design, denoted as p x j. This should result in three variance components: variance due to product differences as the object of the measurement (σ_p^2) as well as two sources of measurement errors, variance due to the judge effect (σ_j^2), and variance due to the product by judge interaction confounded with residual ($\sigma_{pj,e}^2$). We also calculated two types of G coefficients. Absolute G *coefficients indicate the extent to which the judges provided comparable scores or not when evaluating projects; in contrast, relative G coefficients refer to the extent to which the judges ranked the projects consistently or not regardless if some might give much higher scores compared to their peer judges. In the one-facet model of p x j design, relative G coefficients are exactly the same as the Cronbach's alpha coefficients under the CTT approach. Relative G coefficients are more relevant for this project because score reliability only matters regarding how the judges ranked the projects submitted by the competition teams. As we aim to align the specifics of each rubric to the NGSS, it will be important to take the absolute G coefficients into account. Absolute G coefficients allow us to examine the reliability of score inferences when we evaluate how well projects demonstrate participants' mastery of the targeted NGSS standards, especially the ones relevant to the engineering practices.*

Table 1 summarizes the variance components and G coefficients for the total scores for the three rubrics we studied. Our interpretations of results primarily focus on the relative G coefficients. Both poster display and product presentation rubrics were relatively more reliable with coefficients of .970 and .950 respectively compared to the technical report rubric with a coefficient of .885. For these two rubrics, the relative G coefficients mostly remained .80 or greater even when randomly selecting any one judge instead of using two or three judges. For example, among the seven scoring teams, any one judge randomly selected from the six teams could score the product presentation rubric with fewer than 80% of score variation accounted for by the measurement error associated with the judge or rater facet. In other words, only one rater is needed because of the small amount of measurement error involved for both poster display and product presentation rubrics.

Similarly, the technical documentation rubric was reliably applied by the judges as well. All the relative G coefficients were satisfactory with magnitudes greater than or closer to .80 when using at least 2 judges. For this rubric, using only one judge may introduce a sizeable amount of measurement error for scores assigned.

Importantly, for the three rubrics, the reliability coefficients varied greatly across the triads of judges. For Teams 6 (who was assigned to the technical documentation rubric), 10-14 (who were assigned to poster display rubric), and 15-20 (who were assigned to the product presentation rubric), each triad of judges scored consistently the projects to such a degree that only one of the judges would be needed while still maintaining an interrater reliability above .90. Judges within each of these teams scored the ROV projects more consistently compared to how their peer judges from other teams. However, teams who were assigned to the technical documentation rubric more likely disagreed with each other when evaluating the performance of ORV teams. Overall all teams appeared to reach an acceptable level of inter-rater reliability regardless the type of rubric they were assigned to, in contrast to the observation of previous year that at least two teams of judges completely disagreed with each other. This finding could be explained with two possible reasons: (1) the selection and matching of the judges were carefully handled by the project team. Judges were invited with a conference call to share with insights and comments related to the use and interpretation of rubrics as well as rationale for assigning different *scores for example projects. And (2) the use of rating scale format allowed judges to focus on fewer yet bulleted criteria for each level of performance.*

Table 1. Estimated Variance Components and G Coefficients for the Total Scores of Three Rubrics: Technical Report, Poster Display, and Product Presentation

ID of Judge Team[a]	Estimated Variance Components In P x J (Project x Judge) Design			G Coefficients Using 1 Judge (n_j=1)		Using 2 Judges (n_j=2)		Using 3 Judges (n_j=3)	
	EVC (project)	EVC (judge)	EVC (pj,e)	Abs. G Coeff. [b]	Rel. G Coeff. [c]	Abs. G Coeff.	Rel. G Coeff.	Abs. G Coeff.	Rel. G Coeff.
Technical Documentation									
1	183.26	55.73	91.64	0.554	0.667	0.713	0.800	0.789	0.857
2	123.49	142.51	63.35	0.375	0.661	0.545	0.796	0.643	0.854
3	117.51	184.16	45.17	0.339	0.722	0.506	0.839	0.606	0.886
4	283.06	0.00	98.57	0.742	0.742	0.852	0.852	0.896	0.896
5	139.96	36.55	68.95	0.570	0.670	0.726	0.802	0.799	0.859
6	280.09	43.88	60.34	0.729	0.823	0.843	0.903	0.890	0.933
7	129.37	16.13	64.04	0.617	0.669	0.763	0.802	0.829	0.858
Pooled[d]	1256.74	478.95	492.06	0.564	0.719	0.721	0.836	0.795	0.885
Poster display									
8	31.75	0.00	9.28	0.774	0.774	0.873	0.873	0.911	0.911
9	12.29	0.58	8.67	0.571	0.586	0.727	0.739	0.799	0.810
10	25.88	0.82	5.68	0.799	0.820	0.888	0.901	0.923	0.932
11	139.14	0.00	0.39	0.997	0.997	0.999	0.999	0.999	0.999
12	39.13	0.00	6.25	0.862	0.862	0.926	0.926	0.949	0.949
13	60.56	0.00	1.89	0.970	0.970	0.985	0.985	0.990	0.990
14	61.75	0.00	1.99	0.969	0.969	0.984	0.984	0.989	0.989
Pooled	370.487	1.399	34.146	0.912	0.916	0.954	0.956	0.969	0.970
Product presentation									
15	91.37	106.61	21.06	0.417	0.813	0.589	0.897	0.682	0.929
16	165.65	10.71	28.11	0.810	0.855	0.895	0.922	0.928	0.946
17	260.82	1.73	33.64	0.881	0.886	0.937	0.939	0.957	0.959
18	250.34	49.21	47.09	0.722	0.842	0.839	0.914	0.886	0.941
19	163.41	2.45	17.11	0.893	0.905	0.944	0.950	0.962	0.966
20	143.36	3.29	5.05	0.945	0.966	0.972	0.983	0.981	0.988
21	88.75	31.35	32.21	0.583	0.734	0.736	0.846	0.807	0.892
Pooled	1163.70	205.35	184.26	0.749	0.863	0.857	0.927	0.900	0.950

Notes: a. Both of technical report and sales representation rubrics were scored by three judges independently whereas poster display rubric was mostly scored by a pair of judges.

b. Abs. G coeff. refers to absolute G coefficient.

*c. Rel. G Coeff. refers to relative G coefficient. Relative G coefficient when **using 3 judges** was equivalent to the Cronbach's alpha obtained.*

d. Pooled estimates were based on the sum of the estimated variance components. The estimates indicate the averaged reliabilities across multiple judge triads.

In order to understand why the technical report rubric was much more demanding to maintain a reasonable interrater reliability, we took a closer look at the reliabilities of sub-categories (also called as sub-scales that judges needed to assign scores to). We chose to examine the four sub-categories because they were considered as the most relevant in comparison to other ones during the content analysis of the rubric and we had evaluated their technical quality in previous years. Table 2 reports the estimated variance components from the G studies and the G coefficients for running the decision studies when using different number of judges. We continue to summarize the patterns of relative G coefficients across multiple judge teams in Table 3.

As shown in Tables 2 and 3, the reliabilities were found uneven across the sub-categories and across the judge triads. Generally, the G coefficients for sub-scales yielded much lower reliabilities compared to the total score because the latter should contain more score variations as the sum of all the sub-categories. It appeared that only the sub-category of teamwork was still challenging for many of the scoring triads to score consistently. At least half of the seven triads of judges were unable to maintain a relative G coefficient of .75; indeed, three judge triads were fail to reach a relative G coefficient of .40 (see Table 3). In contrast, the remaining three sub-categories were relatively more attainable for the judges because often just one or two triads had experienced some difficulties in maintaining a reasonable reliability coefficient.

Interestingly there was not much improvement in terms of how the rubrics behaved when we compared the G coefficients of the fours sub-categories to the findings of previous years. The coefficients continued to be mostly comparable across the last two years' competitions, with an average relative G coefficient across all the teams and pooled as .668 (SD=.231) for the 2016 competition in contrast to .675 (SD=.234) for the 2015 competition.

Lastly, scoring teams had different level of interrater reliability; that is, members in some teams more likely ranked the projects in a similar way than their colleagues in other teams. The judges within Triads 1, 3, and 4 tended to have a stable interrater reliability virtually for all the sub-categories of the technical documentation rubric. The judges in Triads 2 and 7 had the lowest interrater reliability across multiple sub-categories. The judges were able to score constantly for most sub-categories only except the vehicle systems sub-rubric. This finding probes the research team to explore better ways for ensuring the scoring reliability. For instance, although the triads' scoring performance could be potentially improved by providing additional targeted training support for some sub-categories, it is truly hard to predict how judges may agree or disagree on a specific sub-rubric. The inconsistency found on some sub-rubrics sometimes could not be explained by the content of those sub-rubrics; in other words, judges may disagree with their peers even on some easy sub-rubrics for unknown reason. To address this issue, probably judges may need more substantial training that tap into deep issues to check if they disagree with any fundamental principles that they tend to apply during their scoring processes. In addition, the scoring rubric can be formatted in a way to remind raters about these fundamental principles to determine the performance levels of 0 to 3 instead of offering too much detailed information which may result in the rater fatigue.

Table 2. Estimated Variance Components and G Coefficients for Important Sub-category Codes of the Technical Report Rubric

	Estimated Variance Components In P x J (Project x Judge) Design			G Coefficients					
				Using 1 Judge (n_j=1)		Using 2 Judges (n_j=2)		Using 3 Judges (n_j=3)	
Triad ID of Judges	EVC (σ_p^2)	EVC (σ_j^2)	EVC ($\sigma_{pj,e}^2$)	Abs. G Coeff.	Rel. G Coeff.	Abs. G Coeff.	Rel. G Coeff.	Abs. G Coeff.	Rel. G Coeff. [a]
Vehicle Systems									
1	0.53	0.57	0.73	0.291	0.422	0.451	0.594	0.552	0.687
2	0.70	0.36	0.93	0.351	0.428	0.519	0.600	0.619	0.692
3	0.89	0.51	1.36	0.322	0.394	0.487	0.566	0.587	0.661
4	0.81	0.28	1.39	0.326	0.367	0.492	0.537	0.592	0.635
5	2.09	0.03	0.64	0.758	0.766	0.862	0.867	0.904	0.908
6	0.10	0.38	0.88	0.072	0.100	0.134	0.182	0.188	0.250
7	0.40	0.19	0.51	0.366	0.443	0.536	0.614	0.634	0.705
Pooled	5.51	2.32	6.43	0.386	0.461	0.557	0.632	0.654	0.720
Teamwork									
1	1.02	0.15	0.51	0.604	0.664	0.753	0.798	0.821	0.856
2	0.00	0.37	0.69	0.000	0.000	0.000	0.000	0.000	0.000
3	0.26	1.94	0.28	0.105	0.479	0.190	0.648	0.260	0.734
4	0.47	0.00	0.66	0.414	0.414	0.586	0.586	0.680	0.680
5	0.14	0.47	0.69	0.108	0.169	0.194	0.289	0.266	0.379
6	0.96	0.39	0.79	0.447	0.548	0.618	0.708	0.708	0.784
7	0.26	0.00	1.18	0.180	0.180	0.305	0.305	0.397	0.397
Pooled	3.11	3.33	4.82	0.276	0.392	0.433	0.563	0.534	0.659
Overall Presentation									
1	4.53	1.28	5.46	0.402	0.454	0.574	0.624	0.669	0.714
3	2.31	1.74	5.08	0.253	0.312	0.404	0.476	0.504	0.577
4	5.79	2.88	7.66	0.355	0.431	0.524	0.602	0.623	0.694
6	8.64	0.00	4.32	0.667	0.667	0.800	0.800	0.857	0.857
7	4.28	3.90	4.00	0.351	0.517	0.520	0.682	0.619	0.763
1	8.49	1.92	4.68	0.563	0.645	0.720	0.784	0.794	0.845
2	4.70	0.09	3.45	0.571	0.577	0.727	0.732	0.800	0.804
Pooled	38.74	11.80	34.64	0.455	0.528	0.625	0.691	0.715	0.770
Safety									
1	2.10	0.00	0.97	0.683	0.683	0.812	0.812	0.866	0.866
2	1.30	0.62	1.13	0.426	0.534	0.598	0.696	0.690	0.775
3	0.59	1.51	1.28	0.175	0.317	0.299	0.481	0.390	0.582
3	1.65	0.03	0.76	0.675	0.684	0.806	0.812	0.862	0.867
7	0.89	0.49	0.81	0.405	0.522	0.577	0.686	0.671	0.766
3	2.89	0.00	0.52	0.847	0.847	0.917	0.917	0.943	0.943
4	0.00	0.00	1.46	0.000	0.000	0.000	0.000	0.000	0.000
Pooled	9.43	2.65	6.94	0.496	0.576	0.663	0.731	0.747	0.803

Note: a. Relative G coefficients in the range of .400-.700, when three raters are randomly selected to form a scoring team, are in orange. Relative G coefficients below .400 are in red.

Table 3. Summary of Relative G Coefficients for Sub-categories of the Technical Report Rubric

					Summary of Number of Rel. G Coeff. (ρ^2) across 9 Sub-categories		
Triad ID of Judges	Vehicle Systems	Overall Presentation	Teamwork	Safety	Number of $\rho^2 \geq .40$	Number of $\rho^2 \geq .60$	Number of $\rho^2 \geq .75$
1	0.687	0.856	0.714	0.866	4	4	2
2	0.692	0.000	0.577	0.775	3	2	1
3	0.661	0.734	0.694	0.582	4	3	0
4	0.635	0.680	0.694	0.867	4	4	1
5	0.908	0.379	0.857	0.598	3	2	2
6	0.250	0.784	0.857	0.943	3	3	4
7	0.705	0.397	0.763	0.000	2	2	1
Max score point	6	12	6	6			
Relevance	High	High	High	High			
Summary of patterns of Rel. G Coeff. (ρ^2) Across Seven Triads							
Number of $\rho^2 \geq .40$	6	4	7	6			
Number of $\rho^2 \geq .60$	6	4	6	4			
Number of $\rho^2 \geq .75$	1	2	3	4			

Overall, the reliability analysis based on the CTT and G studies provide supportive evidence for the psychometric quality of the three rubrics that required subjective appraisal of judges: technical documentation, poster display, and product presentation. For all the three rubrics, our analysis indicates that all judge teams were able to successfully maintain a relative G coefficient at least .80 with three judges and at least .74 even with two judges. This piece of finding establishes the empirical evidence for choosing only two judges to score the poster display. Still the research team decided to keep three raters for the other two rubrics because of the technical issue for sub-rubrics. The detailed scores from these sub-rubrics can provide rich information for redesigning the competition handbook, guidelines, training materials, and scoring rubrics. Therefore, the benefit of obtaining scores and comments from additional judges for formative feedback to the research team, instructional support, and participants outweighs the cost of recruiting and scoring.

A detailed analysis with the sub-categories for the technical documentation rubric revealed that the scoring inconsistency mostly was caused by a few challenging rubrics and a couple of teams of judges who tended to score very differently. The G studies with the sub-categories of the product presentation rubric offered three insights for improving the interrater reliability in next project year and any possible future use:

(1) Triads 2 and 7 definitely need additional training to get familiar with the rubrics and close monitoring of their scores in the upcoming competition events. A follow-up analysis confirmed that the inconsistency was caused by all of the three judges rather than the situation that two judges scored reliably but differed from the other judge. The judges may need different types of training activities to discuss the rubrics or the underlying principles and assumptions that they have chosen to apply.

(2) This relatively newer version of the technical report rubric has been revised based on the findings from last year, by using well-articulated language, breaking down and elaborating multiple facets for complex sub-categories if needed, and aligning the detailed descriptions for each performance level (see Figure 1 for an example of the rubric layout). Overall many of the sub-categories were found with acceptable interrater reliability after the clarity and easiness of the rubric was improved during last year. Among the four important sub-categories (i.e., Vehicle Systems, Overall Presentation, Teamwork, and Safety), the overall presentation sub-rubric is the only candidate which might need some tinkering because several of the seven triads of judges failed to apply this rubric consistently, with their interrater reliability coefficients as .40 or lower.

2016 MATE ROV Competition Technical Documentation Rubric **Judge:______________________________**

Category	**Scoring Criteria**				**Points**
Overall Presentation	**3 - Excellent**	**2 - Very Good**	**1 - Good**	**0 – Poor or missing**	
Basic requirements	Technical documentation is 25 pages or less; includes a table of contents; all measurements are in SI units (exceptions include ½ PVC, etc.); excellent attention to grammar; title page includes all elements as specified in the guidelines: Company name, school, club or organization name, city and state, members and roles, name of mentor	Technical documentation is 25 pages or less; includes a table of contents; most measurements are in SI units (exceptions include ½ PVC, etc.); very good attention to grammar; title page includes most elements as specified in the guidelines: Company name, school, club or organization name, city and state, members and roles, name of mentor	Technical documentation is over or less than 25 pages; includes an inaccurate table of contents; some measurements are in SI units; good attention to title page includes some elements as specified in the guidelines: Company name, school, club or organization name, city and state, members and	Technical documentation is over or significantly under 25 pages; table of contents missing or inaccurate; measurements not SI units; poor attention to grammar, many typos, etc.; many specified elements of the title page missing	
Abstract	250 words or less and provides an excellent, clear and concise summary of work	250 words or less and provides a concise summary of work	250 words or less and provides an adequate summary of work	250 words or less but is not clear nor concise	
Understanding of ROV	Clearly describes how the vehicle was designed, clear understanding of the technical and scientific concepts behind designing and building the vehicle	Describes how the vehicle was designed, demonstrates an understanding of the technical and scientific concepts behind designing and building the vehicle	Issues with the description of how the vehicle was designed, demonstrates some understanding of the technical and scientific concepts behind designing and building the vehicle	Poorly written, information missing, does not demonstrate or capture in any way an understanding of the technical and scientific concepts behind the vehicle	
Photos of ROV	Photo of complete vehicle included, includes additional photos which fully capture vehicle design, excellent captions accompany photos, also includes an excellent mechanical drawing or sketch	Photo of complete vehicle included, includes additional photos which somewhat capture vehicle design, captions accompany photos, also includes a mechanical drawing or sketch	Photo of complete vehicle included, no additional photos or additional photos which do not capture vehicle design, captions accompany photos, also includes an adequate mechanical drawing or sketch	Photos missing or not of high quality, captions missing and mechanical drawing or sketch missing or of very poor quality	

Figure 1. One page of the scoring rubric for the product presentation, sub-rubric of overall presentation

(3) Revisions for the overall presentation sub-rubric should be waited after we conduct additional analysis related to construct validity. It remains unclear whether the change will be just formatting by creating bullets of the performance descriptions for each level or by clarifying the criteria or

adding more relevant descriptions. The planned factorial analysis will help us pinpoint how the sub-categories relate to the underlying construct intended to measure. We also plan to conduct a cognitive interview with a small sample of judges to unpack their reasoning processes when they review projects and assign scores as well as possible difficulties or confusions they may experience when applying the rubrics. With new findings we may recommend to collapse or drop a couple of sub-categories to simplify the cognitive process of making scoring decisions.

Improving the Internal Consistency of Existing Surveys

Building on last two years' analysis of the internal structure and consistency of four existing surveys based on the data collected from all the competition events in 2014 and 2015, we were able to refine the survey questions of Student Survey and Parent Survey because both Instructor Survey and Volunteer Survey were found with satisfactory measurement properties. The revision comments for Student Survey and Parent Survey had focused on four issues: (1) change items that did not behave well in the Cronbach's alpha by clarifying the wording of the stem or the labels of the Likert scale or even drop items when they overlapped significantly with others; (2) use similar layouts to present survey questions and the Likert scale across the four surveys; (3) make the use of words and phrases consistent across all the four surveys so that potentially the response patterns can be compared; and (4) revisit and revise items that were grouped around one common stem as multiple statements when appropriate or create new set to organize items into such a set.

For the survey data collected in 2016, we again evaluated the technical quality of each of the surveys. We first examined the descriptive statistics to make sure that responses were properly recoded if needed and checked if items may involve a usually high non-response rate (which was one indication that the wording of the survey question might be confusing or ambiguous to respondent) or may have too small standard deviation that can potentially cause poorly behaving items due to the restricted range of scores. We then performed both exploratory factor analysis (EFA) and confirmation factor analysis (CFA) to verify the internal structure of the construct. Lastly, we calculated the Cronbach's alpha coefficients as indicators for internal consistency for the sub-scales corresponding to the factor model identified by the factor analyses. In what follows we only report the key findings of the psychometric analysis based on the CFA and Cronbach's alpha.

Student Survey. Based on the χ^2 value, we recommend the 3-factor model because it has a much lower χ^2/df ratio and slightly statistically better than the other two factor models (see Table 4 for fitness statistics and Figure 1 for factor loadings for each tested model). Still this 3-factor model is not statistically robust which means there is room to improve the model fit and factor loadings of the survey items.

Table 4. Student Survey Items: Good-of-fit Indices for 1-, 3- and 4-factor Models (n = 2164)

Model	Chi-square	df	CFI	TLI	RMSEA
1-factor	10805.19**	189	0.94	0.94	0.16
3-factor	3941.87**	186	0.98	0.98	0.10
4-factor	3792.69**	183	0.98	0.98	0.10

*** p < .01*

As shown in Table 4, the factor loadings and R-squared statistics of the survey questions provide supportive evidence that the three factors are distinct constructs, differing from each other. The number of items belongs to Factors 1, 2, and 3 is 8, 6, and 7, respectively. Factor 1 items mainly measure student's interest in the STEM field; the six Factor 2 items mainly assess student's skills and knowledge in the STEM field; and Factor 3 items mainly focus on measuring how student can communicate well in the STEM field. Table 5 reports the factor loading based on the three-factor CFA model.

Table 5. Student Survey Items: Statistics of Observed Variables based on the Three-factor CFA Model (n = 2164)

	Factor Loading			
Rubrics	*1*	*2*	*3*	*R-squared*
Q03A	*0.73***			*0.54*
Q03B	*0.79***			*0.63*
Q03C	*0.80***			*0.64*
Q03D		*0.74***		*0.55*
Q03E			*0.70***	*0.49*
Q04A	*0.81***			*0.65*
Q04B	*0.75***			*0.56*
Q04C	*0.71***			*0.50*
Q04D	*0.66***			*0.44*
Q04E	*0.73***			*0.53*
Q05A		*0.85***		*0.72*
Q05B		*0.83***		*0.69*
Q05C		*0.76***		*0.56*
Q05D		*0.84***		*0.70*
Q05E		*0.60***		*0.36*
Q06A			*0.81***	*0.65*
Q06B			*0.77***	*0.59*
Q06C			*0.89***	*0.79*
Q06D			*0.89***	*0.79*
Q06E			*0.69***	*0.48*
Q06F			*0.75***	*0.55*

Table 6 below provides the detailed information for the underlying factors as we chose the 3-factor model. We report the internal consistency for the three factors (i.e., sub-scales) informed by the CFA. All of the three sub-scales yield satisfactory internal consistency as items tended to sufficiently capture the intended construct with less amount of variabilities due to the item variations compared to the previous versions of Student Survey.

Table 6. Descriptions of Sub-scales for Student Survey and Internal Consistency Statistics (i.e., Cronbach's alpha)

Factor	*Questionnaire Item IDs*	*Description of the Factor*	*Cronbach's Alpha Coefficient*
1	*Q03A, Q03B, Q03C, Q04A,Q04B, Q04C, Q04D, Q04E*	*More interested in STEM field because of the ROV project*	*0.87*
2	*Q03D, Q05A, Q05B, Q05C, Q05D, Q05E*	*Increased skills and knowledge in STEM field because of the ROV project*	*0.83*
3	*Q03E, Q06A, Q06B, Q06C, Q06D, Q06E, Q06F*	*Better communicate in STEM field because of the ROV project*	*0.88*

<u>Parent, Instructor, and Volunteer Surveys</u>. Following the same procedure described above, we examined the internal structure of the three other surveys based on the content review of all the survey items, using both EFA and CFA modeling. It is important to reiterate that only Parent Survey had gone through the revision process based on last year's findings whereas the other two surveys appeared to demonstrate acceptable psychometric quality.

Table 7 reports the CFA model fits in the form of χ^2 statistics and recommends one model for each survey that yields the best model fit and theoretically aligns with our assumptions of the underlying constructs. The factor loading for the items appeared satisfactory, which offers another piece of evidence for the hypothesized internal structure of these constructs.

Table 7. CFA Model Fit for Parent, Instructor, and Volunteer Survey Items

CFA Model	*Chi-square*	*df*	*CFI*	*TLI*	*RMSEA*	*Selected Model*
Parent Survey	*(n = 437)*					
1-factor	*2059.94***	*77*	*0.95*	*0.94*	*0.24*	
3-factor	*302.14***	*74*	*0.99*	*0.99*	*0.08*	*x*
Instructor Survey	*(n = 346)*					
1-factor	*588.54***	*44*	*0.92*	*0.91*	*0.19*	
2-factor	*337.71***	*43*	*0.96*	*0.95*	*0.14*	*x*
Volunteer Survey	*(n = 274)*					
1-factor	*1432.98***	*65*	*0.90*	*0.87*	*0.28*	
3-factor	*172.32***	*62*	*0.99*	*0.99*	*0.08*	*x*

For each survey, we continued to calculate the internal consistency of survey items for sub-scales for the recommended model based on the factor analysis results (see Table 8). All of the sub-scales used in these surveys were found with satisfactory internal consistency, even just with three to six items per sub-scale. The results, combined with findings from last competition, also confirmed that the technical quality of all the questionnaires is robust across different cohorts of participants as the statistics of their internal structure and Cronbach's alpha coefficients are similar between the last two competitions, despite the fact that different sample of participants responded the questionnaires.

Table 8. Descriptions of Sub-scales for Parent, Instructor, and Volunteer Surveys and Internal Consistency Statistics

Factor	Survey Item IDs	Description of the Factor	Cronbach's Alpha Coefficient
Parent Survey			
1	Q4A,Q4B, Q4C, Q4D	Grades improved because of the ROV project	.89
2	Q5A, Q5B, Q5C, Q5D	Children love the STEM field more because of the ROV project	.88
3	Q6A,Q6B,Q6C, Q6D, Q6E, Q6F	Children positively affected because of the ROV project	.90
Instructor Survey			
1	Q4A,Q4B, Q4C, Q4D, Q4E, Q4F	Comments on the effects of the ROV competition on student learning	. 84
2	Q5A, Q5B, Q5C, Q5D, Q5E	The ROV project is valuable to me	.80
Volunteer Survey			
1	Q3A, Q3B, Q5A,Q5B, Q5C	Overall comments on the ROV competition experiences	.85
2	Q6A, Q6B, Q6C	Positive comments on ROV program	.84
3	Q7A, Q7B, Q7C, Q7D, Q7E	Skills improved through ROV program	.87

Plan of Additional Psychometric Analysis

Informed by the research activities and analysis we have conducted, we will continue with the psychometric work by replicating the current data analysis procedures and incorporate new empirical evidence to evaluate the validity claims and plan for a measurement publication for the scoring rubrics used for the ROV competition.

<u>Validation of Competition Scoring Rubrics</u>. We will combine the multiple years of sample and apply CTT and G theory to evaluate how different configuration of expertise in scoring teams may influence the score variation. So far it is unclear why certain scoring teams could apply the scoring rubrics reliably and

why other teams were unable to do so. We plan to incorporate the background information of judges (such as amount of scoring experience, professional background, amount of ROV competition experience) into our modeling to understand whether some of variables can explain the scoring variations within a scoring team or on particular sub-rubrics.

Furthermore, we plan to conduct three validity studies with the rubric scores of poster, engineering, and technical report and spend our energy for the last study:

(1) *A set of factor analysis modeling with the sub-category scores. After we establish the interrater reliability, we will run the structural equation modeling (SEM) to perform the factor analysis to identify the internal structure for each rubric. This will allow us to determine whether some sub-categories can be collapsed (or dropped if the sub-category is difficult to maintain interrater reliability and does not contribute to the underlying constructs).*
(2) *Multi-trait multi-method (MTMM) correlation to explore the convergent and discriminant evidence of measuring various sub-scales of students' engineering performance. We hypothesize that the correlation patterns across the three rubrics (e.g., are the Design Rationale scores from both poster and technical report highly correlated?) can be empirically tested, which then can inform us to revise the three rubrics to ascertain what aspects of student performance can be captured by all the three rubrics or uniquely assessed only by one particular rubric.*
(3) *Cognitive interview with the judges via audio-taped conference calls. We will recruit the judges for the Product presentation and Technical Report rubrics to talk aloud their scoring decisions related two projects selected from a previous competition and provide comments about the rubric pages. We will ask each judge to explicitly explain why one judge gave a certain category, etc. The transcriptions will be analyzed to understand how judges made evaluative decisions when applying the rubrics, what evidence they chose to focus on, and which parts or pieces of rubrics caused inconsistency or difficulties in scoring. In addition to the research purpose, the transcribed excerpts of the cognitive interview can be augmented with annotations as training examples for new or returning judges.*

<u>Analysis of Existing Survey</u>. We will start to streamline four surveys, including student, parent, instructor, and volunteer, to ensure that information can be triangulated across multiple surveys by running regression analysis and SEM with latent variables in order to compare the response patterns. Examples of the research questions include: did students who participated in the ROV competition for multiple times report higher scores in academic outcomes than their peers who only participated once? Did students or instructors report the same obstacles for the ROV competition?

<u>Analysis of Knowledge Tests</u>. Our psychometric analysis has not taken a closer look at the knowledge tests used with teachers' professional development. If time allows, we would like to evaluate the psychometric properties of the knowledge tests. We envision that we can start with two important studies to address the reliability and validity issues:

(1) *Item analysis with both the pretest and posttest items to produce item parameters of difficulty index and discrimination index as well as the Cronbach's alpha. This will enable us to flag items that may have technical problems, e.g., too difficult to differentiate high performing examinees from those with little understanding.*

(2) *Potentially if the sample size allows, we can also execute the factor analysis to examine the dimensionality of the measured construct, the differential item functioning analysis to determine whether items may bias against particular sub-groups (e.g., gender, or socio-economic status), and the index of instructional sensitivity to evaluate whether items reflect the amount of opportunities to learn involved in the PD activities.*

***

Regional Workshops

PRE AND POST TEACHER WORKSHOP SURVEYS

Pre and post paper surveys were administered to teacher workshop attendees at the beginning of the workshop day and at the end of the training. The surveys addressed issues of teacher confidence facilitating STEM learning experiences, commitment to bringing a team to competition, concerns about mentoring students in designing and building an ROV, expectations of the workshops, and additional ways that the regional coordinators and the MATE Center could support the participants.

Summer Institute

IMMEDIATE FEEDBACK AND SIX-MONTH FOLLOW-UP SURVEYS

The evaluation of the Summer Institutes is a two-step process, collecting feedback from the participants immediately after the Institutes (using the Institute feedback surveys) then again a year later (using the Institute follow-up surveys). The follow-up surveys intend to measure the Institutes' longer-term impact and, in particular, to compare participants' actions once they returned to their classrooms with the intentions they had expressed at the close of the Institute. Because of the timing of the Summer Institute and the evaluation reporting, this evaluation covers the 2016 Summer Institute.

Other Data Sources

Additional data sources informing the evaluation include the annual reports turned in by the regional coordinators to the ITEST grant PI, observations of the Pacific Northwest regional competition, review of participation data, unsolicited letters sent from students, parents and teachers, website review and document review, including supporting technical materials and the MATE Center's annual report.

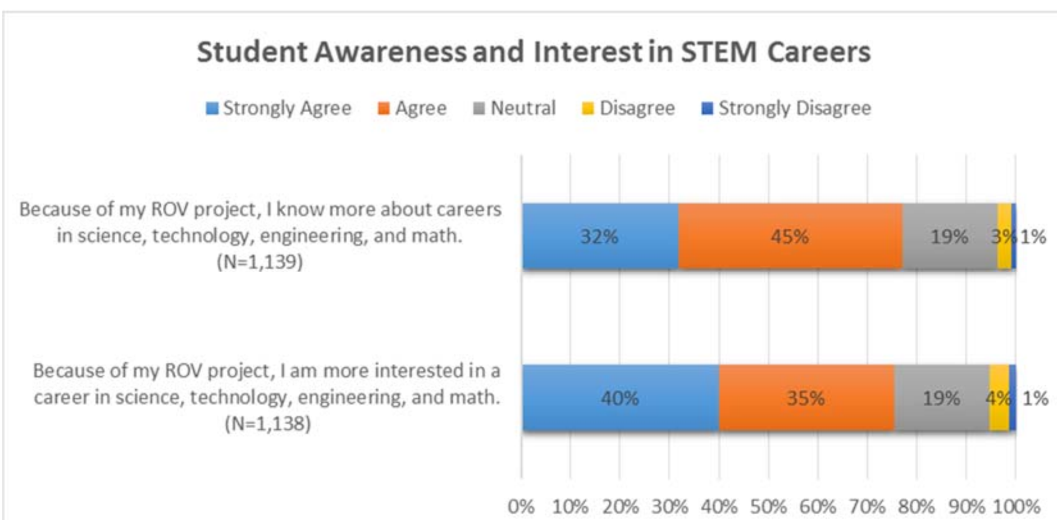
FINDINGS

Project Goal 1: Increase middle and high school students' interest in STEM and STEM careers, as well as their knowledge of STEM and understanding of how science and engineering work together to solve real-world problems

Evaluation Question(s) 1.1. To what extent did the MATE robotics activities lead to an increase in the students' interest in and knowledge of STEM content and STEM careers? Did educators and parents observe an increase in the students' interest in STEM content and STEM careers as a result of the robotics activities? An increase in the students' STEM knowledge and skills and 21st Century workplace skills?

Increased Awareness of and Interest in STEM Careers: In the post-competition surveys, over three-quarters of the students (77%, N=1,139) indicated that due to their ROV project, they knew more about careers in science, technology, engineering, and math (STEM), and three-quarters (75%, N=1,138) stated that their ROV project made them more interested in pursuing a STEM career.

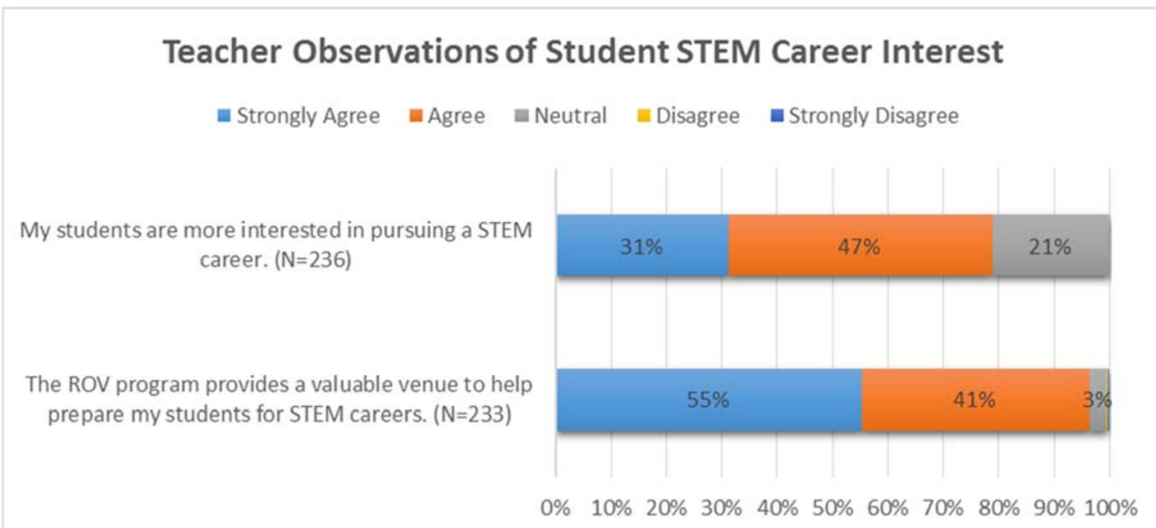
Figure 1: Effect of ROV Project on Students' Awareness and Interest in STEM Careers



Overall, 76% of the students (N=1,135) were interested in a STEM career; 20% were not sure, and 5% were not interested in a career in these fields. Students mentioned wanting careers such as computer engineer, marine scientist, robotics engineer, mechanical engineer, aeronautical engineer, environmental scientist, and astronaut. Students noted that their experience in the ROV program sparked their interest in STEM careers, with comments such as the following: "It has opened me up to other career opportunities because I am more educated in areas relating to engineering and science," and "Greatly influenced my choice in career."

Among the teachers/mentors who completed post-competition surveys, 90% (N=236) indicated that they had observed that their students were more interested in pursuing a STEM career since they began designing and building their ROVs. Ninety-six percent (96%, N=233) agreed that the ROV program provided a valuable venue to help prepare their students for STEM careers.

Figure 2: Teacher Observations of Student STEM Career Interest



Parents also noted an increased awareness and interest in STEM careers; 88% (N=281) agreed or strongly agreed that due to the ROV project, their child(ren) know more about STEM careers (53% strongly agreed, 35% agreed, 11% neutral, 0% disagreed, 0% strongly disagreed, and 1% don't know).

Also, 85% (N=277) agreed that participating in the ROV project has led their children to be more interested in pursuing a STEM career (53% strongly agreed, 32% agreed, 13% neutral, 1% disagreed, 0% strongly disagreed, 0% don't know). Parents described their children's interest in STEM careers in comments such as the following:

> *He has changed his career path- because of the program is studying electro mechanical tech engineering.*

Increased Interest in STEM: Eighty-three percent (83%) of the students (N=1,138) stated that their ROV project made them want to learn more about science, technology, engineering, and math (46% strongly agreed, 37% agreed, 14% neutral, 2% disagreed, and 1% strongly disagreed). As one student explained his experience, "This was a great learning experience, especially when things go wrong. This is highly recommended for people who like sciences and hands on projects. This is also a great program for those who are interested in learning basics in the science, technology, engineering and mathematic fields."

Students indicated that their ROV projects increased their desire to take courses in engineering (77%, N=1,139), science (73%, N=1,135), computer science (69%, N=1,129), math (64%, N=1,132), and other hands-on classes or club activities like robotics, electronics and shop courses (88%, N=1,129). (See *Figure 3* below.) One student saw the connection between the competition and his or her courses in the following quote:

It heightened my confidence in college engineering courses.

ROV Program Testimonials

Students

The MATE competition is a wonderful program that lets students find themselves and what they would like to do in life.

I'm confident I'm speaking for all my team members when I say that this has been the experience of a lifetime and is something we will never forget.

Parents

This opportunity was the best part of his high school years.

This has been a catalyst for: learning technology, learning about teamwork and collaboration, learning about projects.

It helped us as a family support our student's learning. Much better than textbook homework.

All the hard work and after school time was NEVER looked at as burdensome by my son. He enjoyed and even thrived in the environment and gained so much confidence and knowledge for the program.

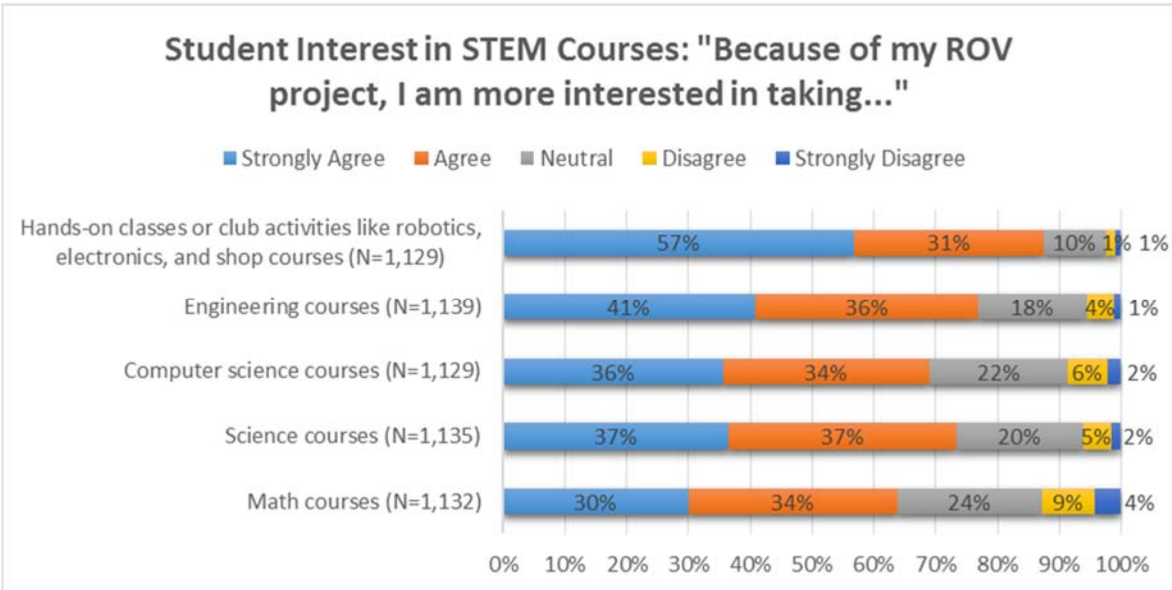
Faculty/Mentors

Excellent program for students of all ages.

The ROV program lets my students be recognized for their hard work in the STEM field. My students have been presented awards and scholarships.

In the post-competition survey, 90% of the teachers/mentors (N=236) indicated that their students were more interested in learning about science, technology, engineering and math (44% strongly agreed, 46% agreed, 10% neutral, 0 % disagreed, 0% strongly disagreed).

Figure 3: Effect of ROV Project on Students' Interest in STEM Courses



Parents concurred with the other sources reporting increased student interest in STEM. Ninety-four percent (94%) of the parents surveyed (N=280) stated that building an ROV has made their child more interested in learning about science, technology, engineering or math (59% strongly agreed, 35% agreed, 5% neutral, 0% disagreed, 0% strongly disagreed).

Increased STEM Knowledge and Skills: In the post-competition surveys, students reported increased skills and knowledge due to their ROV project in several subjects: engineering (90%, N=1,131), technology (89%, N=1,125), science (79%, N=1,129), the competition theme (75%, N=1,114), and math (61%, N=1,124). Students noted their increased STEM skills in comments such as the following:

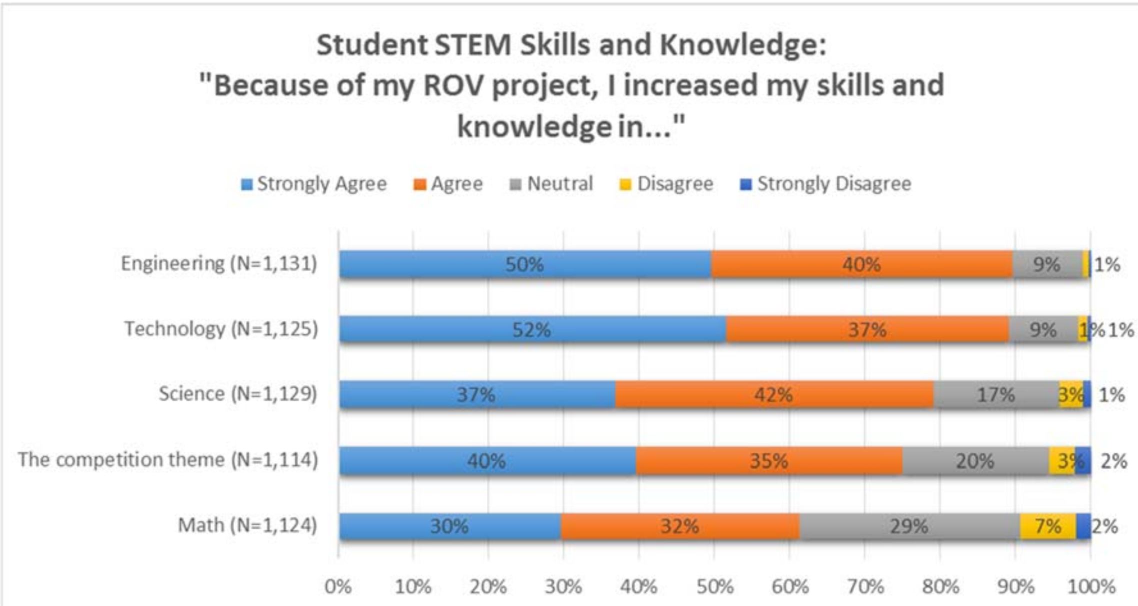
> *I have learned a lot of things in relation to working as a team and getting things done before the due date, as well as more education in the math, science, and engineering field. I hope that this competition can continue to bring many more students into STEM related fields.*

> *Fantastic. Learned a lot about designing and coding!*

> *I am extremely interested and excited about marine science because of this competition.*

I loved competing in the MATE competition! I learned both technical skills and leadership skills, and it inspired me to pursue a career in engineering.

Figure 4: Effect of ROV Project on Students' STEM Skills and Knowledge



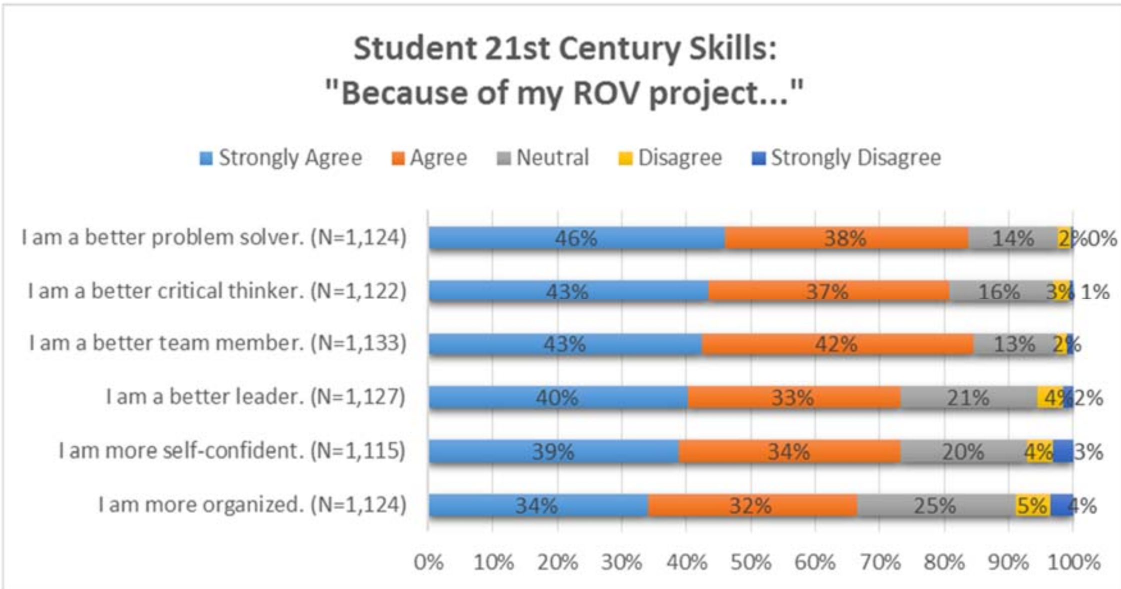
Among the teachers/mentors who completed post-competition surveys (N=236), 98% of the respondents reported that they observed improvements in their students' STEM knowledge and skills (60% strongly agreed, 38% agreed, 2% neutral, 0% disagreed or strongly disagreed).

Parents reported that building an ROV contributed to improving their children's grades in engineering/robotics (47%, N=276), science (55%, N=279), math (48%, N=278) and computers (43%, N=276).[2]

Increased 21st Century Skills: Students reported that participating in the ROV project improved their problem solving (84% agreed or strongly agreed, N=1,124), teamwork (85%, N=1,113), critical thinking (81%, N=1,122), leadership (73%, N=1,127), and organization skills (66%, N=1,124). Students also reported increased self-confidence (73%, N=1,115).

[2] Percentages are calculated among students studying each topic.

Figure 5: Effect of ROV Project on Students' 21st Century Skills



In responses to open-ended survey questions, students also described gaining 21st Century skills through their experiences building an ROV, such as the following:

> *I learned perseverance, teamwork and critical thinking skills.*
>
> *I think ROV has taught more than just academics, it also taught teamwork and cooperation in the business world.*
>
> *ROV has strengthened my ability to make connections and to work w/ other ideas/opinions.*

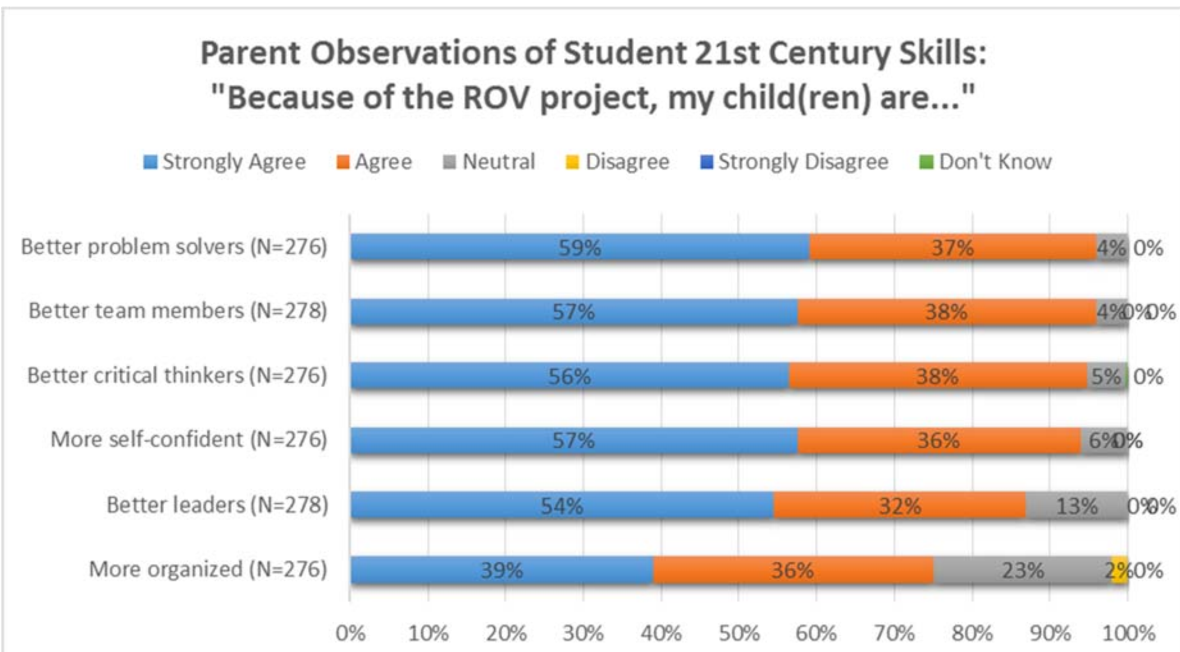
In the post-competition surveys, 97% of the teachers/mentors (N=235) mentioned that they observed increases in their students' skills in team building, problem solving, and/or critical thinking (54% strongly agreed, 43% agreed, 3% neutral, 0% disagreed, 0% strongly disagreed). Teachers/mentors saw skill development in many areas, as evidenced by their written comments:

> *This program truly helps with troubleshooting, problem solving skills, team cooperation and confidence to speak in front of a large audience.*
>
> *Wonderful program that teaches all the lessons needed to build a student's grit, problem solving, collaboration, teamwork and technical skills.*

Parents were asked about their observations of changes in their children due to the ROV program. The vast majority agreed or strongly agreed that because of the ROV program, their children were better problem solvers (96%, N=276), critical thinkers (94%, N=276), team members (95%, N=278), and/or leaders (86%, N=278). Ninety-three percent (93%) of the parents (N=276) reported increased self-confidence in their children. In the open-ended comments, parents noted other changes that they observed in their children, including public speaking, leadership, prioritizing, working under pressure, resiliency, focus, time management, and self-confidence. Comments in this theme include the following:

> *My daughter got to learn more about herself which will help her in the future with leadership. Her self-confidence soared.*
>
> *He is driven and excited about ROV but also about his future. He knows he needs to excel in math and science to do what he wants to do in life.*
>
> *Maturity, critical thinking, independence, "thinking outside the box", problem-solving.*
>
> *This gave my son confidence and a chance to feel part of a team.*

Figure 6: Parent Observations of Effect of ROV Competition on Students' 21st Century Skills



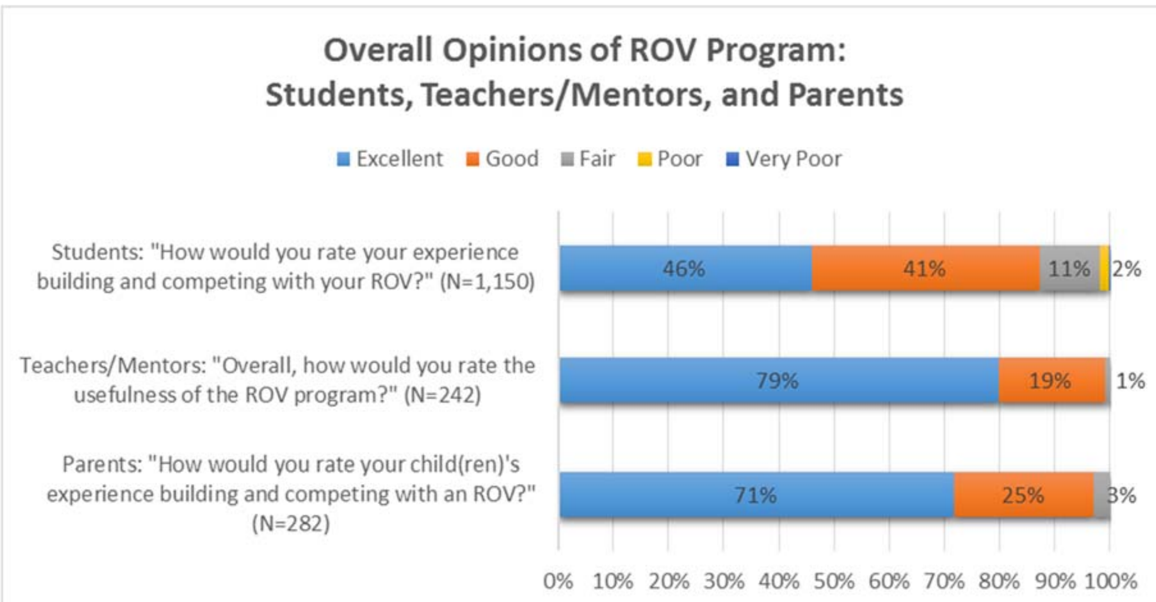
Overall Opinions of ROV Program: Overall, students (N=1,150) rated their experiences building and competing with their ROV very positively, with 46% rating their experience as excellent, and 41% providing a rating of good. Eleven percent (11%) thought their experience was fair. Two percent (2%) gave the experience a poor rating, and 0.3% rated it as very poor. (See Figure 7) In the post-competition surveys, students wrote comments such as the following:

> *The experience, knowledge, skills, and friendships that have been gained and developed through this program have proved invaluable and life changing.*

> *This is a very educational experience that allows many fun opportunities to meet new people/learn new criteria. Helps to grow skills.*

> *Great program, great to be a part of, had lots of fun. Great starting place for anyone interested in a career in robotics.*

Figure 7: Overall Opinions of the ROV Program from Students, Teachers/Mentors, and Parents



Teachers/mentors (N=242) gave nearly uniformly positive ratings of the usefulness of the competition, with 79% stating that it was excellent and 19% indicating that it was good (1% rated the competition as fair.) Teachers/mentors also rated the support provided by the MATE program highly (50% excellent, 38% good, 10% fair, 3% poor, and 0% very poor). Teachers/mentors stressed the importance of the program in comments such as the following:

The competition was so valuable, adding lessons in soft skills that the classroom/build time might not touch on.

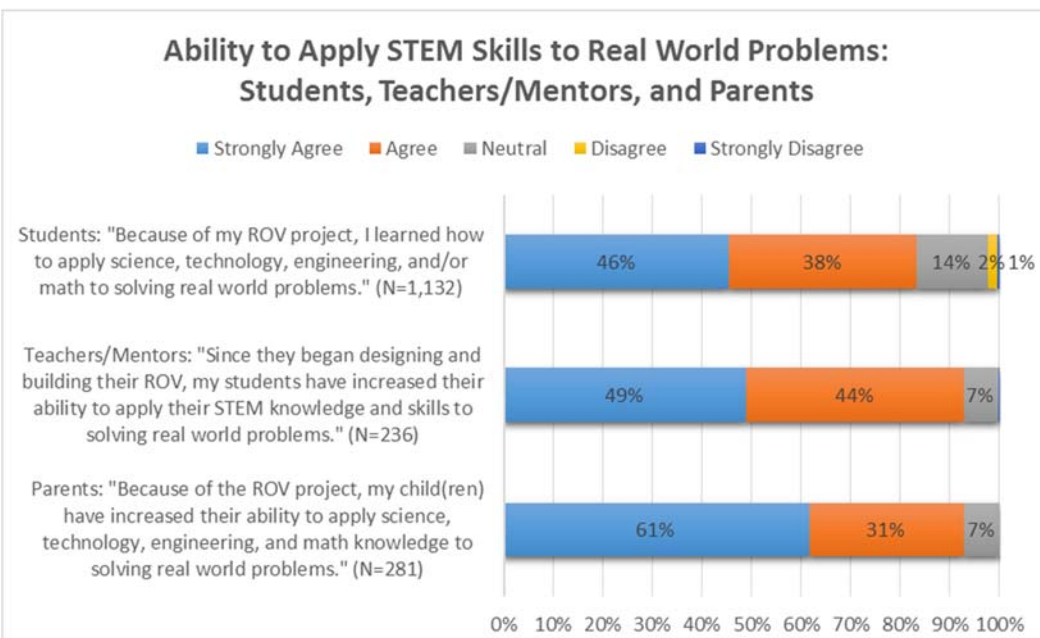
MATE is a highly valuable program for our students. MATE engages our students with practical engineering in a way that no other program or organization offers.

Overall, parents gave extremely positive ratings to their children's experience building and competing with an ROV. Seventy-one percent (71%, N=282) rated it as excellent, 25% gave a rating of good, 3% marked fair, and 1% were not sure. When asked whether the competition has been valuable for the educational development of their child, 73% strongly agreed that it was (N=278), 24% agreed with the statement, and 2% were neutral. No respondents disagreed or strongly disagreed.

Evaluation Question 1.2. How did the robotics activities affect students' ability to apply STEM knowledge and skills to finding solutions to real-world problems?

In the post-competition surveys, 83% of the students (N=1,132) indicated that participating in the ROV project helped them learn to apply STEM to real world problems. Ninety-three percent (93%) of the teachers/mentors (N=236) observed improvements in their students' ability to apply STEM knowledge and skills to real world problems, as did 92% of the parents (N=281). (See Figure 8)

Figure 8: Effect of ROV Program on Ability to Apply STEM Skills to Real World Problems: Students, Teachers/Mentors, and Parents



Students recognized the connection between the competition and real-world application of their science and technology skills in comments such as the following:

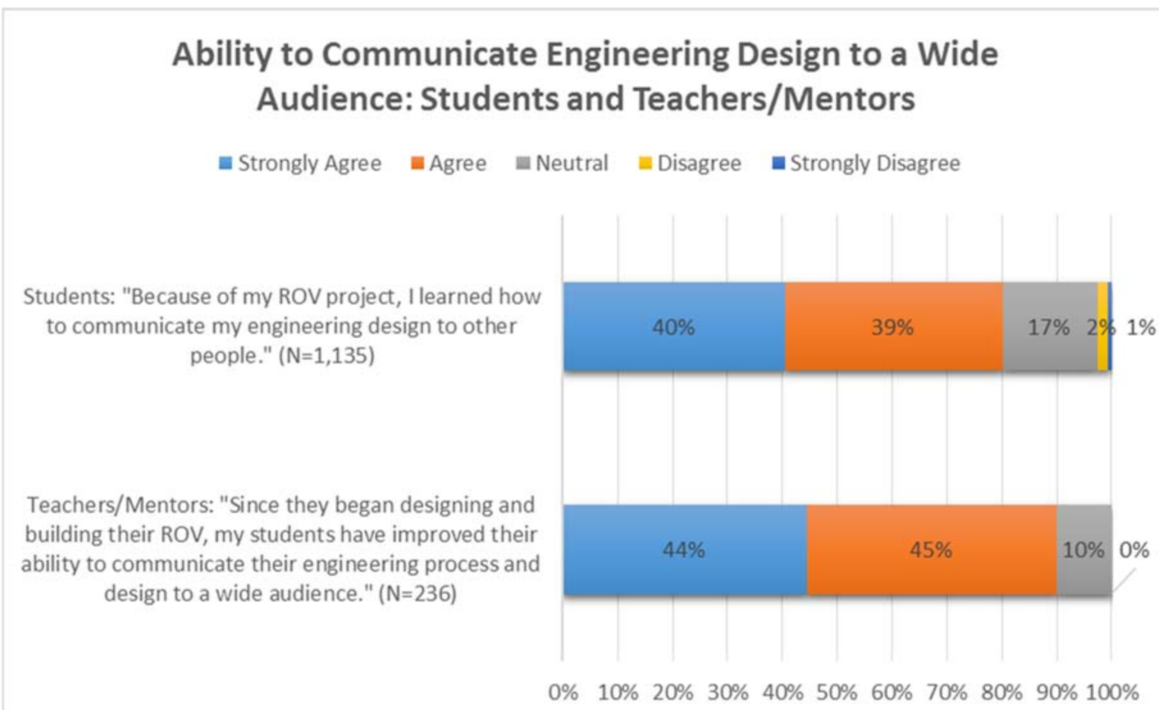
> *The MATE program simulated a very realistic work environment, which made it an especially unique experience.*

Evaluation Question 1.3. How did the robotics activities affect students' ability to communicate their engineering process and designs to a wide audience (from engineers and technicians to the general public)?

Seventy-nine percent (79%) of the students in the post-competition surveys (N=1,135) stated that participating in the ROV project helped them learn how to communicate their engineering design to other people.

Eighty-nine percent (89%) of the teachers/mentors (N=236) indicated that their students had improved their ability to communicate their engineering process and design to a wide audience. (See Figure 9)

Figure 9: Effect of ROV Program on Ability to Communicate Engineering Process and Design



Evaluation Question 1.4. How did participation in the robotics activities influence students' educational and career paths?

There are two sources of educational and career information on former competition participants: 1) the ROV Competition Student Alumni Survey and 2) the National Student Clearinghouse match.

ROV COMPETITION STUDENT ALUMNI SURVEY

ROV Competition Student Alumni Survey results include the following:

- Among the 220 alumni who earned a college degree, 85% earned a degree in a STEM discipline.
- Among the 236 current college and university students, 85% are studying towards a STEM degree.
- Among the employed alumni (N=320), 73% are currently working a STEM-related job, and 22% currently or previously worked a job related to ROVs or other underwater technologies.
- Two-thirds (67%, N=432) of the alumni credit the ROV competition with influencing their educational or career path "to a great extent" or "somewhat".
- The ROV competition played a role in alumni attaining employment (37%), admittance into educational programs/college/university (36%), internships (30%), awards (21%), and scholarships (21%).

Please see the 2014-2015 evaluation report for further results from the ROV Competition Student Alumni Survey.

NATIONAL STUDENT CLEARINGHOUSE MATCH

Background

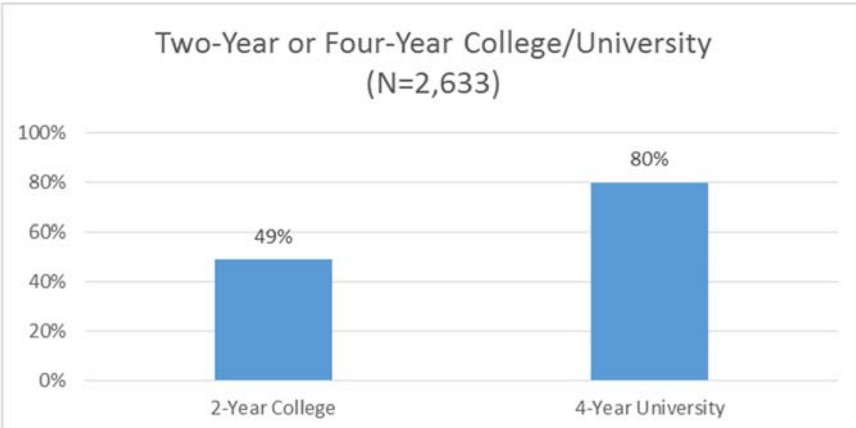
The National Student Clearinghouse (NSC) is a nationwide source of higher education information. Colleges and universities, numbering over 3,400 institutions enrolling over 96% of college students, share their enrollment data with NSC. The NSC database includes over 130 million students. See http://www.studentclearinghouse.org/ for further information about the NSC.

A total of 3,974 former competition participants (aka "alumni") were identified for matching with National Student Clearinghouse (NSC) data in the May of 2017. These alumni have birthdates that indicate that they are at least 18 years of age as of May 2017, and their addresses suggest that they live within the United States. Two-thirds of the alumni (66%, or 2,633) were found in the NSC database.

Two-Year or Four-Year Colleges/Universities

Among the 2,633 alumni who were found in the NSC database, 49% attended a 2-year college, and 80% attended a 4-year college or university. Note that this is not unduplicated. Some alumni attended both.

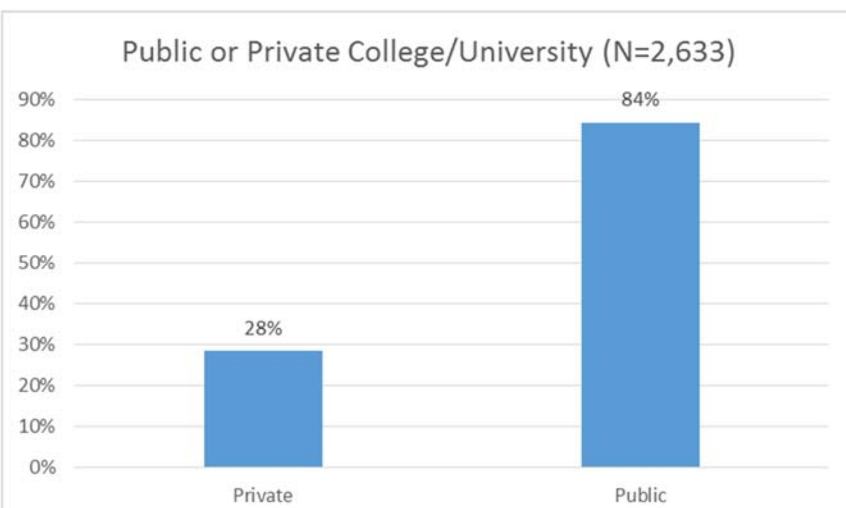
Figure 10: NSC Data: Two-Year or Four-Year University Attendance



Public or Private Colleges/Universities

Among the 2,633 alumni in the NSC database, 84% attended a public college/university, and 28% attended a private college/university. Note that this is not unduplicated, and students may have attended both public and private institutions.

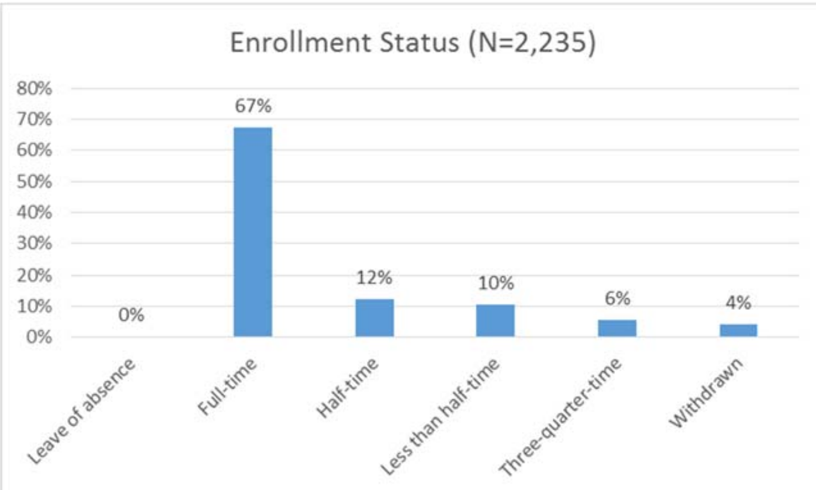
Figure 11: NSC Data: Public or Private College/University Attendance



Enrollment Status

Examining the most recent enrollment status reported per alumni (N=2,235), 67% of the alumni were attending college or university full-time. Six percent (6%) were enrolled three-quarter-time. Twelve percent (12%) were attending half-time, and another 10% attended less than half-time.

Figure 12: NSC Data: Most Recent Enrollment Status

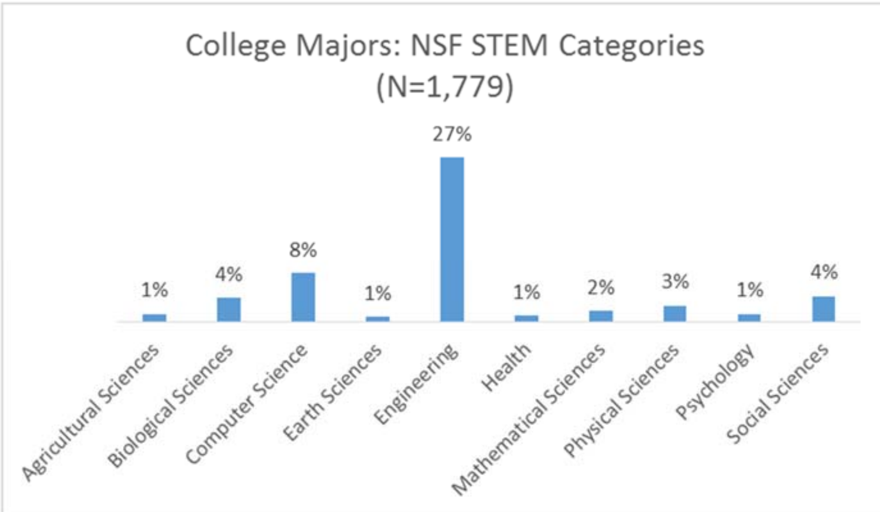


Majors

A total of 1,779 alumni had a major of study in the database, and 1,668 had a Classification of Instructional Programs (CIP) code. The majors were explored with two methods. First, the CIP codes were matched to a list of NSF-designated STEM CIP codes. [3] With this approach, 901 (51%) were designated as having a STEM CIP code. The most common NSF STEM Categories were engineering, with 27% of the alumni, and computer science, with 8%.

[3] See http://aaude.org/documents/public/reference/crosswalk-nsf-cip.xls for CIP code list.

Figure 13: NSC Data: College Majors: NSF STEM Categories

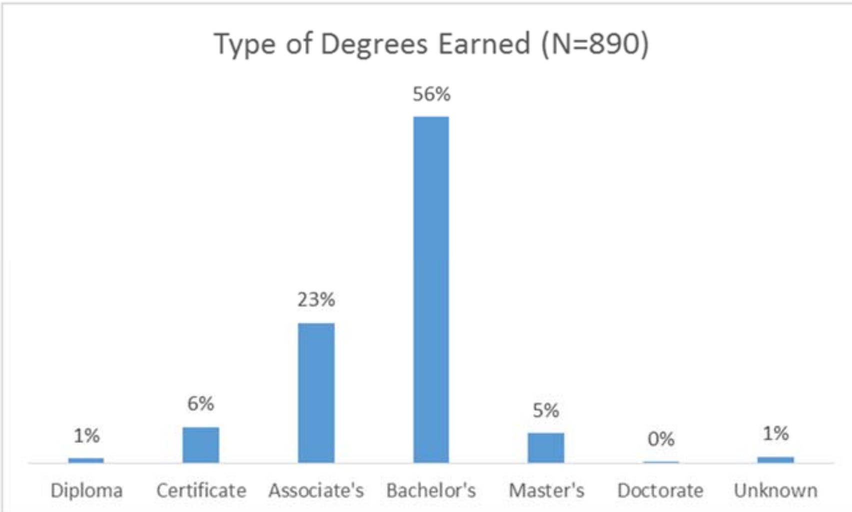


The downside to using the NSF STEM CIP Codes is that certificates and applied associate of science degrees are not included in this list, and these types of completions are common among two-year community and technical colleges, which are a particular area of focus among ATE Centers. Therefore, a second analysis method was employed: the evaluator hand-coded the majors as STEM or non-STEM. With this approach, 66% of the 1,798 competition alumni were enrolled in a STEM major.

Degrees

A total of 708 alumni earned 975 degrees and certificates. Within the degree names, it was possible to determine the type of degree for 890 degrees. The most common type of degree earned by the alumni was Bachelor's degree (56%), followed by Associate's degree (23%). Six percent (6%) of the degrees were certificates, and another 5% were Master's degrees. Two individuals earned doctorates. (See Figure 14)

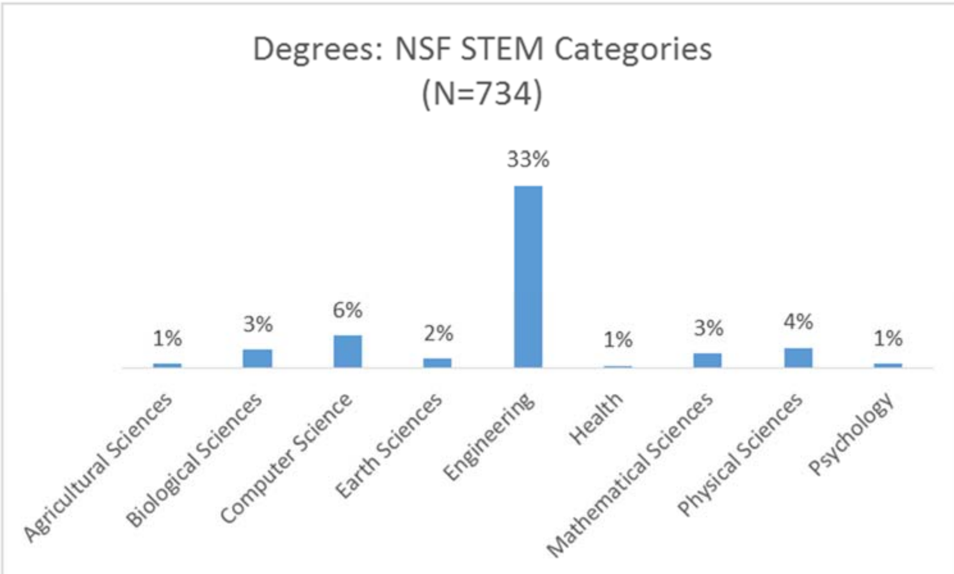
Figure 14: NSC Data: Type of Degrees Earned



The same two analysis methods were used for the degrees as the majors. First, the CIP codes were matched to a list of NSF-designated STEM CIP codes. There were 734 degrees with CIP codes. With this approach, 416 (57%) were designated as having a STEM CIP code. The most common NSF STEM Categories were engineering, accounting for 33% of the degrees with CIP codes, and computer science, with 6%. In comparison, according to the National Center for Education Statistics, in 2013-2014, only 1% of the Bachelor's degrees were in engineering technologies.[4]

[4] SOURCE: U.S. Department of Education, National Center for Education Statistics, Higher Education General Information Survey (HEGIS), "Degrees and Other Formal Awards Conferred" surveys, 1970-71 through 1985-86; Integrated Postsecondary Education Data System (IPEDS), "Completions Survey" (IPEDS-C:91-99); and IPEDS Fall 2000 through Fall 2014, Completions component. (This table was prepared September 2015.) http://nces.ed.gov/programs/digest/d15/tables/dt15_322.10.asp?current=yes

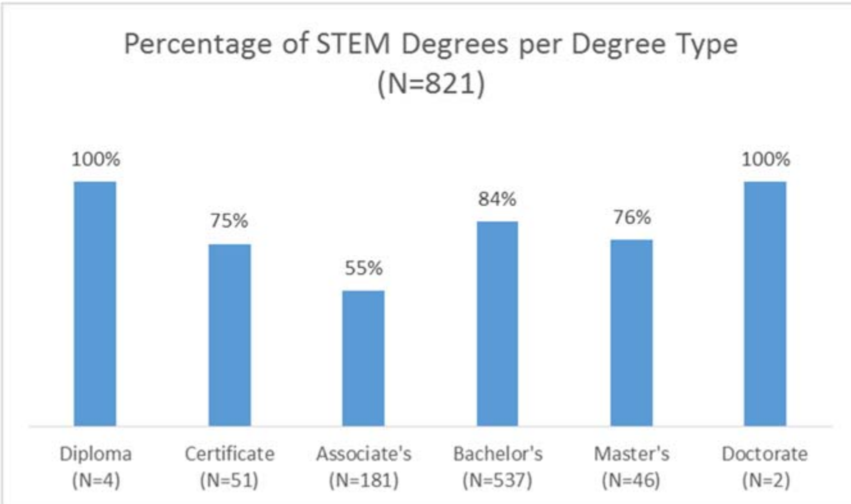
Figure 15: NSC Data: Degrees: NSF STEM Categories



With the hand-coding degrees into STEM and non-STEM, a total of 76% of the degrees (N=828) were in STEM disciplines.

The hand-coded STEM degrees were examined by degree type. (See Figure 16) Compared to other degrees and certificates, it was less common for Associate's Degrees to be in a STEM discipline (55%). Some of this difference is due to the fact that many Associate's degrees were given generic names that did not designate a specific discipline, such as General Studies Associate Degree or Oregon Transfer Associate of Arts.

Figure 16: NSC Data: Percentage of STEM Degrees per Degree Type



Evaluation Question 1.5. What effect did multi-year participation have on the above evaluation questions?

There were several statistically significant differences between the first year and multi-year competition participants. Multi-year participants reported that their participation in the ROV program resulted in higher levels of awareness of and interest in pursuing STEM careers, gains in interest in taking STEM courses, improvements in STEM knowledge and skills, increased 21st Century skills, and the receipt of awards, honors, and new educational and career opportunities.

Specifically, multi-year participants were significantly more likely to report the following:

Table 2: Statistically Significant Differences between First-Year and Multi-Year Participants

	First-Year Participants: Percentage Strongly Agreeing	Multi-Year Participants: Percentage Strongly Agreeing
Increased awareness of STEM careers due to ROV program**	27.1%	39.1%

Increased interest in pursuing a STEM career due to ROV program*	37.6%	43.6%
Learned how to apply STEM knowledge to solving real world problems due to ROV program*	41.6%	51.8%
Learned how to communicate their engineering designs due to ROV program*	37.3%	45.9%
More interested in taking engineering courses due to ROV program*	37.5%	46.3%
More interested in taking science courses due to ROV program**	33.2%	41.7%
More interested in taking math courses due to ROV program**	27.0%	34.9%
Increased skills and knowledge in science due to ROV program**	33.4%	41.8%
Increased skills and knowledge in math due to ROV program**	25.6%	35.9%
Increased leadership skills due to ROV program**	34.8%	48.9%
Increased problem solving skills due to ROV program*	43.5%	50.2%
Increased critical thinking skills due to ROV program**	40.4%	48.1%
Received an award or honor due to ROV program**	23.1%	39.8%
ROV program participation opened educational or career opportunities*	29.5%	36.6%

* $p < 0.05$
** $p < 0.01$

DEMOGRAPHIC BREAKDOWNS OF FIRST-YEAR AND MULTI-YEAR PARTICIPANTS

There were no statistically significant differences between the first-year and multi-year participants in gender (first-year: 68.1% male; multi-year: 69.8% male) or whether the participants had a disability that required accommodations (first-year: 3.0%, multi-year: 1.8 %).

There were statistically significant differences between the first-year and multi-year participants in ethnicity (first-year: 62.7% white; multi-year: 72.5% white). In addition, the multi-year participants were significantly more likely to live in a low socioeconomic status neighborhood (first-year: 51.1%, multi-year: 60.7 %).

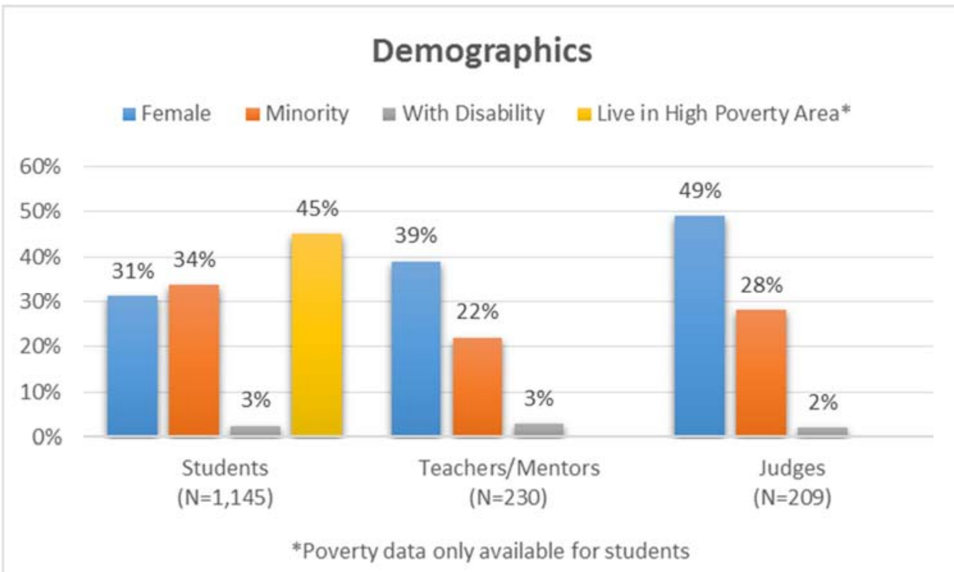
Compared to the multi-year participants, a greater proportion of the first-year participants were in the SCOUT (entry-level) competition class (first-year: 46.2%, multi-year: 26.2 %). The bulk of the multi-year students were in the RANGER (intermediate) class (first-year: 38.0%, multi-year: 55.6%) and NAVIGATOR (beginner-intermediate) classes (first-year: 13.9%, multi-year: 18.2 %). Two percent (1.9 %) of the first-year participants were in the EXPLORER (advanced) competition class. No multi-year participants in the EXPLORER competition class this year. The low proportion of participants in the EXPLORER class is due to the fact that the competition survey dataset did not include the international competition, where most of the EXPLORER teams compete.

Evaluation Question 1.6. Did the robotics activities create the same impacts among underrepresented groups (by gender, ethnicity, socio-economic status, disability) as were found among students who traditionally participate in these types of activities?

BACKGROUND: DEMOGRAPHICS OF STUDENTS, TEACHERS AND INDUSTRY REPRESENTATIVES

According to the demographic data in the surveys, the students (N=1,145) were 31% female; 34% percent were of minority backgrounds[5]; 45% came from high poverty areas[6]; and 3% reported that they had disabilities requiring accommodations. (See Figure 17)

Figure 17: Student, Teacher/Mentor, and Judge Demographics



The project has made efforts to include the participation of teachers, college students, staff, and competition judges (industry professionals) of diverse backgrounds who can serve as role models for the

[5] The sample size of participant surveys from each ethnicity was not large enough to do analysis by individual ethnicity. Instead, all non-white respondents were coded as "minority", and results were analyzed by this "minority status" variable.

[6] High poverty areas were defined as zip codes where the percentage of families with children under age 18 in poverty was higher than the nationwide average of 18.1%. This calculation is based on data from the 2014 American Community Survey at the ZTCA level.

middle school students. Thirty-nine percent (39%) of the teachers (N=230) were female, 22% were of minority backgrounds, and 3% indicated that they had a disability.

Among the judges completing surveys (N=209), 49% were female, 28% were of minority ethnic backgrounds, and 2% marked that they had a disability.

ANALYSIS

As the MATE Center is a longstanding center, the evaluation has improved over time. In the 2009-2010 evaluation report, preliminary results presented the trends by gender and ethnicity only. In 2010-2011, the analysis took a different approach. Rather than simply look at trends, the changes in survey administration methods helped us produce a dataset more suitable for more sophisticated analysis. Thus, we looked for statistically significant differences between the under-represented students and the students who more typically participate in these types of STEM events.

This new analysis begged the question: how should success be defined? In consultation with project managers, the evaluators decided that the measure of successfully engaging under-representative students would be that their results were not statistically different from the other students' results. In other words, the under-represented students made the same gains as the other students.

FINDINGS BY STUDENT DEMOGRAPHICS

The analysis focuses on whether there were statistically significant differences between the groups (gender, ethnicity, socioeconomic status, and disability status) in the following topics:

- Awareness of and interest in STEM careers
- Interest in STEM topics
- STEM skills and knowledge

Statistically significant differences existed between the groups in the following measures:

Awareness of and Interest in STEM Careers

- **Gender:** Between the male and female students, there were no statistically significant differences in gains in knowledge about STEM careers (male: 34% strongly agreed, female: 29% strongly agreed); however, male students were significantly more likely to cite the competition as influencing them towards a STEM career (male: 43% strongly agreed; female: 34% strongly agreed).
- **Ethnicity:** There were no significant differences in gains in knowledge of or interest in STEM careers by ethnicity.

- **Socioeconomic status:** Students living in high poverty areas were less likely to indicate that the competition led to gains in knowledge about STEM careers (low poverty: 37%; high poverty: 30%) or interest in STEM careers (low poverty: 46%; high poverty: 38%).
- **Disability status:** There were no significant differences in gains in knowledge of or interest in STEM careers by disability.

Interest in STEM Topics

- **Gender:** Male students were more likely to report increased interest in computer science (male: 38% strongly agreed, female: 29% strongly agreed) and engineering (male: 46%, female: 31%) courses. There were no significant differences between the genders in the interest in taking science courses, math, or hands-on classes or clubs like robotics, electronics, or shop.
- **Ethnicity:** There were no significant differences in interest in STEM topics by ethnicity.
- **Socioeconomic status:** Students in low poverty areas were more likely to indicate that the competition increased their desire to take science courses (low poverty: 40%; high poverty: 35%), math courses (low poverty: 34%; high poverty: 28%), and hands-on classes or club activities (low poverty: 61%; high poverty: 57%). There were no significant differences between the responses of the students living in high and low poverty areas in the increased interest in other STEM topics.
- **Disability status:** Students with disabilities were less likely to report increased interest in computer science courses (students with disabilities: 26%, students without disabilities: 35%). There were no significant differences between the responses of the students with and without disabilities in regards to interest in other STEM courses.

STEM Skills and Knowledge

- **Gender:** Male students were more likely than female students to report increased skills and knowledge in engineering (male: 52% strongly agreed, female: 46% strongly agreed). There were no significant differences by gender in gains in knowledge in science, math, technology, or knowledge about the competition theme.
- **Ethnicity:** White students were more likely to strongly agree that due to the ROV competition, they increased their skills and knowledge in engineering (white: 51%; minority: 47%). There were no significant differences by ethnicity in the increased skills in other STEM topics.
- **Socioeconomic status:** Students living in a low poverty area were more likely to report gains in skills and knowledge in engineering (low poverty: 55%, high poverty: 47%) and technology (low poverty: 57%, high poverty: 51%). There were no other significant differences in the STEM skills and knowledge according to socioeconomic status.

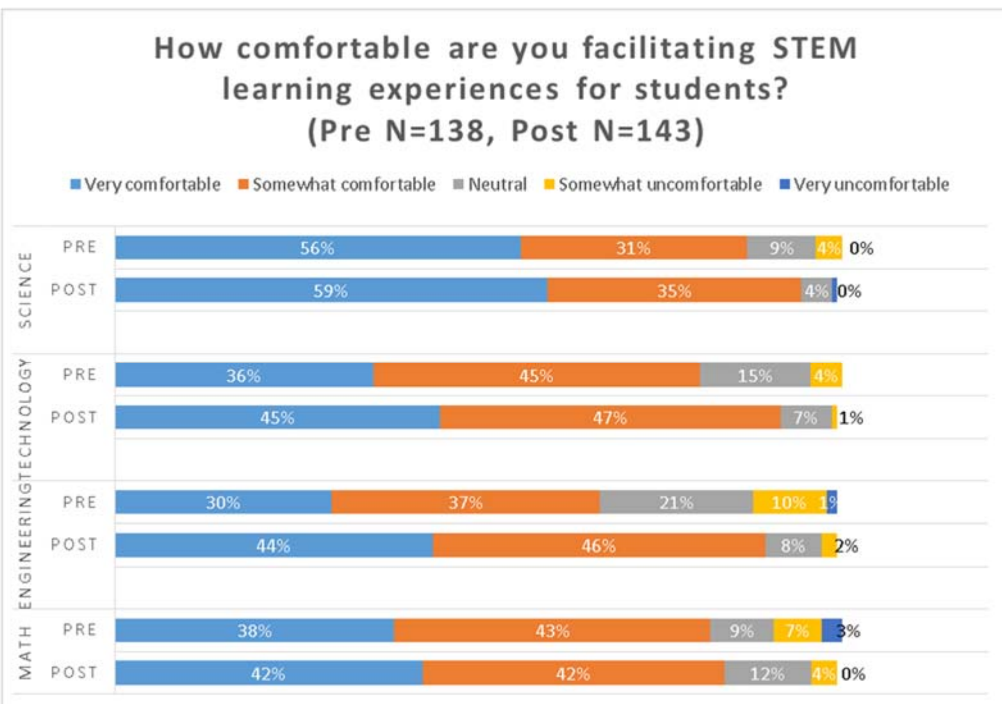
- **Disability status:** There were no significant differences in increased skills in any STEM topics and knowledge by disability status.

Project Goal 2: Provide teachers with professional development, instructional resources, and mentors to support and sustain the delivery of STEM career information and learning experiences.

Evaluation Question 2.1. Are teachers more confident delivering STEM learning experiences? Delivering career information and outlining career pathways?

Pre and post workshop surveys, post competition surveys, and Summer Institute feedback surveys demonstrate that the participants gained confidence facilitating STEM learning experiences through the training and support provided by MATE. The percentage of respondents who rated themselves as "very comfortable" facilitating STEM learning experiences for students rose between the pre and post workshop surveys for science (pre: 56%, post: 79%), technology (pre: 36%, post: 52%), engineering (pre: 30%, post: 58%), and math (pre: 38%, post: 54%).

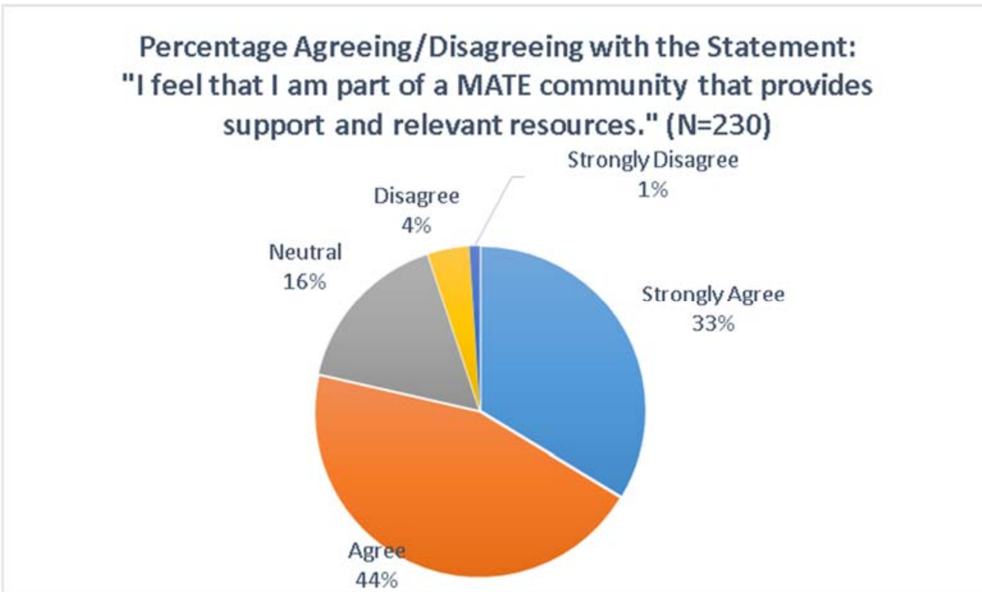
Figure 18: Level of Teacher Confidence Facilitating STEM Learning Experiences: Pre and Post Workshops



Evaluation Question 2.2. Do teachers feel they are a part of a larger MATE community that provides support and relevant, necessary resources?

Among the post-competition surveys, 77% of the teachers (N=230) agreed that they felt they were part of a MATE community that provides support and relevant resources. Sixteen percent (16%) felt neutral about the statement; 4% disagreed, and 1% strongly disagreed.

Figure 19: Percentage Agreeing/Disagreeing that They Feel a Part of a MATE Community



Evaluation Question 2.3. Do teachers incorporate MATE robotics activities/curriculum into courses and afterschool programs? Are the courses and/or curriculum adopted by school districts?

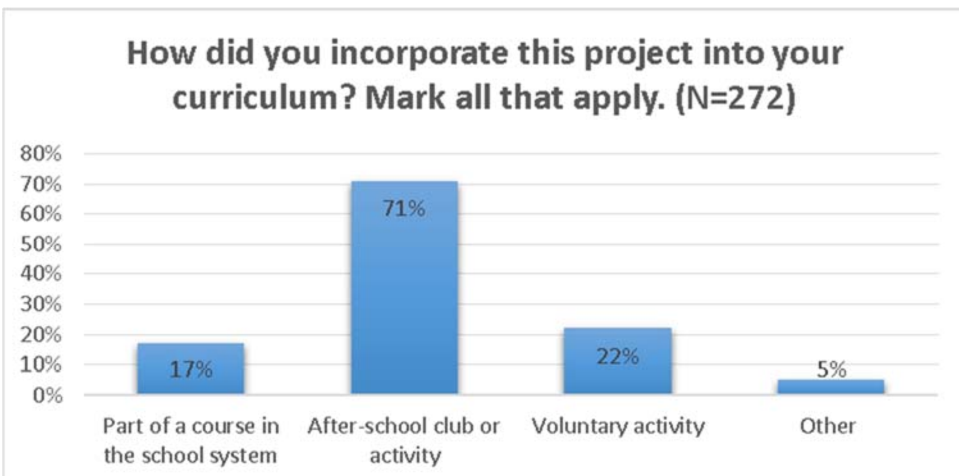
In the post-competition surveys, over three-quarters (76%) of the teachers (N=230) stated that they used MATE materials and resources to incorporate ROV building into their course or club, and over one-third (41%) modified their curriculum and teaching based on MATE resources. Two teachers mentioned that how (s)he incorporated the competition components and rubrics into his or her classroom:

> *Used the ROV as a tool to teach about the science involved in an ROV to my science classes.*

> *We have added 2 more semesters of robotics courses to next academic year that I will teach.*

Seventy-one percent (71%) of the post-competition survey respondents (N=272) incorporated building ROVs into an after-school club. Seventeen percent (17%) built ROVs as part of a course; 22% built ROVs as a voluntary activity; and 5% built ROVs in another venue.

Figure 20: ROV Projects in Courses, After-School Clubs, Voluntary Activities, and Other Activities



In the teacher post-workshop surveys, 24% of the respondents (N=180) were planning to incorporate ROVs into 56 courses in the school system. According to the teachers, these courses fall into the categories of engineering, science, technology, physics, math, and Electronics. The course names include the following, as written by the respondents:

- Marine science
- Physics
- Technology elective
- Specific aptitude program
- Electronics
- Physical science
- Physics - motion, electricity, and magnetism
- Science inspired/underwater robotics
- Computer/science
- Achieve physics
- Mathematics enrichments
- Advanced STEM
- Pre-Engineering
- Autonomous Learning model (ALM)
- Robotics
- Technology design
- Oceanography, Biology, Earth science
- ROV elective
- System design and process
- Marine biology
- Angelfish
- Mechatronics

In the one-year follow-up survey of the week-long, intensive 2016 Pufferfish Summer Institute, all 19 participants who responded to the survey question built ROVs with students. A total of 109 ROVs were built with 1,687 students. Twenty-nine (29) of the ROVs were used in the 2017 MATE ROV competition.

Evaluation Question 2.4. Are teachers able to access classroom mentors as needed? Do the classroom mentors help them successfully incorporate robotics activities into the course? Are the classroom mentors adequately prepared?

In several regions, the regional coordinator matched up college and high school students – in many cases, former ROV competitors themselves – with middle school ROV teams to work with them throughout the competition season. College students also acted as helpers at the workshops. In some cases, the college students received a small stipend (though they stated that they would have done the work without it), and in other cases they received service learning credit, Presidential Volunteer Service Awards, or simply volunteered their time with no recompense. This arrangement worked well for the regional coordinators, college students and middle school students and teachers.

Anecdotal reports suggest that the involvement of college students as mentors can lead to profound experiences for both the college and middle school students. Many sources reported that the middle school students found the college students to be approachable representatives of science. These young adults modeled the paths that the middle school students could take to a STEM career.

For 25% of the post-competition teacher survey respondents (N=230), a classroom/club mentor came to their site to help their teams. Among these teachers, the mentor helped them incorporate robotics into their course or club to "a great extent" for 37% of the respondents, a "moderate extent" for 26% of the respondents, and a "small extent" for 20% of the respondents; 1% stated that the mentors were not helpful at all, and 4% were not sure.

The majority of the respondents (79%) indicated that their mentors were adequately prepared to help them and their students through the ROV design and building process. One percent (1%) marked that the mentors were not adequately prepared, and 4% were unsure.

Project Goal 3: Increase parental involvement in order to support and encourage students to pursue STEM education and careers.

Evaluation Question(s) 3.1. Did the MATE robotics activities lead to an increase in the parents' support of their children's interest in STEM careers?

Ninety-two percent (92%) of the parents surveyed (N=276) agreed or strongly agreed that participation in the ROV program changed how they envisioned their child's future, making it easier to picture their child with a STEM career (67% strongly agreed, 25% agreed). Seven percent (7%) were neutral, one percent disagreed, and one percent strongly disagreed.

Evaluation Question 3.2. Did the enhanced parent online resources lead to an increase in the parents' ability to provide assistance and support for their children's involvement in the MATE robotics activities?

The online Parent Resource Center (http://www.marinetech.org/parent-resource-center) was launched in the spring of 2015. The resource center offers a "welcome" note for to parents, as well as links to information and resources. Resources include links to videos from the international and regional competitions, links to information such as competition timeline and costs, links to student learning objectives for ROV building, and "Getting Started with MATE Underwater Robotics," a section aimed at helping newcomers navigate finding information (e.g. what kit to purchase, what competition class to enter, etc.) on the MATE web site.

Anecdotal evidence indicates that parents found the resources to be helpful, such as comments including the following from 2015:

> *Great overall experience, I have no experience with ROV's but the MATE educational PowerPoints really helped!*
>
> *This was a great first year. The online resources were very helpful.*

Evaluation Question 3.3. Did the regional parent advisory committees provide feedback and advice to help improve the competitions and ensure that the program is inclusive of all participants?

Advisory committees were broadened to invite participation from parents as well as industry representatives, professional organizations (e.g., Marine Technology Society), government agencies (e.g., NOAA) 6-12th grade educators, community college faculty, and university faculty. The committees were implemented at the regional level so the recommendations would be applicable to the local community needs.

In 2016-2017, some regions held an annual advisory meeting (e.g., New England), while others held monthly advisory meetings (e.g., Florida Region), biannual meetings (Gray's Reef), or communicated on an ongoing basis with advisory members outside of formal meetings (e.g., Pennsylvania, Oregon, Great Lakes, and Wisconsin Regions). Advisory groups communicated face-to-face, and via video meetings (e.g., ZOOM), email, and web surveys. Advisory committees provided recommendations in the following areas:

- Timing and content of workshops and regional events
- Funding sources
- Outreach
- Competition logistics

The regional coordinators were responsive to their committees' recommendations.

Broader Impacts

The MATE Center's ITEST activities have been leveraged by regional coordinators and participants in ways that were unanticipated during the writing of the proposal. Thus, they don't fit under any particular evaluation question. Since the evaluation was not set up to monitor these activities, the findings presented here should be considered preliminary. Next year, the evaluation tools will be modified to capture more of this data.

These "broader impacts" fall into three main categories:

1. Leveraging ITEST activities/funding to raise additional funding by regional coordinators, teachers, schools, and student teams;
2. Using ROVs and ROV-based activities outside of the competition by teachers and students;

3. Broader impacts on teachers and institutions: new careers, new classes, deeper relationships with students, improved STEM knowledge, increased motivation and engagement with their discipline, and increased professional development opportunities.

LEVERAGING ITEST ACTIVITIES/FUNDING

Faculty who led ROV teams and/or attended the Summer Institute reported that they have applied for and won funding from grants and school boards and have received equipment donations from local industry. Additionally, ROV competition regions outside of the United States have leveraged news of the ITEST grant to raise additional funds.

USING ROVS OUTSIDE THE COMPETITION

Many faculty have reported using ROVs or ROV-based activities outside of the competition, incorporating these tools and topics into their classes or clubs in order to bring science to life.

BROADER IMPACTS ON TEACHERS AND INSTITUTIONS

Teachers report a broad variety of positive results from their participation in the ROV competition and professional development, including the following:

- New careers
- New classes
- Deepened relationships with students
- New collaborations with industry, research orgs, and other educational institutions
- Improved STEM knowledge
- Increased motivation and engagement with their discipline
- Increased professional development opportunities (in addition to that offered by MATE)

CONCLUSIONS

Overall, the MATE Center successfully implemented the 2016-2017 year of ITEST grant activities. The 2016 MATE ROV Competition was held, with ITEST funding helping to support 16 of the 20 US-based regional events. A total of 413 regional workshops were held for teachers and students, and 20 teachers attended the intensive Summer Institute professional development. The focus on a formal curriculum was changed to development of a suite of online instructional materials, which were disseminated, including videos, PowerPoints, ROV kits, and an online course.

The evaluation's student follow-up efforts made progress in several areas. The alumni survey was completed in 2015. The National Student Clearinghouse data match was completed late in the summer of 2016 and again in the spring of 2017. The collaboration with Dr. Li from the University of Washington continues to produce improvements to the survey instruments and competition scoring rubrics and manuals. The Washington State Follow-up slowed due to communication issues with ERDC; though we hope that the study will move forward in the next year.

Overall, evaluation results continue to show strong positive outcomes for students and teachers. Involvement in the ROV competition generated greater awareness of and interest in pursuing STEM careers, increased interest in studying STEM topics, improved STEM knowledge and skills, and increased 21st Century skills, including teamwork, critical thinking, leadership, problem solving and self-confidence. Participating in the ROV competition helped students learn how to apply STEM skills to real world problems. Students also learned how to communicate their engineering process and design to a wide audience.

Parents were passionate supporters of their children's involvement in the program, with comments such as, "This opportunity was the best part of his high school years." Educational research has stressed the importance of family support in a student's choice to follow a STEM career path. Evaluation results show that the ROV program impacted the participants' parents as well, making it easier for them to picture their child in a STEM career.

ROV competition student alumni survey results and National Student Clearinghouse match analysis suggests that the majority of ROV competition participants go on to study STEM topics, earn STEM degrees, and work in STEM fields. In fact, roughly one in five former participants have worked in a job related to ROVs or other underwater technologies. The majority of ROV competition alumni credit the ROV competition with influencing their educational and career paths, including playing a role in attaining internships, scholarships, admittance to educational programs, and employment.

These findings suggest that the MATE ROV Competition is effective in increasing the STEM workforce, especially related to underwater technologies.